

Dissertation presented to the Instituto Tecnológico de Aeronáutica, in partial fulfillment of the requirements for the degree of Master of Science in the Graduate Program of Electronics Engineering, Field of (Area).

**Rodrigo Aragão Santos**

**SPATIAL TEXTURES SAR DATA FOR FOREST  
MAPPING AND TARGET CHANGING DETECTION**

Dissertation approved in its final version by signatories below:

Prof<sup>a</sup>. Dr<sup>a</sup>. Marcelo Pinho

Advisor

Prof<sup>a</sup>. Dr<sup>a</sup>. Renato Machado

Co-advisor

Prof. Dr. Marcelo Gomes da Silva Bruno  
Dean for Graduate Education and Research

Campo Montenegro  
São José dos Campos, SP - Brazil  
2021

A todos que tornaram um pouco mais  
fácil a jornada até aqui

# Acknowledgments

*"There's some good in this world, and it's worth fighting for."*  
— J.R.R. TOLKIEN

# Resumo

# **Abstract**

# List of Figures

FIGURE 2.1 – Illustration of SAR geometry. $r_0$ is the shortest distance to target, $\theta_a$ stands for the azimuth beam width and $v$ for the satellite speed. (MOREIRA <i>et al.</i> , 2013) . . . . .	18
FIGURE 2.2 – Temporal separation of pulses for target resolution. $\tau$ is the pulse temporal width . . . . .	19
FIGURE 2.3 – Summary of SAR compression in range and azimuth directions (MOREIRA <i>et al.</i> , 2013) . . . . .	20
FIGURE 2.4 – Normalization areas for SAR backscatter (SMALL, 2011) . . . . .	21
FIGURE 2.5 – Reference Geometry for Across-Track interferometry. (RIZZOLI, 2018) . . . . .	23
FIGURE 2.6 – Reference Geometry for flat earth component. (RIZZOLI, 2018) . .	25
FIGURE 2.7 – (a) Amplitude of image of SAR image acquired by ERS-1/2 at C-Band with a baseline of 133m. (b)Interferogram before the flat earth component removal. (c)Interferogram after the flat earth component removal. The color on the images represent the value of the interferometric phase of the pixel, such that pixels connected with the same color should have the same topographical height (RIZZOLI, 2018) .	26
FIGURE 2.8 – Explanation about phase unwrapping. (a) Topographical phase before the phase unwrapping. (b) Topographical phase after the unwrapping of the phase. The images are obtained from a repeat pass acquisition SIR-C-X SAR over Mount Etna, Italy. (RIZZOLI, 2018)	27
FIGURE 2.9 – Bias of the MLE coherence estimator (BAMLER; HARTL, 1998). .	28
FIGURE 3.1 – GlobCover classification map . . . . .	32
FIGURE 3.2 – Mean and Standard deviation of the coherence estimate as number of looks (filter size) for different classes . . . . .	33

FIGURE 4.1 – Fishbone pattern of deforestation . . . . .	36
FIGURE 4.2 – TanDEM-X dataset available in Rondonia Area . . . . .	37
FIGURE 4.3 – SENTINEL-1 dataset available in Rondonia Area . . . . .	38
FIGURE 5.1 – Classification Boundaries for Different Algorithms . . . . .	42
FIGURE 6.1 – Figura . . . . .	53
FIGURE 6.3 – Reference Map provided by DLR. . . . .	58
FIGURE 7.1 – Volumetric Correlation and Classification . . . . .	73
FIGURE 7.2 – $\gamma$ acquisition . . . . .	76
FIGURE 7.3 – $\sigma^0$ acquisition . . . . .	76
FIGURE 7.4 – Histograms of Coherence Textures . . . . .	78
FIGURE 7.5 – Histograms of $\sigma^0$ Textures . . . . .	80
FIGURE 7.6 – Reference Map provided by DLR . . . . .	81
FIGURE 7.7 – Classification Results without textures . . . . .	82
FIGURE 7.8 – Classification Results with textures . . . . .	83
FIGURE 7.9 – Google Earth image of analysed area . . . . .	84
FIGURE 7.10 –Accuracy versus number of estimators for different tree depths . . .	85
FIGURE 7.11 –Accuracy versus minimum number of samples for leaf parameter . .	86
FIGURE 8.1 – The CARABAS-II VHF SAR mounted in front of a Sabreliner airplane	90
FIGURE 8.2 – Military vehicles used as target. Left:TGB11. Middle:TGB30. Right:TGB40. This picture also depicts the vegetation characteristics of the test area	92
FIGURE 8.3 – Examples of CARABAS-II images where (a) is mission 2, (b) is mission 3, (c) is mission 4 and (d) is mission 5 . . . . .	93
FIGURE 9.1 – Example of a SAR image acquired by the CARABAS-II system. The red circle shows the position of the cars hidden under a canopy of trees. . . . .	97
FIGURE 9.2 – Entropy image. The red circle shows the position of the vehicles. . .	99
FIGURE 9.3 – Variance image. The red circle shows the position of the vehicles . .	101
FIGURE 9.4 – UNET architecture . . . . .	102

FIGURE 9.5 – Inference Phase for the corresponding DI . . . . .	102
FIGURE 9.6 – Diagram of the proposed algorithm. . . . .	103

# List of Abbreviations and Acronyms

SAR	Radar de abertura sintética (Synthetic Aperture Radar)
ML	Machine Learning
DLR	Centro Aeroespacial Alemão (Deutsche Zentrum für Luft- und Raumfahrt)
InSAR	SAR Interferometry
PDF	Função Distribuição de Probabilidade (Probability Distribution Function)

# Contents

1	INTRODUCTION . . . . .	15
2	THEORETICAL BACKGROUND . . . . .	17
2.1	Remote sensing and SAR imaging . . . . .	17
2.1.1	SAR Image Geocoding . . . . .	21
2.2	SAR Interferometry . . . . .	21
2.2.1	Height acquisition and interferogram equations . . . . .	22
2.2.2	The Complex Interferometric Coherence . . . . .	27
2.2.3	Co-registration . . . . .	29
2.2.4	Decorrelation in Vegetated Areas . . . . .	29
3	CLASSIFICATION MAPS GENERATED WITH SAR IMAGES . . . . .	31
4	MATERIAL . . . . .	35
4.1	Test Area: The Rondonia State . . . . .	35
4.2	TanDEM-X datasets . . . . .	36
4.3	Sentinel-1 datasets . . . . .	37
5	TEXTURES METHODS FOR IMAGE ANALYSIS . . . . .	40
5.1	Machine Learning and the problems for classification . . . . .	41
5.2	Textures Methods . . . . .	43
5.3	The GLCM Method . . . . .	43
5.4	The Laws Textures . . . . .	46
5.5	The Sum and Difference Histograms Texture . . . . .	47
5.5.1	The Single Value Decomposition . . . . .	47

---

5.5.2	Approximation of Probability Density Functions . . . . .	49
5.5.3	Textures Features . . . . .	50
<b>6</b>	<b>ANALYSIS OF TEXTURE RESULTS . . . . .</b>	<b>52</b>
<b>6.1</b>	<b>The GLCM Textures . . . . .</b>	<b>52</b>
6.1.1	The Displacement Vector . . . . .	52
6.1.2	The GLCM textures results . . . . .	55
<b>6.2</b>	<b>The Laws Textures results . . . . .</b>	<b>61</b>
<b>6.3</b>	<b>Sum And Difference Histograms Textures . . . . .</b>	<b>66</b>
<b>7</b>	<b>CLASSIFICATION RESULTS . . . . .</b>	<b>72</b>
<b>7.1</b>	<b>TANDEM-X Data Analysis . . . . .</b>	<b>72</b>
<b>7.2</b>	<b>Sentinel1 Data Analysis . . . . .</b>	<b>75</b>
<b>7.3</b>	<b>Performance of Random forest and parameters tuning . . . . .</b>	<b>84</b>
<b>8</b>	<b>THE CARABAS SYSTEM AND CHALLENGE PROBLEM . . . . .</b>	<b>88</b>
<b>8.1</b>	<b>Summary . . . . .</b>	<b>88</b>
<b>8.2</b>	<b>Wavelength Resolution SAR Systems . . . . .</b>	<b>88</b>
<b>8.3</b>	<b>The CARABAS-II Challenge Problem . . . . .</b>	<b>89</b>
<b>8.4</b>	<b>The CARABAS-II System . . . . .</b>	<b>90</b>
<b>8.5</b>	<b>The CARABAS-II Dataset . . . . .</b>	<b>91</b>
<b>8.6</b>	<b>Traditional Change Detection in Wavelength-Resolution SAR Images . . . . .</b>	<b>94</b>
<b>9</b>	<b>TEXTURES APPLICATIONS FOR CHANGE DETECTION ALGORITHMS . . . . .</b>	<b>95</b>
<b>9.1</b>	<b>Summary . . . . .</b>	<b>95</b>
<b>9.2</b>	<b>Introduction . . . . .</b>	<b>95</b>
<b>9.3</b>	<b>Relevant Textures for the Change Detection Algorithm . . . . .</b>	<b>97</b>
<b>9.4</b>	<b>UNET Based Algorithm for Target Change Detection . . . . .</b>	<b>100</b>
<b>9.5</b>	<b>Results . . . . .</b>	<b>104</b>
<b>9.6</b>	<b>Conclusion . . . . .</b>	<b>105</b>
	<b>BIBLIOGRAPHY . . . . .</b>	<b>106</b>

# 1 Introduction

Land Cover classification is a fundamental research topic with applications geography, ecology, geology, forestry, land policy and planning etc... With that in mind the focus of this work is to introduce the reader to the usage of Synthetic Aperture Radars (SAR) and its applications to Remote Sensing and Land Cover Classification.

This work will focus on the usage of SAR data to the detection of forest and non forest areas and change target detection. Since Forest preservation is something crucial for the environment preservation nowadays it is very important to have technology that can perceive fast changes in the scenario of forest, specially if these changes are due to illegal deforestation. Change Target detection methods are also very useful to identify rapid changes in a scene and it is of great importance in remote sensing, monitoring environmental changes and land use.

This work will provide useful information on the current state of the art methods of creating Landcover maps from SAR data and change target detection using SAR data. It will be also be shown how the textural information can improve these methods.

The work presented here is part of a work that was developed at the German Aerospace Center (DLR) during the year of 2019. The rest of the work was done as a Master Student at ITA during the years of 2020 and 2021.

The work developed at DLR was done under the supervision of Paola Rizzoli and Andrea Pulella during the time I was a exchange student there (from february 2019 until november 2019). After that period the work developed at DLR was used for this master thesis (which was written under the supervision of ITA Professor Marcelo Pinho).

# PART I

## 2 Theoretical Background

### 2.1 Remote sensing and SAR imaging

Remote Sensing (RS) is the science of obtaining and interpreting information from a distance using sensors that are not in physical contact with the object being observed. Through remote sensing it is possible to study several processes that occur on Earth. For example, remote sensing has applications in geology, oceanography, meteorology, forest monitoring and even monitoring urban areas subsidence. Physically, remote sensing works by transmitting and receiving electromagnetic waves, and by studying the reflected waves, it is possible to extract parameters about the process of interest. There are two classes of remote sensing: Passive RS and Active RS. Passive RS works by receiving natural radiation emitted or reflected by Earth, while active RS uses sensors that produce their own electromagnetic radiation (e.g.,RADAR).

The Synthetic Aperture Radar (SAR) is an active RS sensor that has several applications in the modern world and several advantages. Two major advantages is that it is possible to monitor the Earth 24 hours per day and it has little dependence of weather conditions while optical sensors can only be used at daylight and cannot be used for monitoring if there are clouds in the area at the time of the study.

A SAR is a side looking radar used in remote sensing for imaging purposes. It consists of an antenna mounted on a spaceborne or airborne platform, that transmits sequentially electromagnetic waves to the Earth's surface. The backscattered echoes are then collected by the same antenna a fraction of a second after the transmission. In a SAR the time of transmission of the pulse and the time of reception of the pulse are related to different antenna positions due to the platform movement. By coherent combination of the received echoes it is possible to synthesize a virtual aperture that is much greater than the antenna length. This basic principle of a SAR is what gives origin to the name "synthetic aperture" and makes possible for the SAR to be an imaging radar. The image obtained is just a reflectivity scene of the area, that due to the coherent property of the SAR, allows to record amplitude and phase of the received signal. For a more detailed explanation regarding SAR it is recommended (CUMMING; WONG, 2005).

Figure 1.1 demonstrates the basic geometry for image acquisition of a SAR. The flight direction is called the along-track dimension and is also called the azimuth direction. The  $x$  axis is called the across-track direction or the ground range direction. The slant range (also called range) direction is the direction between the sensor and the target point on the ground.

The sensor is an antenna with length  $L_a$  along the azimuth direction.

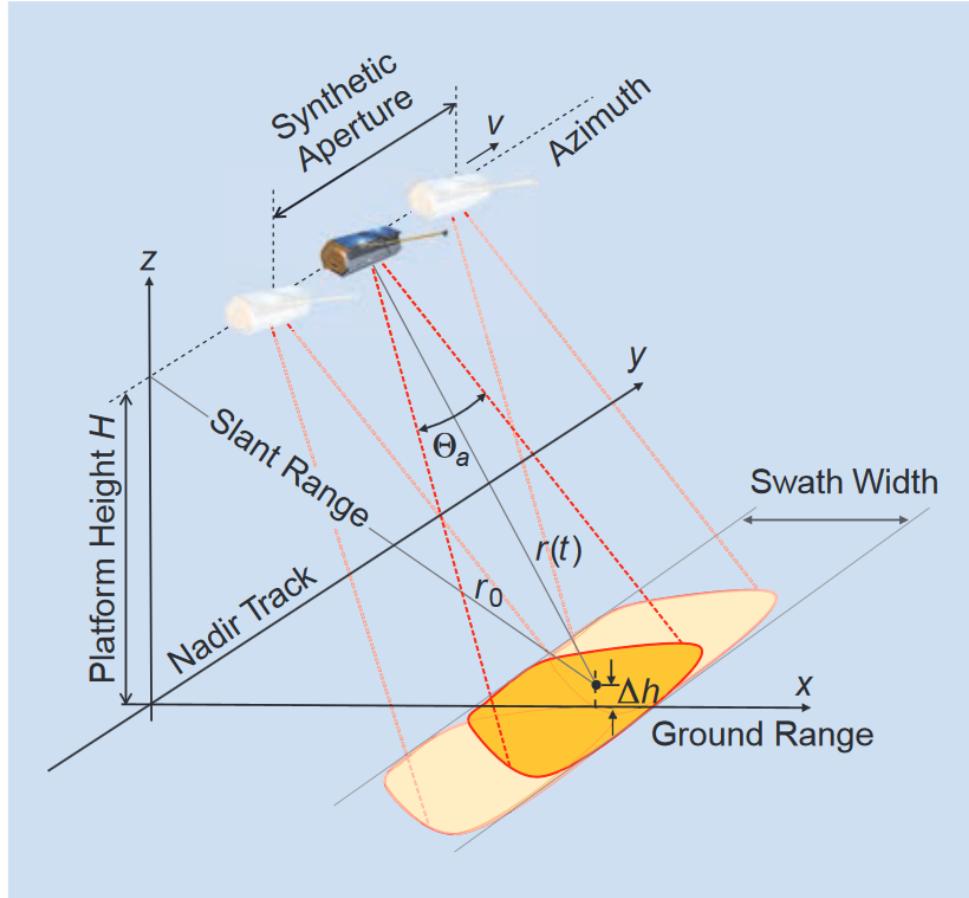


FIGURE 2.1 – Illustration of SAR geometry.  $r_0$  is the shortest distance to target,  $\theta_a$  stands for the azimuth beam width and  $v$  for the satellite speed. (MOREIRA *et al.*, 2013)

The resolution of an image in a given direction is the minimal distance between two points so it is still possible to identify the two different points on the image (also called to resolve the two different targets). In SAR there are resolution on range direction and azimuth direction, which is the minimal distance to separate two targets on range direction and azimuth direction respectively.

For resolving two separate targets, it is normally considered that the spatial separation of the pulses should be higher than half of the Bandwidth of the pulse (this is visually explained in Figure 2.2).

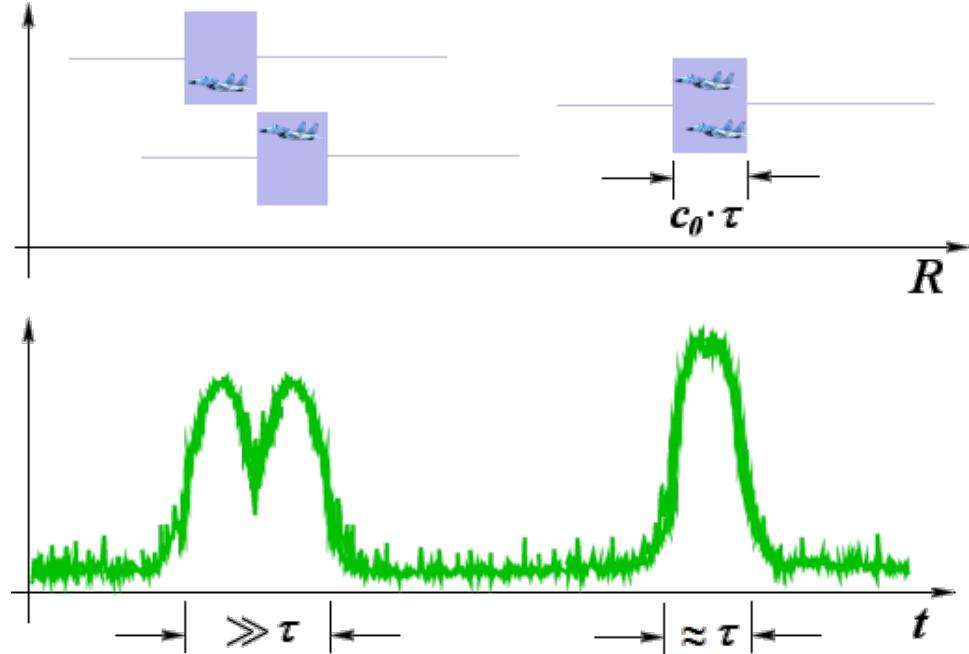


FIGURE 2.2 – Temporal separation of pulses for target resolution.  $\tau$  is the pulse temporal width

From Figure 2.2, it is possible to understand why it was adopted that the temporal separation should be greater than half  $\tau$ . If it is greater it's reasonable to assume that it is possible to separate the targets, but if it is less than that then it's not clear the target separation.

For SAR systems, the resolution in azimuth is equal to half the antenna length. And the resolution in range direction is given by:

$$\rho_{rg} = \frac{c\tau_{rg}}{2} = \frac{c}{2B_{rg}} \quad (2.1)$$

where  $\rho_{rg}$  is the range resolution,  $c$  is the speed of the light,  $\tau_{rg}$  is the temporal resolution and  $B_{rg}$  is the bandwidth.

After acquiring the image of the area, it is necessary to filter the image, so it is clearer for interpretation and information extraction. This is done by using two matched filters (to maximize the signal-to-noise ratio of the image), one in the azimuth direction and the other in the range direction. This process of applying these two filters is called range compression and azimuth compression.

Figure 2.3 gives a summary of how the azimuth and range compression works. The focusing process is done by making convolutions (usually carried out in frequency domain) with the reference function for the received signal in the range direction and in the azimuth direction.

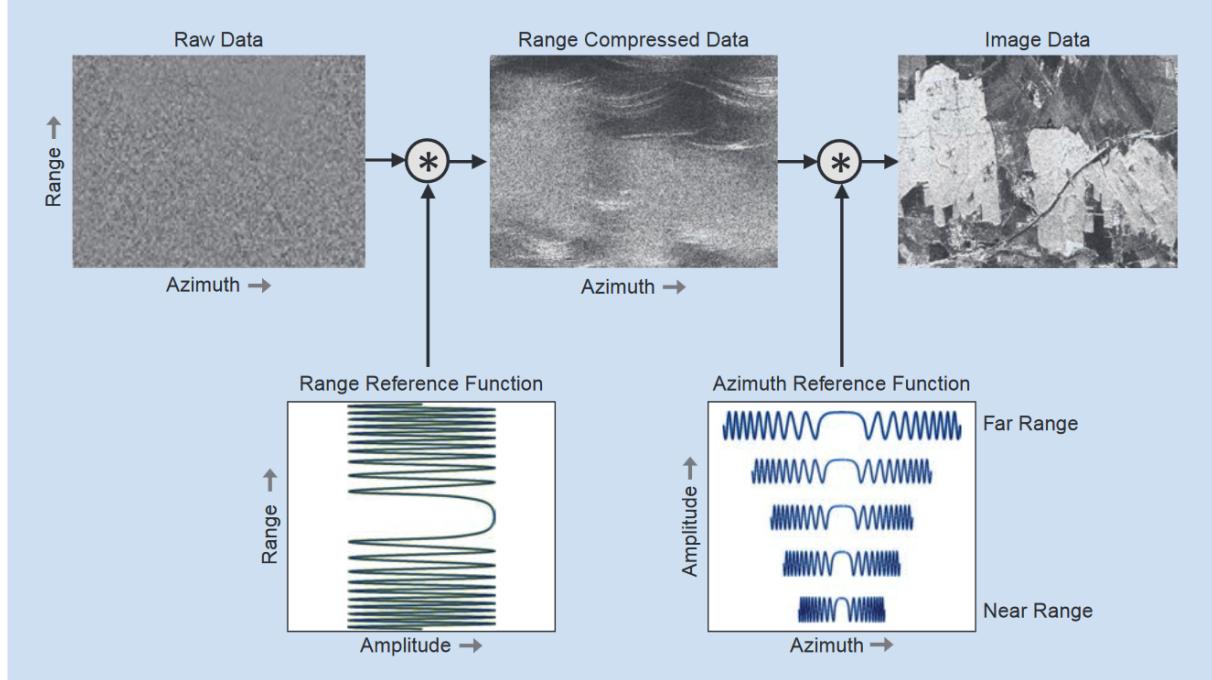


FIGURE 2.3 – Summary of SAR compression in range and azimuth directions (MOREIRA *et al.*, 2013)

After the compression process it is finally obtained a complex matrix called Single Look Complex (SLC) matrix. The amplitude of each element of this SLC matrix represents the power reflected to the antenna. After the compression, it is possible to obtain the backscatter coefficient by multiplying the absolute value of the SLC by a calibration constant obtained experimentally. The backscatter value acquisition is important because accurate backscatter values enable more accurate results in applications like deforestation monitoring, land-cover classification and delineation of wet snow covered area (SMALL, 2011). It is also important to mention that the local terrain has an important role in the determination of the backscatter value, and if a Digital Elevation Model (DEM) of the area is not available then the final result might be compromised.

According to (SMALL, 2011) the radar backscatter  $\beta$  is expressed as the ratio between scattered power  $P_s$  and incident power  $P_i$  at ground level:  $\beta = \frac{P_s}{P_i}$ . When one chooses a reference area  $A_\beta$  in the slant range plane, it is possible to define the radar brightness or  $\beta^0$  backscatter according to (RANEY *et al.*, 1994) as  $\beta^0 = \frac{\beta}{A_\beta}$ .

If the reference area is the ground area ( $A_\sigma$ ), then the result is the *sigma naught* and it is defined as  $\sigma^0 = \frac{\beta^0 \cdot A_\beta}{A_\sigma} = \beta^0 \cdot \sin(\theta)$ :

Instead, if the reference area is perpendicular to the line of sight  $A_\gamma$ , then the *gamma naught*  $\gamma^0$  is the result:

$$\gamma^0 = \frac{\beta^0 \cdot A_\beta}{A_\theta} = \beta^0 \cdot \tan(\theta) \quad (2.2)$$

A visual description of these normalized backscatter coefficients is shown in Figure 2.4

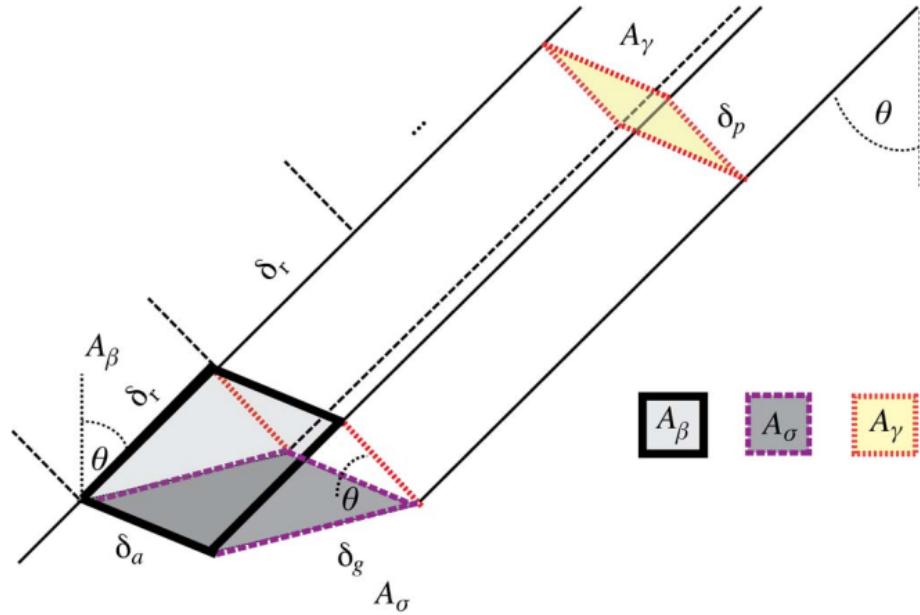


FIGURE 2.4 – Normalization areas for SAR backscatter (SMALL, 2011)

### 2.1.1 SAR Image Geocoding

When an image of a SAR is acquired, there is no direct correspondence on the coordinates  $(x, y, z)$  on the ground reference, since the coordinates of a pixel are in coordinates of range and azimuth. The process of transforming a SAR image from range and azimuth coordinates to local cartesian coordinates on the ground is called geocoding. By using the SAR image acquired, and a reference Digital Elevation Model of the area it is possible to find the points of the SAR image that correspond to the points of the Earth model used as a reference. It is important to make clear that it is absolutely necessary a DEM of the area since a SAR image is a 2D photo of a 3D area, and therefore it is not possible to make the inversion of coordinates since information is lost when making the image acquisition.

## 2.2 SAR Interferometry

The SAR interferometry process uses the fact that two points on the ground with different heights when looked at two distinct positions in the plane perpendicular to the azimuth direction, will give a change in the phase difference of the received signal coming from these points.

By comparing acquisitions obtained at different antennas positions it is possible to extract additional information of a scene. The first time this was done was in 1974 by Graham (GRAHAM, 1974), who obtained a pattern of interference fringes by vectorially adding the signals received from two SAR antennas. One very important remark shown by (GRAHAM, 1974) is that the type of information extracted depends on the implementation of the system. If the images are acquired at different positions but at the same time, then it is possible to extract information about the topography of the area: this is called the across-track SAR interferometry (XTI-SAR). On the other hand, if the images are taken at different times it is possible to get a map of the surface velocities: this technique is called Differential InSAR(D-InSAR).

By taking the product of the first image with the conjugate of the second image it is possible to obtain the interferogram of that area. This interferogram is still not very useful for information extraction since there are more factors that affect the interferogram that must be compensated. The first step important is the flat earth component removal of the interferogram, which is compensate the effect of the phase added to the interferogram due to the horizontal distance of the target. After the flat earth removal what is left is called the interferometric phase. (ROSEN *et al.*, 2000) and (BAMLER; HARTL, 1998) give a summary of how to compensate the flat earth component of the interferogram.

There are two kind of acquisitions that can be used for extracting interferometric information.

- Repeat-pass interferometry: Where the images are acquired during different passes. Normally this means that there are a temporal interference and the quality is inferior, since it is subject to changes in the scene, atmospheric changes.
- Single-pass interferometry: Where the images are acquired at the same time. This means that the quality of the interferogram is superior, since there are less factors that can affect the final product. Since the temporal decorrelation does not affect this type of interferogram, it is chosen as the product for creation of a Digital Elevation Model (DEM) due to its superior quality.

### 2.2.1 Height acquisition and interferogram equations

Consider two SAR images acquired over the same area using sensors S1 and S2. The distance of the sensors is called the baseline  $B$ . Its projections, perpendicular and parallel to the slant range dimension, are the normal baseline  $B_{\perp}$  and the parallel baseline  $B_{\parallel}$ , respectively. They are oriented at an angle  $\beta$  with respect to the local horizontal.

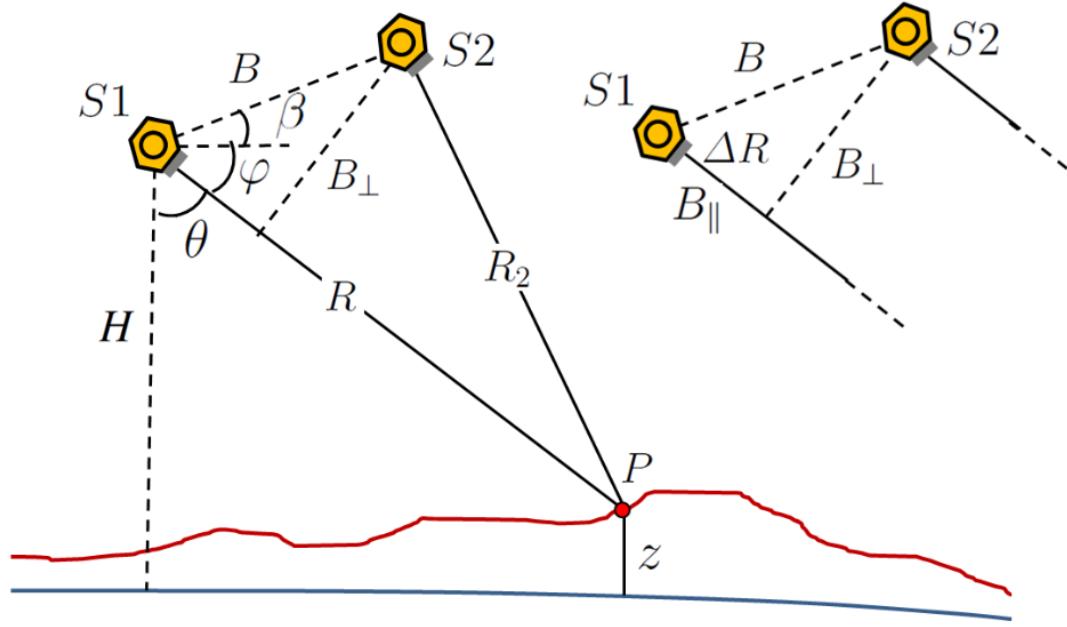


FIGURE 2.5 – Reference Geometry for Across-Track interferometry. (RIZZOLI, 2018)

Since the slant-range distance is normally much greater than the baseline the following approximation holds true:  $|R - R_2| \ll R$ , where  $R$  is the distance from the Sensor S1 to the point on ground,  $R_2$  is the distance from sensor S2 to the point on ground.

The angle  $\varphi$  is given by  $\varphi = \frac{\pi}{2} - \theta$ , and by applying the cosine law on triangle S1 S2 P it yields:

$$(R - \Delta R)^2 = R^2 + B^2 - 2BR \cos(\varphi + \beta) \quad (2.3)$$

Using the parallel-ray aprroximation we get the following result:

$$\Delta R = B \sin(\theta - \beta) \quad (2.4)$$

And the topographical height  $z$  is obteined as:

$$z = H - R \sin(\varphi) \quad (2.5)$$

where  $H$  is the height of S1.

By assuming a monostatic configuration (where both sensor acquire independently an image of the same area) we get finally that the phase of a pixel within the SAR image on sensor S1 is:

$$\Phi_1 = -\frac{4\pi}{\lambda} R + \Phi_{obj,1} \quad (2.6)$$

and the phase of a pixel within the SAR image on sensor S2 is:

$$\Phi_2 = -\frac{4\pi}{\lambda}(R - \Delta R) + \Phi_{obj,2} \quad (2.7)$$

where  $\Phi_{obj,1}$  and  $\Phi_{obj,2}$  are the phases in the images of Sensor S1 and S2. These phases in the images are assumed to be the same.

Therefore, the interferometric phase is obtained as the difference between the phases, such that:

$$\Phi_{int} = \Phi_1 - \Phi_2 = -\frac{4\pi}{\lambda}\Delta R \quad (2.8)$$

and for a bistatic configuration (where only one radar transmits and both receive) it equals to:

$$\Phi_{int} = -\frac{2\pi}{\lambda}\Delta R \quad (2.9)$$

Since the phase difference is half the phase difference for a bistatic configuration since the transmit path does not affect the phase difference.

The interferogram is the two dimensional image of the  $\Phi_{int}$ . It is worth mentioning that before the computation of the interferogram it is needed to make the coregistration of the images, that is merely to project one image on the geometry of the other image. Once the interferogram has been computed the interferometric phase can be decomposed into two parts: the topographical phase component and the flat earth component.

$$\Phi_{int} = \Phi_{topo} + \Phi_{fe} \quad (2.10)$$

where  $\Phi_{topo}$  is the topographic component and  $\Phi_{fe}$  is the flat earth component. The flat earth component accounts for the interferometric phase difference occurring between two scatterers at the same topographical height but at different horizontal positions. The flat earth component is visually explained in the Figure 2.5.

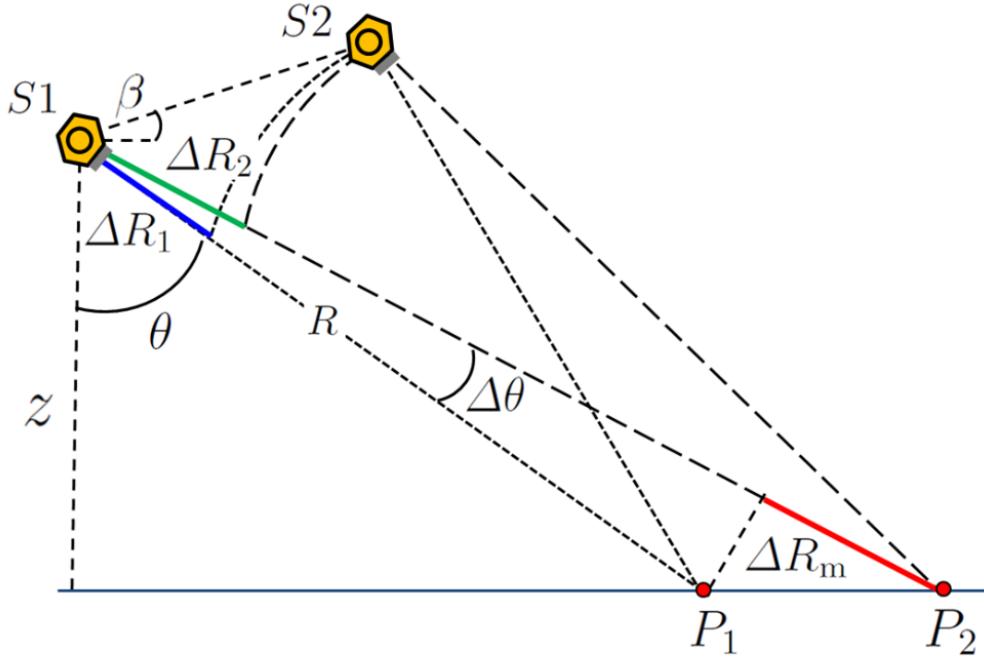


FIGURE 2.6 – Reference Geometry for flat earth component. (RIZZOLI, 2018)

Consider now two points \$P\_1\$ and \$P\_2\$ on the same topographical height. The interferometric phase for each one is given by:

$$\Phi_1 = -\frac{4\pi}{\lambda} \Delta R_1 = -\frac{4\pi}{\lambda} B_\perp \sin(\theta - \beta) \quad (2.11)$$

and for \$P\_2\$

$$\Phi_2 = -\frac{4\pi}{\lambda} \Delta R_2 = -\frac{4\pi}{\lambda} B_\perp \sin(\theta + \Delta\theta - \beta) \quad (2.12)$$

The difference between the two phases is the flat earth component that should be removed. Therefore the flat earth component equals to:

$$\begin{aligned} \Phi_{fe} &= \Phi_1 - \Phi_2 = -\frac{4\pi}{\lambda} [B \sin(\theta + \Delta\theta - \beta) - B \sin(\theta - \beta)] = \\ &\approx \frac{4\pi}{\lambda} B \cos(\theta - \alpha) \Delta\theta \end{aligned} \quad (2.13)$$

where \$R\_m\$ is the slant range difference between \$P\_1\$ and \$P\_2\$.

The results of the extraction of the flat earth component can be visualized in the figure below.

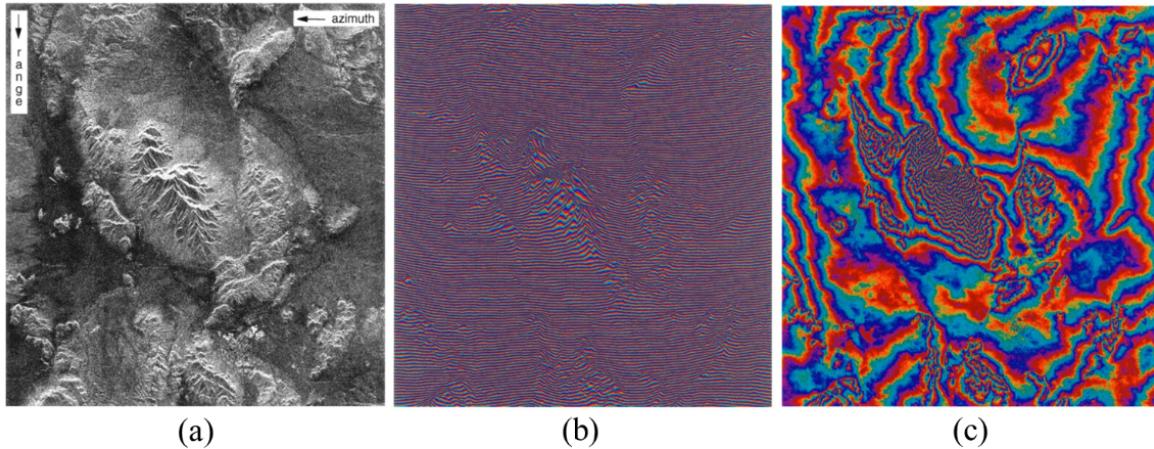


FIGURE 2.7 – (a) Amplitude of image of SAR image acquired by ERS-1/2 at C-Band with a baseline of 133m. (b)Interferogram before the flat earth component removal. (c)Interferogram after the flat earth component removal. The color on the images represent the value of the interferometric phase of the pixel, such that pixels connected with the same color should have the same topographical height (RIZZOLI, 2018)

The altitude difference  $h_{amb}$  corresponding to a phase variation of  $2\pi$  is called height of ambiguity. It is inversely proportional to the perpendicular baseline  $B_{\perp}$  and is defined as:

$$h_{amb} = \frac{\lambda R \sin(\theta)}{n B_{\perp}} \quad (2.14)$$

where  $n = 2$  for monostatic configuration and  $n = 1$  for a bistatic configuration. The idea is that the higher the perpendicular baseline the lower the height of ambiguity and therefore the more accurate the height estimation is.

It is also important to mention the process of phase unwrapping for the processing of the topographical phase component.

Since the topographical phase obtained is modulo  $2\pi$  it is important to reconstruct the correct phase by adding multiples of  $2\pi$  to the phase obtained. This process is called phase unwrapping and it is show in Figure 1.8.

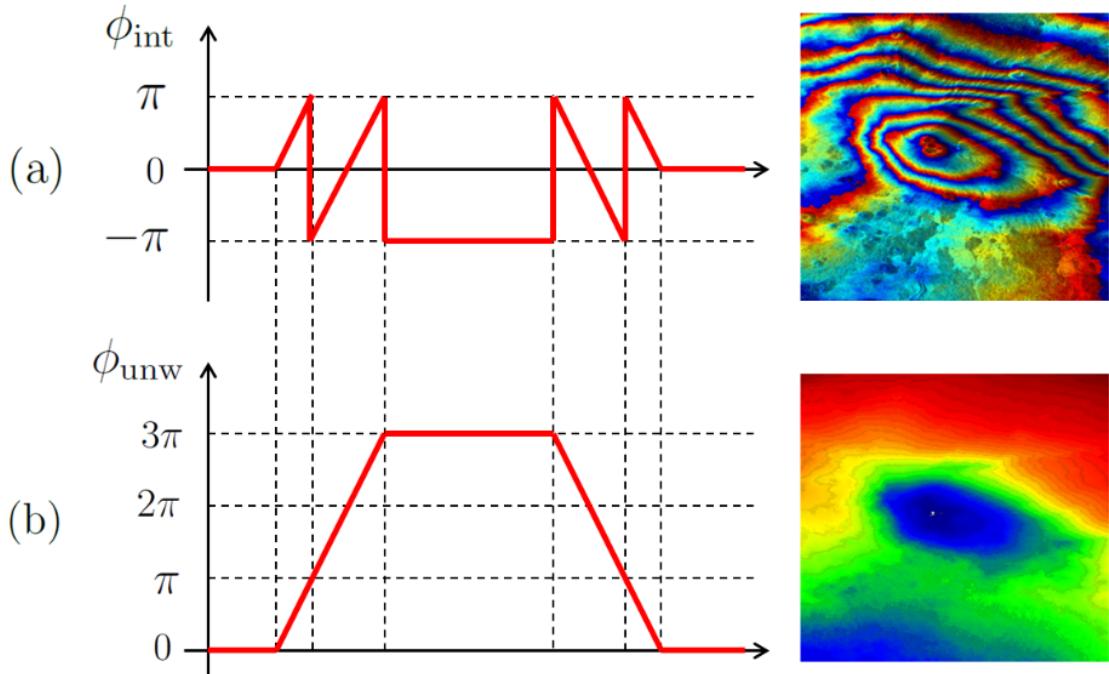


FIGURE 2.8 – Explanation about phase unwrapping. (a) Topographical phase before the phase unwrapping. (b) Topographical phase after the unwrapping of the phase. The images are obtained from a repeat pass acquisition SIR-C-X SAR over Mount Etna, Italy. (RIZZOLI, 2018)

After the flat earth removal and phase unwrapping the interferogram is finally ready for analysis and information extraction.

### 2.2.2 The Complex Interferometric Coherence

The interferometric coherence( $\gamma_t$ ) is a measure of the quality of the interferogram. It is, for each pair of corresponding pixels, the correlation between these two pixels:  $u_1$  and  $u_2$ . It is given by:

$$\rho = \frac{E[u_1 u_2^*]}{\sqrt{E[|u_1|^2] E[|u_2|^2]}} \quad (2.15)$$

where  $E[\cdot]$  is the expectation operator. In practice it is not possible to obtain the correlation between the random variables that model the complex value of the pixels, so in practice this correlation is obtained by a moving average filter along the image. What is obtained is then a estimation of the real value of the coherence, by using this moving average filter what is obtained is the Maximum Likelihood Estimation of the coherence.

This is obtained as:

$$\rho_{MLE}[i, j] = \frac{\sum_{i=1}^N [u_1[i] u_2^*[i]]}{\sqrt{\sum_{i=1}^N |u_1[i]|^2 \sum_{i=1}^N |u_2[i]|^2}} \quad (2.16)$$

Where  $N$  is the number of independent samples used to compute the coherence estimation. The absolute value of the coherence, gives information about the decorrelation level and it is also a measure of the quality of the interferogram. The phase of the coherence, also called interpherometric phase, gives information about the path difference between the two antennas used for the acquisition.

It is important to say that this estimation of the coherence is not unbiased, even though it is asymptotically unbiased (BAMLER; HARTL, 1998). In fact, if the values of the pixels are assumed to follow a gaussian distribution, then it is possible to find the relationship between the estimated coherence, the real value of the coherence and the number of samples used to make the estimation. This function is show in the Figure 2.9 below.

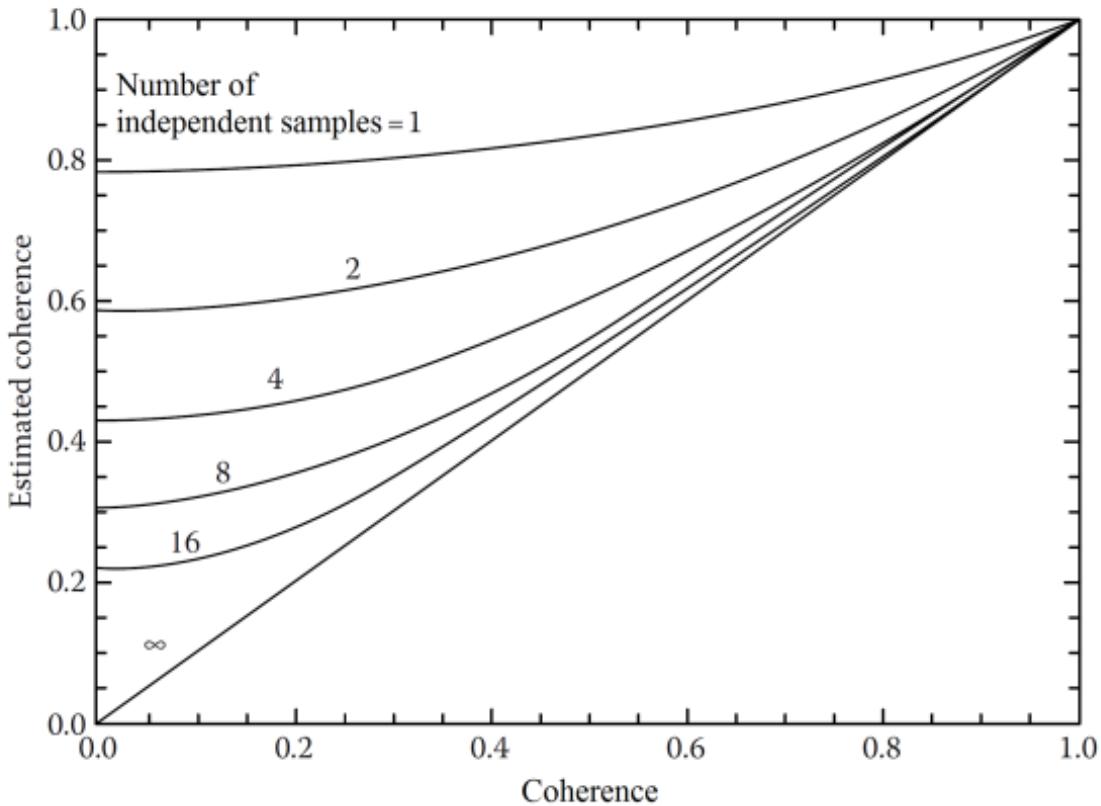


FIGURE 2.9 – Bias of the MLE coherence estimator (BAMLER; HARTL, 1998).

It is important to notice that the estimator is always biased towards higher values than it actually is, and by increasing the number of samples the estimation gets every

time closer to the real value. It is also important to notice that for high values of the coherence the estimator provides a better estimation than it would if the coherence were of low value.

This coherence number is crucial to the textures work that will be developed on the rest of the work, since there is much useful information for classification that can be extracted by looking at this number and by making the textures of the coherence image.

### 2.2.3 Co-registration

Before taking the complex interferometric coherence between the images, a step called co-registration is necessary. After making the focusing step of the two images, it is necessary to transform the slave image to what it would look like if it was taken with the acquisition geometry of the master image. Since the images are taken with different geometries, it produces effects of rototranslations that can scale differences between the images and induce errors in the coherence process. Therefore, this co-registration is a process that correctly aligns the two images to the same geometry and reduces the effects of errors in the calculation of the interferometric coherence(PULELLA, 2017).

### 2.2.4 Decorrelation in Vegetated Areas

According to (MARTONE; RIZZOLI, 2018) different factors contribute to the total coherence and the coherence can be written as a product of the different contributions to the coherence as follows:

$$\rho = \rho_{SNR} \cdot \rho_{Quant} \cdot \rho_{Amb} \cdot \rho_{Range} \cdot \rho_{Azimuth} \cdot \rho_{Vol} \cdot \rho_{Temp} \quad (2.17)$$

The terms on the right hand side describe the errors contribution due to: limited Signal to noise ratio ( $\rho_{SNR}$ ), quantization errors( $\rho_{Quant}$ ), ambiguities( $\rho_{Amb}$ ), baseline decorrelation ( $\rho_{Range}$ ), errors due to relative shift of the doppler spectra ( $\rho_{Azimuth}$ ), and temporal decorrelation( $\rho_{Temp}$ ).

The term ( $\rho_{Vol}$ ) is the volume correlation factor and it is due to volume scattering. On forests this is mainly due to vegetation and therefore is crucial for forest land-cover classification.

A more detailed explanation about each contribution to the coherence is given by (MARTONE; RIZZOLI, 2015) and (RIZZOLI *et al.*, 2017), but a summary of each contribution is given following:

- $\rho_{SNR}$ : Is the coherence loss due to finite sensitivity of the radar system. It can be calculated from the Signal to Noise Ratio of the SAR System as follows:

$$\rho_{SNR} = \frac{SNR}{1 + SNR} \quad (2.18)$$

According to (MARTONE *et al.*, 2012) for the TANDEM-X satellite system it holds that:

$$\rho_{SNR} > 0.8$$

- $\rho_{Quant}$ : It represents the error due to quantization of the recorded received signal. From (MARTONE; RIZZOLI, 2015) it is expected that this value is lower than 10%.
- $\rho_{Range}$ : Possible baseline decorrelation effects are estimated by deriving a local slope map from orbit and elevation information.
- $\rho_{Temp}$ : Is the decorrelation due to the temporal baseline between both acquisitions. On TANDEM-X, since both acquisitions are taken almost at the same time it holds that  $\rho_{Temp} \approx 1$ .
- Also according to (MARTONE *et al.*, 2012) and (KRIEGER *et al.*, 2007) the contributions of  $\rho_{Amb}$  and  $\rho_{Azimuth}$  are very small (less than 2%) and are solved in the TANDEM-X by adaptative selection of the azimuth processing bandwidth and total independent zero Doppler steering.

In (KRIEGER *et al.*, 2007) there is a more detailed explanation about the different correlation factors.

In (SICA *et al.*, 2019) it was also shown that a study about the evolution of the temporal correlation of an area can also provide useful information for land-cover classification. The temporal decorrelation is modelled in (SICA *et al.*, 2019) as:

$$\rho_{Temp} = (1 - \rho_{LT})e^{-(\frac{t}{\tau})^2} + \rho_{LT} \quad (2.19)$$

where  $\gamma_{LT}$  is the long term coherence,  $t$  is the time and  $\tau$  is the target decorrelation factor. The determination of these parameters also provide important information that might help improve the land cover classification algorithms. The work developed in (SICA *et al.*, 2019) is crucial for this thesis, since this thesis is strongly based on this work and it lays the foundation for an improvement of land cover classification that is proposed on the end of this work.

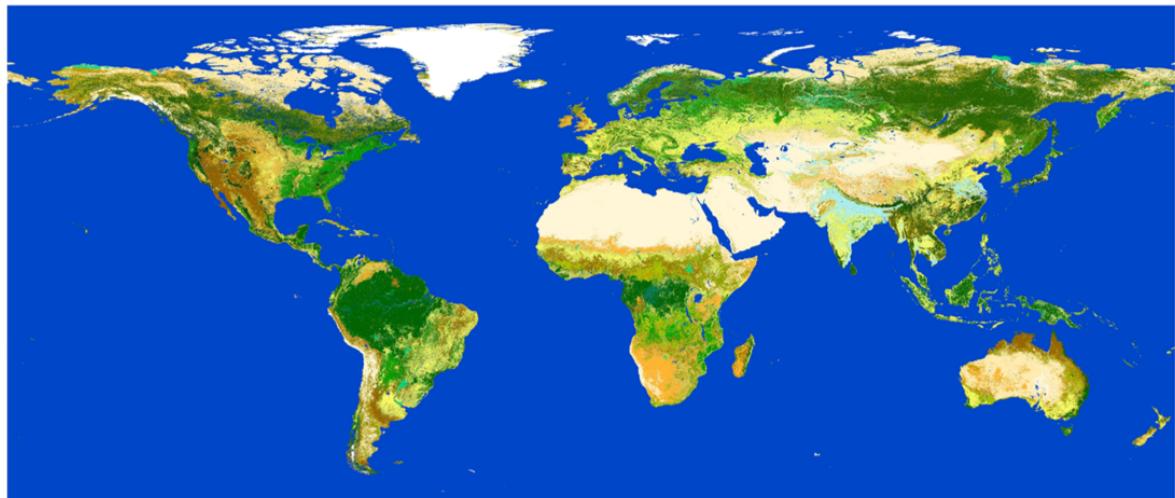
### 3 Classification Maps generated with SAR Images

The goal of land cover mapping is to make a classification of different areas the Earth's global surface. Since land cover classification has several economical/social applications, such as, monitoring and environmental planning, it is crucial that classification maps be generated with high accuracy.

Classification maps can be generated in two different ways: by field measuring or through remote sensing. Field measuring provides a high accuracy classification map but there is the obvious problem that it does not scale well to large cover areas, since it takes a lot of time, money and it is limited to areas that can be easily accessed.

Because of these problems mentioned above, classification maps are normally generated by using remote sensing data. The GlobCover (European Space Agency GlobCover Portal) map is one example of a classification map that was generated using MERIS data.

Some of the different areas of interest can be classified as: Vegetated areas, artificial surfaces (human made surfaces like constructions, buildings, roads, houses, etc...), water bodies, permanent snow or ice, among others. Some of these classes can be subclassified among different classes, for example, vegetation can mean: irrigated croplands, deciduous forests, mosaic grassland among others. The GlobCover classification map, provided by the European Space Agency is used to perform classification (European Space Agency GlobCover Portal). The GlobCover is composed of 23 different classes, which can be seen in the figure below.



1	Post-flooding or irrigated croplands (or aquatic)	
2	Rainfed croplands	
3	Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%)	
4	Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%)	
5	Closed to open (gt 15%) broadleaved evergreen or semi-deciduous forest (> 5m)	
6	Closed (>40%) broadleaved deciduous forest (>5m)	
7	Open (15-40%) broadleaved deciduous forest/woodland (>5m)	
8	Closed (>40%) needleleaved evergreen forest (>5m)	
9	Open (15-40%) needleleaved deciduous or evergreen forest (>5m)	
10	Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m)	
11	Mosaic forest or shrubland (50-70%) / grassland (20-50%)	
12	Mosaic grassland (50-70%) / forest or shrubland (20-50%)	
13	Closed to open (>15%) (broadleaved or needleleaved) shrubland (<5m)	
14	Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)	
15	Sparse (<15%) vegetation	
16	Close to open (>15%) broadleaved forest regularly flooded	
17	Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water	
18	Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil	
19	Artificial surfaces and associated areas (Urban areas >50%)	
20	Bare areas	
21	Water bodies	
22	Permanent snow and ice	
23	Invalid pixels - No data (burnt areas, clouds)	

FIGURE 3.1 – GlobCover classification map

Among the different classes mentioned above forests are a class that plays a key role in Earth's ecosystem. Therefore it is very important to analyse the deforestation or change in forest areas in order to assess the impact of the ecosystem.

Nowadays, optical and lasers sensors are used for land cover classification and forest

changes, each with its own advantages. For example, optical sensors provide an easy way to make classification, but it has the disadvantages that it is weather dependent and cannot provide a high resolution map. SAR mapping has none of these disadvantages but it is harder to make an accurate classification. Given the necessity to generate classification maps throughout the year, independent of cloud coverage, detected SAR backscatter is widely used for forest mapping (KRIEGER *et al.*, 2007).

Some of these maps are very successful in the task of making an accurate forest map, such as the global forest/nonforest map L-Band ALOS PALSAR which was generated creating a threshold on the cross-polarization levels of the detected backscatter. Even though it is possible to generate classification maps using the signature backscatter, this work will focus on the creation of said maps by analyzing the interferometric coherence.

The first use of interferometric data for forest mapping is reported on (WERTMULLER; WERNER, 1995). This work was used using ERS-1 SAR data and not only proved that forests can be clearly discriminated from other land categories, but it also showed that it is possible to distinguish among different forest types. This first approach consisted of analyzing the different values for the mean and standard deviation of the coherence for different areas, and by verifying that these values are separated. This method is visually explained in the picture 3.2.

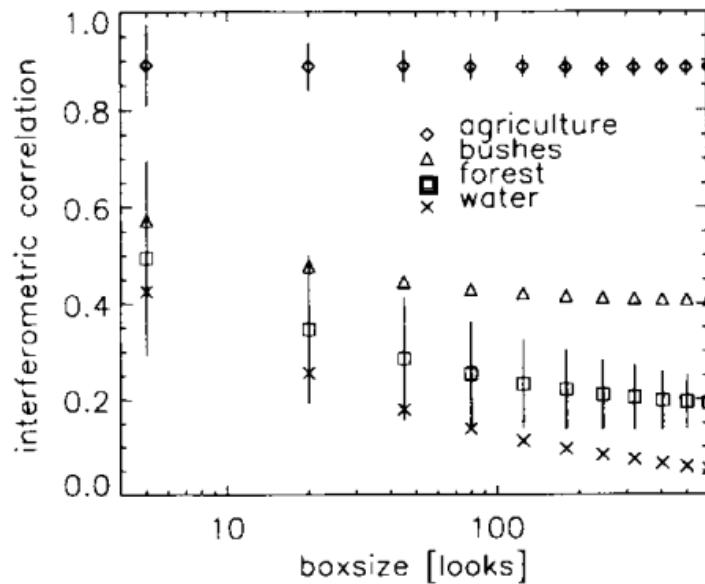


FIGURE 3.2 – Mean and Standard deviation of the coherence estimate as number of looks (filter size) for different classes

Nowadays, interferometric data can be used to monitor not only deforestation, but a series of other phenomena like topographical change and ice melting on polar caps (RIZZOLI, 2018).

On this work it will be presented the use of the ESA Sentinel-1 mission and its applications for forest mapping using the interferometric acquired data.

The Sentinel-1 mission consists of two satellites which acquire data on a small temporal baselines (normally between 6 and 12 days) which use the C-band frequency range to acquire data in swaths of over 250km in range direction. Even though, like every SAR, the main product is the backscatter received from the target, the focus of this work is to rely only on the interferometric information for classification purposes.

Some works rely on acquiring the backscatter data over large periods of time (months or years), and then classifying the target area based on the temporal change of it (BRUZ-ZONE *et al.*, 2004), this method is named long-time-series classification. Since the coherence is subjected to more variation over small periods of time, it is possible to make a classification based on its value, which don't have to be acquired over months, but over a few weeks is enough, this is why this method is called short-time-series classification, and since it has the capability of generate more updated maps, it was chosen as the method for this analysis.

On this work it will be demonstrated that the short time series interferometric information is valuable resource for classification. Besides using the temporal model for classification, the result will also be combined with state of the art machine learning algorithms in order to even more improve the accuracy of the result obtained.

# 4 Material

## 4.1 Test Area: The Rondonia State

The test area chosen for this thesis is the Amazon Rainforest, focusing specifically in the Rondonia state area. Rondonia topography is composed mainly by flat lands, which also have a lot of rivers throughout them, but also has regions of plateau with low altitude. The climate in Rondonia is what is called "humid equatorial" climate, which means that the temperature variation thorough the year is very little, something positive, since it is preferable to have no variations in the scene between SAR acquisition. The pluviometric indexes in the state can reach up to 2100mm per year, with most of the rains happening between May and September. The vegetation in the state (the focus of this thesis), just like most of north of Brazil, is the Amazon Rainforest, even though in some areas the *Cerrado* vegetation is more common. The *Cerrado* area is known for a very humid season and a very dry season, when the trees lose their leaves as a form of adaptation.

The goal of this thesis is to improve algorithms to identify deforestation areas using SAR images. Because of that the natural choice of area to study was the Rondonia state, since it is the state with the most deforestation in the Amazon Rainforest. The Rondonia state already lost over 31% of its forests and most of the remaining areas are degraded. For comparison, Acre, the state which borders Rondonia on the west, has 91% of its original forest cover and a greater part of it is still intact.

Even though there are a lot of efforts being made to control the illegal deforestation, the results are not yet satisfactory and recent data even shows that the pace of deforestation have increased in the years of 2018 and 2019.

The deforestation in the Rondonia state can be easily seen with optical satellite data acquired with Google Earth. Deforestation follows a fairly predictable pattern, as seen in figure 4.1. The pattern of deforestation is known as fishbone pattern due to its similarity with a fish skeleton. This pattern arises from the fact that deforesting is normally done by penetrating the forest and then deforesting along the edges of the road firstly created.



FIGURE 4.1 – Fishbone pattern of deforestation

Due to recent fires that happened in the year of 2019 in the Amazon Rainforest, there is much concern about studying and monitoring the deforestation that happens in that area, and considering that the Rondonia state is the area in which the deforestation is most critical, it was a natural choice of area to study for this thesis.

## 4.2 TanDEM-X datasets

TerraSAR-X (TSX) and TanDEM-X (TDX), launched in June 2007 and June 2010, respectively, are two German SAR satellites operating in X-band, developed within a public/private partnership between the German Aerospace Center (DLR) and Airbus Defence and Space.

The goal of both satellites is to provide SAR products for commercial purposes and scientific purposes, and the TanDEM-X mission has the primary goal of generating a global, high precision, and consistent digital elevation model (DEM) with full coverage and no gaps. The relevance of the mission lies in that, until now, the available DEMs of large parts of Earth are of low resolution, inconsistent, incomplete and commonly based on different data sources and survey methods.

TanDEM-X has offered, for the first time, a global, high accuracy and homogeneous DEM. Besides the main goal, other secondary missions based on along-track interferometry have been defined as well as new techniques with bistatic SAR (MOREIRA *et al.*, 20-24 Sept. 2004; HAJNSEK; MOREIRA, May 2006). The work presented on this thesis is one of the possible exploitation of the activities of the TanDEM-X data.

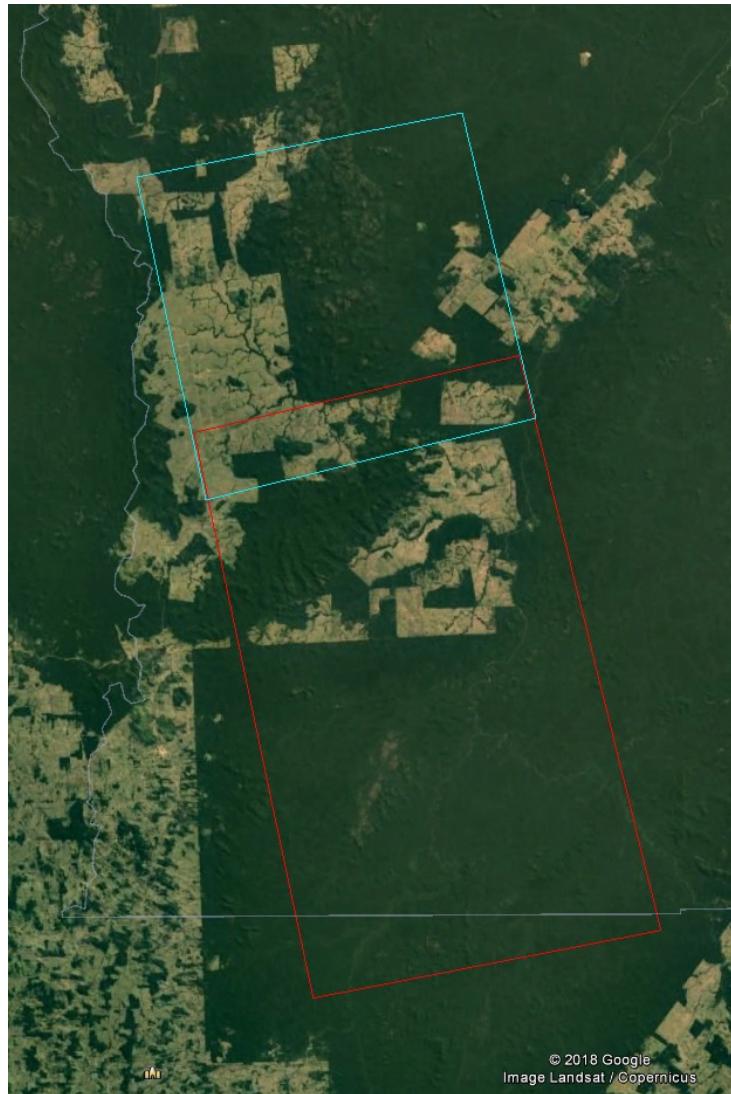


FIGURE 4.2 – TanDEM-X dataset available in Rondonia Area

TanDEM-X has the advantage that it can provide very high resolution data (each pixel has dimensions of 1.5m X 1.5m). Since the data is very resolution, and therefore very big in terms of memory size, it was chosen to work with small areas in the Rondonia state. The dataset available is on the east of the Rondonia State. It was collected images from two areas in east Rondonia as can be seen in the figure 4.2. Each rectangle represents one SAR image acquired with the TanDEM-X.

### 4.3 Sentinel-1 datasets

The SENTINEL-1 mission is a joint initiative of the European Commission (EC) and the European Space Agency (ESA). Copernicus, previously known as GMES, is a European initiative for the implementation of information services dealing with environment and security. It is based on observation data received from Earth Observation satellites

and ground-based information.

Sentinel-1 is an imaging radar mission operating in C-band consisting of a constellation of two satellites aimed at providing day and night supply of imagery for users, independent of the weather at the time. It also provides dual polarisation capability, very short revisit times and rapid product delivery. The constellation of two satellites, SENTINEL-1A and SENTINEL-1B, shares the same orbital plane.

According to (SNOEIJ *et al.*, December 2008) there are five priorities for the SENTINEL-1 mission: To monitor sea ice zones and the arctic environment, surveillance of marine environment, to monitor land surface motion risks, mapping of land surfaces and mapping in support of humanitarian aid in crisis situations.

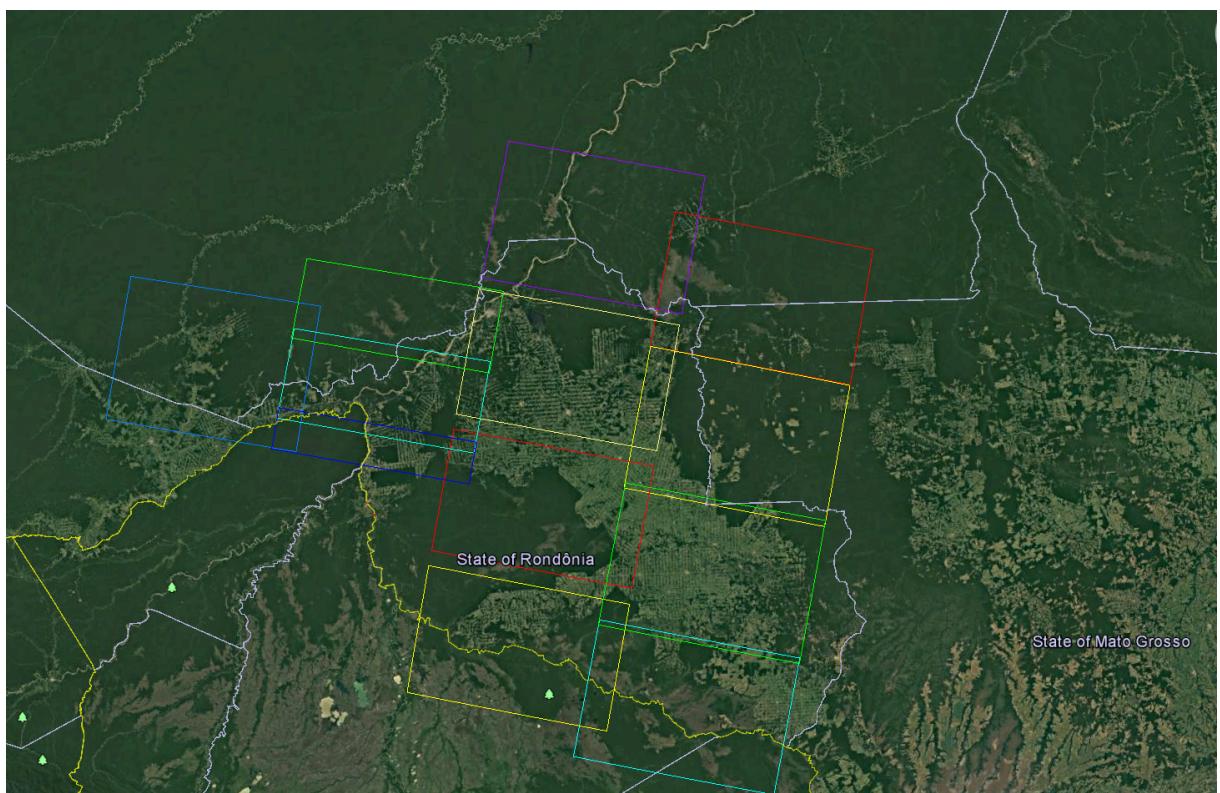


FIGURE 4.3 – SENTINEL-1 dataset available in Rondonia Area

SENTINEL-1 provides area coverage much greater than the TanDEM-X satellite with coverage up to 400km, but the resolution of the image is lower compared to the TanDEM-X mission. While the pixel resolution for SENTINEL-1 mission is 3.7mX14m, the TanDEM-X can provide pixels with resolution of 1.5mX1.5m. Since SENTINEL-1 provides area coverage much greater than TanDEM-X, it was chosen as main option for area monitoring in the Rondonia state as a whole. With SENTINEL-1 it was possible to cover almost the entire Rondonia state although the entire coverage was not done because there are some gaps between acquisitions as seen on image 4.3. Even though the resolution for the SENTINEL-1 is much lower than the resolution of the TanDEM-X, it will be investigated

on this thesis if this resolution is enough to detect new areas of deforestation. The SENTINEL-1 dataset for the Rondonia state can be seen on the image 4.3, where each rectangle is one acquistion made by SENTINEL-1.

# 5 Textures Methods for Image Analysis

## Summary

The objective of this work is to find techniques that might help improve classification and segmentation of SAR images.

Given an image of the radar brightness ( $\sigma^0$ ) of an area or the image of the coherence between two SAR acquisitions it is possible to segment the image into different areas of interest, for example, given an image of the Amazon Forest someone might be interested in generating a map of the area that was deforested and the area that it is still preserved. Even though there are many machine-learning algorithms that can do this with an acceptable accuracy, these algorithms are not perfect and have a error that is noticeable. To improve the result of those algorithms the textures of the image will be analyzed and then it will be validated if it is possible to extract useful information from these textures that otherwise would be hidden to the traditional algorithms for classification, like random forest or neural networks.

In some images, the texture might be a defining feature of a specific region and critical to obtain a correct classification. The main idea of the looking at the textures of an image is to extract information based not only on the value of intensity of the pixel, but also looking at the pixels that are around that one, such that the value of a texture in a single pixel is a function that depends on the value of the original pixel and the pixels that are nearby it.

The textures are a function of the spatial arrangement of intensities and colors in an image. Even though different images might have the same histograms of pixel intensities, they can be very different, and sometimes, analyzing just the histograms of the pixel intensities might not be enough to define in which region a pixel is in. There are two main approaches to analyze the textures of an image: the structural approach and the statistical approach. In this work it will be only used the statistical approach for analyzing textures.

## 5.1 Machine Learning and the problems for classification

Machine Learning is one of the most important and influential technologies in today's world. It is also one of the areas that has the most focus on research since the demand is very high, therefore it is very likely that there is still much room for improvement and its full potential has not yet been reached.

Machine learning is a tool that was developed to deal with the excessive amount of data that has started to appear from 50 years ago until now. But this excessive amount of data is useless if there is no way to extract useful information from it. Focusing on this problem, machine learning is a field of study that tries to analyze data and find the hidden patterns hidden within it. Finding these patterns might be useful to make predictions of a problem and even help taking decisions on how to act when presented a situation.

The name machine learning comes from the fact that the algorithms have to go through a learning process, trying different rules and seeing how well they perform to describe the problem presented. Among the different forms of machine learning, on this work there will be a focus on the supervised machine learning, which is a type of machine learning system that is presented with data which is labeled, that means that each data has a correct label. The goal of this machine learning is that when presented with new data it is possible to predict in which label that data fits.

The are five steps to this process: Data collection, data preparation, model fitting, model evaluation and parameter tuning.

1. Data collection: Collect the data the algorithm will learn from (in this case it is the SAR image)
2. Data preparation: Format the data into optimal format for the algorithm.
3. Training: The stage in which the algorithm derives a model by fitting the output and the input using some method.
4. Evaluation: Check how well the algorithm performs
5. Parameter Tuning: Tune the parameters to improve performance

The types of Supervised learning are two: Classification and regression.

1. Classification: The output is a category, like a color, or in the case of a SAR image, if the pixel indicates a area of forest or a deforested area.
2. Regression: The output is a real number, such as height or price.

On this work, classification machine learning will be used to classify a SAR image into regions of forest and deforested areas.

There are several machine learning algorithms, such as: C-support vector classifier, K-nearest neighbors classifier, multi-layer perceptron, Gaussian process classification, Radial-basis function kernel, Decision tree classifier, AdaBoost classifier, Random forest classifier, Gaussian Naive Bayes classifier, Quadratic discriminant analysis classifier and neural network classifier.

All those different algorithms try to solve the same problem: given a set of input labeled data, what is the model that best fits the labeled input data and the given label, in order to predict labels for unlabeled data with least error?

The difference between those algorithms is that the model that each one will find is different. Some algorithms, like the neural network, will try to derive a model using linear functions on the input data and trying to minimize the mean square error(MSE) between the predicted output and the label for example, as the decision trees for example will try to create a series of binary questions and by combining the answers of these questions will give the label of the input data.

Either way, the model of each algorithm can be geometrically understood by creating decision regions in a  $R^n$  space (n is the number of input variables), and by looking in which region a input is located, it is possible to give the output answer.

For example, if the number of input variables is 2, and different algorithms are run to classify labeled data between two colors (blue and red) the result can be seen in the figure below.

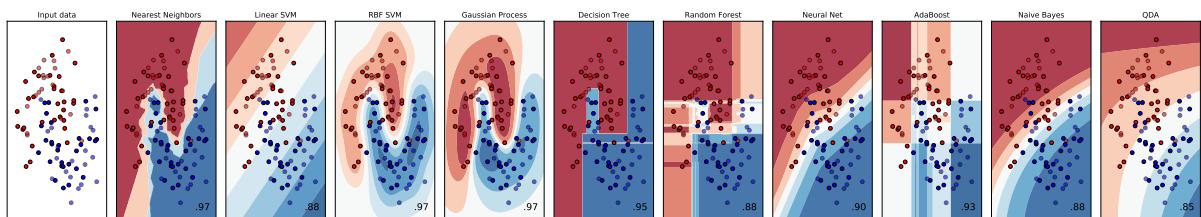


FIGURE 5.1 – Classification Boundaries for Different Algorithms

In the image 5.1 the color intensity illustrates how certain the algorithm is of the classification and the number on each figure is the time it took to create that region. For different algorithms it is clear that the classification regions it derives are different and how the algorithm creates these regions determines how well the algorithm classifies. The problem of using machine learning for classifying forest and deforested areas is that the algorithm relies too much on creating decision boundaries and by checking in which decision region a certain pixel is, but it does not check the surroundings of that pixel, which is also a important information for classification.

For example, in the image 5.1 above, if a red pixel falls in a blue region it will be wrongly labeled blue. For a SAR image it might happen that due to noise a forest pixel looks like a deforested pixel, but if all the pixels around it looks like forest pixels then this might be an indicator that that pixel is a wrongly labeled forest pixel. Even though it is not possible to make these machine learning algorithms consider this question in the learning process (this would mean interfering in the modelling process, which would be equivalent to creating a new algorithm), it is possible to create more input data by looking at the surroundings of a pixel. This new data information can help this algorithm not to make a wrong classification. This new input data is exactly the textures of a image.

## 5.2 Textures Methods

As previously mentioned, the textures might give useful information for classification because sometimes the neighborhood of a pixel also provides valuable information for the classification. In forest landscapes, for example, the texture value might depend on the size and distance between trees, such that in high-resolution images if two pixels fall in the same tree then they will have similar value, resulting in a small local variance of intensities, something that will be indicated by the texture value. According to (WOODCOCK; STRAHLER, 1987) if the resolution is increased to a size comparable to the size to of trees, then the local variance also increases, something noticeable in tropical forests with a high species diversity. Therefore is important to mention that the texture is dependent of the resolution of the image, and the texture of a high resolution image of an area might be different than a low resolution image of the same area.

On this work it will be used three different textures methods for improving classification results: the Grey level co-occurrence matrix(GLCM) method, the Laws textures method and the sum and different histograms methods.

## 5.3 The GLCM Method

The first method of texture creation is the grey-level-co-occurrence matrix method. A co-occurrence matrix is a matrix extracted from an image in which the values in the rows and columns of this matrix represent the set of possible grey scale values of the image. For example, the co-occurrence matrix  $C$  is a matrix in which the elements represent the possible co-occurrence of values for the image  $I$  given a spatial relationship on a set of possible values  $V$ . For example, given a image, the co-occurrence matrix  $C$  indicates how many times the value  $i$  co-occurs with the value  $j$  given a spatial relationship. This spatial

relationship is given by a displacement vector  $d = (d_r, d_c)$  that dictates the distance of the pixels that one wants to analyze the co-occurrence of values. A more mathematical way to express this matrix is given by the following definition:

$$C_d(i, j) = \#\{(r, c) | I(r, c) = i \text{ and } I(r + d_r, c + d_c) = j\} \quad (5.1)$$

For example, if the grey scale image  $I$  is equal to:

$$I = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{bmatrix} \quad (5.2)$$

Then three co-occurrence matrices for different displacement vectors  $d = (0, 1)$ ,  $d = (1, 0)$  and  $d = (1, 1)$  are:

$$C_{(0,1)} = \begin{bmatrix} 4 & 0 & 2 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (5.3)$$

$$C_{(1,0)} = \begin{bmatrix} 4 & 0 & 2 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (5.4)$$

$$C_{(1,1)} = \begin{bmatrix} 2 & 0 & 2 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.5)$$

In  $C_{(0,1)}$  the position  $(0, 0)$  equals 4, which means that  $j = 0$  appears to the right of  $i = 0$  four times in the image. The position  $(1, 0)$  equals 2 because  $j = 0$  appears to the right of  $i = 1$  two times in the image. The position  $(2, 1)$  equals 0 because  $j = 1$  never appears to the right of  $i = 2$ .

From this grey level co-occurrence matrix it is important to generate the *normalized* grey level co-occurrence matrix  $N_d$  given by:

$$N_d(i, j) = \frac{C_d(i, j)}{\sum_i \sum_j C_d(i, j)} \quad (5.6)$$

Co-occurrence matrices capture the texture properties, but are not directly useful for further analysis, such as comparing different textures. To solve this problem, is possible to compute numeric features from this normalized co-occurrence matrix that might give useful information of the image that otherwise would be hidden for machine learning classification algorithms. These numeric features computed from the normalized co-occurrence matrix represent more compactly the textures of the image. The following textures are features that can be obtained from normalized co-occurrence matrix:

Contrast	$\sum_{i,j=0}^{N-1} i \frac{N_d(i,j)}{1+(i-j)^2}$
Dissimilarity	$\sum_{i,j=0}^{N-1} i N_d(i,j)  i - j $
Homogeneity	$\sum_{i,j=0}^{N-1} \frac{N_d(i,j)}{1+ i-j }$
Angular Second Moment	$\sum_{i,j=0}^{N-1} i N_d(i,j)^2$
Correlation	$\sum_{i,j=0}^{N-1} \frac{(i-\mu)(j-\mu)N_d(i,j)}{\sigma^2}$
Energy	$\sum_{i,j=0}^{N-1} N_d(i,j)^2$
Entropy	$-\sum_{i,j=0}^{N-1} N_d(i,j) \log_2(N_d(i,j))$

TABLE 5.1 – Formula for texture calculations for GLCM Method

Where  $\mu$  is the GLCM mean and  $\sigma$  is the standard deviations of the GLCM given by:

$$\mu = \sum_{i,j=0}^{N-1} i N_d(i,j) \quad (5.7)$$

$$\sigma^2 = \sum_{i,j=0}^{N-1} N_d(i,j) (i - \mu)^2 \quad (5.8)$$

A question that rises from this method is how to choose the displacement vector  $d = (d_r, d_c)$  such that the result is the best. A solution proposed by (ZUCKER; TER-ZOUPOULOS, 1980) is to run a hypothesis test to check the displacement vector  $d$  that most rejects the hypothesis that the pair of pixels separated by the displacement vector  $d$  are independent, therefore accepting the hypothesis that this pair of pixels give the most information one of the another and are better suited to compute a texture feature. The hypothesis test suggests to run a  $\chi^2$  statistical test to find the value  $d$  that give the most information, that is, to maximize the value:

$$\chi^2(d) = \left( \sum_i \sum_j \frac{N_d^2(i,j)}{\sum_j N_d(i,j) \sum_i N_d(i,j)} - 1 \right) \quad (5.9)$$

## 5.4 The Laws Textures

A different way to create textures is to use 2D filters to detect different kinds of textures. Laws developed a method to measure the amount of variation of the image on a fixed sized window. The method consists of making 16 different convolutions of the original image with convolution masks and use the result to compute the texture. The 16 different filters are obtained by taking the product of the following vectors.

$$L5(Level) = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (5.10)$$

$$E5(Edge) = \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \end{bmatrix} \quad (5.11)$$

$$R5(Ripple) = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \end{bmatrix} \quad (5.12)$$

$$S5(Spot) = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \end{bmatrix} \quad (5.13)$$

Then the 2D convolution masks are obtained by taking the product of these vectors. For example the R5S5 mask is computed as the product of R5 and S5 as following:

$$\begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \\ -1 \end{bmatrix} \times \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ 4 & 0 & -8 & 0 & 4 \\ -6 & 0 & 12 & 0 & -6 \\ 4 & 0 & -8 & 0 & 4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix} \quad (5.14)$$

After applying the 16 5x5 masks to the image it is necessary to generate the texture energy map on a fixed size window. Let the window size be 15x15 and  $F_k[i, j]$  the result of the  $k$ th mask on the pixel [i,j]. Then the texture energy map  $E_k$  is defined by:

$$E_k[r, c] = \sum_{j=c-7}^{c+7} \sum_{i=r-7}^{r+7} |F_k[i, j]| \quad (5.15)$$

After creating the 16 energy maps it is needed to group together similar images that perceive the textures in different directions. For example, if R5S5 detects at ripples in the vertical direction, then S5R5 detects ripples in the horizontal direction, so it is desirable to group them together to detect ripple in both directions. The average of these two images measure then the total ripple content. It is possible to group the 16 energy maps

in 9 texture images as following:

$$\begin{array}{ll}
 L5E5/E5L5 & L5S5/S5L5 \\
 L5R5/R5L5 & E5E5 \\
 E5S5/S5E5 & E5R5/R5E5 \\
 S5S5 & S5R5/R5S5 \\
 & R5R5
 \end{array}$$

## 5.5 The Sum and Difference Histograms Texture

### 5.5.1 The Single Value Decomposition

Let  $I[k, l]$  be a discrete image which is the realization of a bi-dimensional stationary process and let  $G = \{1, 2, \dots, N_g\}$  be the set of the  $N_g$  quantized grey levels. Just as in the GLCM method, consider a pair of pixels  $y_1$  and  $y_2$  separated by a displacement vector  $d = (d_1, d_2)$  such that:

$$\begin{cases} y_1 = I[k, l] \\ y_2 = I[k + d_1, l + d_2] \end{cases} \quad (5.16)$$

The discrete joint probability function of these two pixels is  $P(y_1, y_2)$ . And the probability of observing a pair of grey-level occurrence  $i$  and  $j$  separated by a displacement vector  $(d_1, d_2)$  is given by:

$$Prob\{y_1 = i, y_2 = j\} = P(i, j, d_1, d_2) = P(i, j) \quad (5.17)$$

which does not depend on the pixel position  $[k, l]$ .

Pay attention to the fact that an estimate of this joint distribution  $P[i, j]$  is the normalized co-occurrence matrix given by 5.6, such that:

$$\hat{P}(i, j) = N_d[i, j] \simeq P(i, j) \quad (5.18)$$

Let us consider the random vector of the two random variables  $y_1$  and  $y_2$ . Let  $Y$  be this random vector such that  $Y = [y_1, y_2]^T$ . Suppose that this random vector  $Y$  has a co-variance matrix  $C_{yy}$ . The goal is to make a linear transformation on this vector such that the new vector has uncorrelated variables. Suppose the linear transformation is given

by a matrix  $H$  and the new random vector is  $Z$  such that:

$$Z = HY \quad (5.19)$$

It is known from (PAPOULIS; PILLAI, 2002) that the co-variance matrix  $C_{zz}$  of the random variable  $Z$  is obtained by the following formula.

$$C_{zz} = HC_{yy}H^T \quad (5.20)$$

If we make a single value decomposition on the matrix  $C_{yy}$  such that  $C_{yy} = U\Lambda^2U^T = U\Lambda\Lambda U^T$  and we choose the matrix  $H$  such that  $H = \Lambda^{-1}U^T$  then the co-variance of the random vector  $Z$  is given by:

$$C_{zz} = HC_{yy}H^T = \Lambda^{-1}U^T U\Lambda\Lambda U^T U\Lambda^{-1} = I \quad (5.21)$$

Therefore the components of  $Z$  are uncorrelated.

Let us model the random vector  $Y$  to have similar cross-correlation between the variables, such that:

$$C_{yy} = \sigma_y^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \quad (5.22)$$

where

$$\sigma_y^2 \rho = E\{(y_1 - \mu)(y_2 - \mu)\} \text{ and } \mu = E\{y_1\} = E\{y_2\} \quad (5.23)$$

And due to stationarity

$$E\{(y_1 - \mu)^2\} = E\{(y_2 - \mu)^2\} = \sigma_y^2 \quad (5.24)$$

Let us make then the transformation given by

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (5.25)$$

That results in the new random vector  $Z = [z_1, z_2]^T$  that:

$$\begin{cases} z_1 = \frac{y_1 + y_2}{\sqrt{2}} \\ z_2 = \frac{y_1 - y_2}{\sqrt{2}} \end{cases} \quad (5.26)$$

Such that the new co-variance matrix for  $Z$  is given by:

$$C_{zz} = \begin{bmatrix} \sigma_y^2(1 + \rho) & 0 \\ 0 & \sigma_y^2(1 - \rho) \end{bmatrix} \quad (5.27)$$

### 5.5.2 Approximation of Probability Density Functions

Let  $\alpha \in N_g$  and  $\beta \in N_g$ . If we assume that the variables for  $y_1$  and  $y_2$  are gaussian, then  $z_1$  and  $z_2$  are also gaussian and uncorrelated, therefore the joint probability function can be calculated from:

$$P(y_1 = \alpha, y_2 = \beta) = P(z_1 = \alpha + \beta, z_2 = \alpha - \beta) = P_s(z_1 = \alpha + \beta) \cdot P_d(z_2 = \alpha - \beta) \quad (5.28)$$

If the variables are not gaussian then the equality is not valid, but it is still a valid approximation to the probability function  $P(y_1, y_2)$ .

$$\hat{P}_y(i, j) = c_0 \cdot P_s(i + j) \cdot P_d(i - j) \simeq P_y(i, j) \quad (5.29)$$

Where  $c_0$  is a normalization constant that assures that

$$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \hat{P}_y(i, j) = 1 \quad (5.30)$$

The relative error of this approximation can be computed analyzing the Kullback-Leibler Divergence between the two variables  $P$  and  $\hat{P}$  as follows

$$I(P, \hat{P}) = \sum_i \sum_j P_y(i, j) \cdot \log\left(\frac{P_y(i, j)}{\hat{P}(i, j)}\right) = H_s + H_d - H_y - \log(c_0) \geq 0 \quad (5.31)$$

Where  $H_s$ ,  $H_d$  and  $H_y$  are the entropies given by:

$$\begin{cases} H_y = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_y(i, j) \cdot \log(P_y(i, j)) \\ H_s = -\sum_{k=2}^{2N_g} P_s(k) \cdot \log(P_s(k)) \\ H_d = -\sum_{l=-N_g+1}^{N_g-1} P_d(l) \cdot \log(P_d(l)) \end{cases} \quad (5.32)$$

This divergence is a measure of the error of approximating one distribution for another, and is equal to zero only when the two variables have the same distribution, therefore

it will only be zero if  $P(i, j) = \hat{P}(i, j)$ . The importance of 5.31 is that it is possible to see the independence of the new variables by comparing only the entropy of the sum and difference to the entropy of the co-occurrence matrix. The closer the mutual information  $I(P, \hat{P})$  is to zero, the better the approximation defined by 5.29

### 5.5.3 Textures Features

It was shown that the sum and difference are a linear transformation that uncorrelates the random variables of the pair of the pixels. Therefore it is suggested by (UNSER, 1986) to replace the regular co-occurrence matrix by the associated histograms of the sum and difference estimated from the original image. The non-normalized sum and difference associated with the displacement vector  $d = (d_1, d_2)$  are given by:

$$\begin{cases} s_{k,l} = I[k, l] + I[k + d_1, l + d_2] \\ d_{k,l} = I[k, l] - I[k + d_1, l + d_2] \end{cases} \quad (5.33)$$

And then we make the normalized histograms of these new vectors. The histograms of the sum and difference are given by:

$$\begin{cases} h_s(i, d_1, d_2) = h_s(i) = |\{(k, l) \in G, s_{k,l} = i\}| \\ h_d(j, d_1, d_2) = h_d(j) = |\{(k, l) \in G, d_{k,l} = j\}| \end{cases} \quad (5.34)$$

And the normalized sum and difference histograms are given by:

$$\begin{cases} \hat{P}_s(i) = \frac{h_s(i)}{\sum_i h_s(i)} & (i = 2, \dots, 2N_g) \\ \hat{P}_d(j) = \frac{h_d(j)}{\sum_j h_d(j)} & (j = -N_g + 1, \dots, N_g - 1) \end{cases} \quad (5.35)$$

And these normalized sum and difference histograms are estimates of the sum and difference probability functions given by:

$$\begin{cases} P_s(i) = Prob \{s_{k,l} = i\} & (i = 2, \dots, 2N_g) \\ P_d(j) = Prob \{d_{k,l} = j\} & (j = -N_g + 1, \dots, N_g - 1) \end{cases} \quad (5.36)$$

From these estimates of the probability density functions it is possible to compute textures. There are 9 important textures that will be used through this work. On the table below there are the computation formulas for them:

mean	$\frac{1}{2} \sum_i i \hat{P}_s(i)$
variance	$\frac{1}{2} (\sum_i (i - 2\mu)^2 \hat{P}_s(i) + \sum_j (j)^2 \hat{P}_d(j))$
energy	$\sum_i \hat{P}_s(i)^2 + \sum_j \hat{P}_d(j)^2$
correlation	$\frac{1}{2} (\sum_i (i - 2\mu)^2 \hat{P}_s(i) - \sum_j (j)^2 \hat{P}_d(j))$
entropy	$-\sum_i \hat{P}_s(i) \log_2(\hat{P}_s(i)) - \sum_j \hat{P}_d(j) \log_2(\hat{P}_d(j))$
contrast	$\sum_j j^2 \hat{P}_d(j)^2$
homogeneity	$\sum_j \frac{1}{1+j^2} \hat{P}_d(j)$
cluster shade	$\sum_i (i - 2\mu)^3 \hat{P}_s(i)$
cluster prominence	$\sum_i (i - 2\mu)^4 \hat{P}_s(i)$

TABLE 5.2 – Formula for texture calculations for Sum and Difference Histograms Method

It is also important to know which size the window should be to compute the histograms for the sum and difference. Through this work the window size chosen was to be a regular window of size 10x10. The number of gray levels is also important. Through this work, unless specified, the images will be normalized and scaled manually to have 10 different grey levels.

Theoretically, if the variables were gaussians, then this method would give the same result as the Gray-Level Co-Occurrence matrix method, since if the variables are gaussians are uncorrelated then they are independent and the formula given by (4.29) holds exactly and is not just an approximation. If they are not gaussians then the approximation is not perfect, but there is a great computational advantage to use this method, since making histograms of 1-D vectors is much faster than extracting a 2-D co-occurrence matrix. On the next section it will be analyzed if this method gives a nice result compared to the other 2 and if it is worth it to use this method due to the computational advantage.

# 6 Analysis of Texture Results

## Summary

On this section it will be demonstrated the results of different textures being applied to a SAR image and it will be analyzed which textures, if any, might help improve classification algorithms for forest areas and deforested areas. On this section it will be presented textures results of images of the coherence between 2 SAR images, and radar brightness ( $\beta^0$ )

### 6.1 The GLCM Textures

#### 6.1.1 The Displacement Vector

The first step to compute the textures of a image using the GLCM method is to choose the displacement vector  $d = (d_1, d_2)$  that will be used to compute the grey level co-occurrence matrix. As said in the previous chapter, one way to choose this value is to use a statistical test proposed by (ZUCKER; TERZOUPOULOS, 1980) to find the value of  $d$  that gives more texture information. The process to compute to find the best  $d$  consists of computing the grey level co-occurrence matrix for different displacement vectors and extracting the chi-squared statistical value for that occurrence matrix. The chi-square value is given by:

$$\chi^2(d) = \left( \sum_i \sum_j \frac{N_d^2(i, j)}{\sum_j N_d(i, j) \sum_i N_d(i, j)} - 1 \right) \quad (6.1)$$

It is expected that the closer the distance between the pixels the better, since pixels that are more close together in the image have bigger correlation between them. This statistical test will be run on coherence images over Amazon forest and  $\beta^0$  images over Amazon forest to confirm that pixels that are closer have greater correlation between them.

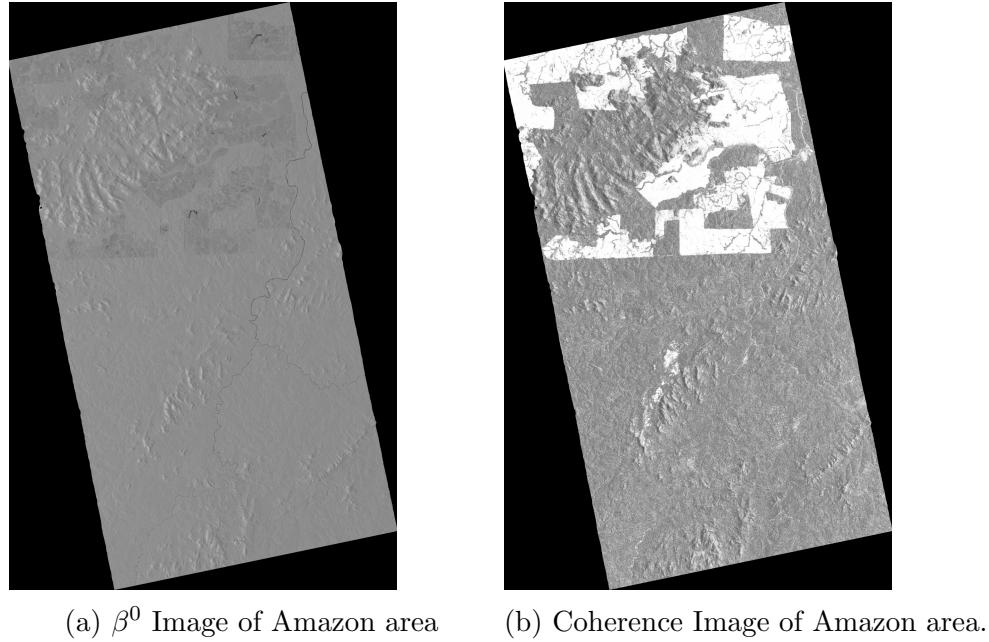


FIGURE 6.1 – Figura

Given these two images it will be calculated the result of (6.1) for different displacement vectors  $d = (d_1, d_2)$ . It is chosen that the displacement vectors have possible angles  $\theta$  equals to  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ , and magnitudes of 1, 2, 3, 4, 5, 6 pixels. The following tables show the result of the test for the images for areas of forest and deforested areas.

Distance	Average Coherence			
	$\theta = 0$	$\theta = 45$	$\theta = 90$	$\theta = 135$
1	13.46	7.95	13.56	7.96
2	5.13	4.06	5.21	4.13
3	4.04	4.40	4.13	4.38
4	4.19	5.30	4.80	5.18
5	5.42	6.71	5.50	6.50
6	6.22	7.22	6.39	6.99

TABLE 6.1 –  $\chi^2$  result for coherence image

Coherence of forest area				
Distance	$\theta = 0$	$\theta = 45$	$\theta = 90$	$\theta = 135$
1	13.01	7.36	12.68	7.10
2	4.52	3.75	4.48	3.77
3	3.65	4.32	3.91	4.34
4	4.14	5.20	4.58	5.10
5	5.12	6.57	5.21	6.41
6	6.91	7.06	6.08	6.87

TABLE 6.2 –  $\chi^2$  result for coherence in forest areas

Coherence of deforested area				
Distance	$\theta = 0$	$\theta = 45$	$\theta = 90$	$\theta = 135$
1	10.58	7.28	10.84	7.16
2	4.94	3.33	5.37	3.24
3	3.26	2.71	3.78	2.63
4	3.11	3.20	3.64	3.11
5	3.73	4.04	4.09	3.95
6	4.29	4.33	4.51	4.29

TABLE 6.3 –  $\chi^2$  result for coherence in deforested areas

$\beta^0$ of forest area				
Distance	$\theta = 0$	$\theta = 45$	$\theta = 90$	$\theta = 135$
1	1.26	0.63	1.59	0.73
2	0.28	0.20	0.42	0.24
3	0.22	0.24	0.26	0.23
4	0.29	0.29	0.28	0.27
5	0.35	0.39	0.31	0.35
6	0.39	0.44	0.37	0.39

TABLE 6.4 –  $\chi^2$  result for  $\beta^0$  in forest areas

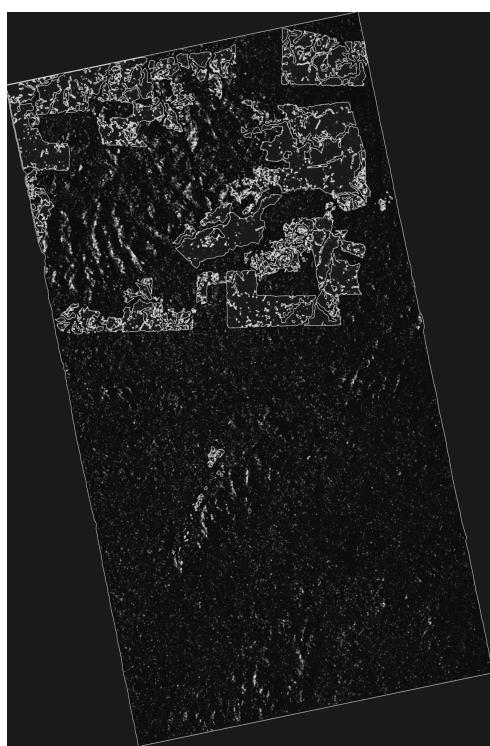
Distance	$\beta^0$ of deforested area			
	$\theta = 0$	$\theta = 45$	$\theta = 90$	$\theta = 135$
1	0.82	0.43	1.07	0.50
2	0.22	0.14	0.33	0.15
3	0.15	0.14	0.20	0.13
4	0.17	0.17	0.19	0.16
5	0.20	0.22	0.02	0.21
6	0.23	0.26	0.23	0.25

TABLE 6.5 –  $\chi^2$  result for  $\beta^0$  in deforested areas

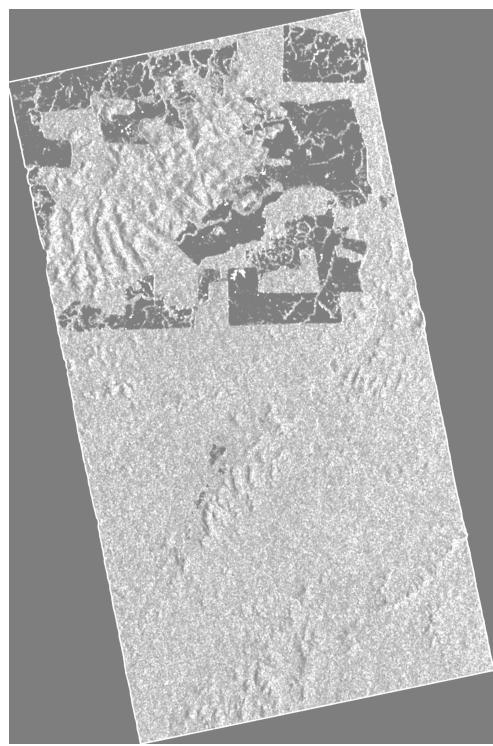
Analyzing the tables above we can see that choosing pixels that are closer in the image gives the maximum information for texture calculations (this means that there are no hidden patterns/structures in the images analyzed). According to the tables, until the rest of this work the displacement vector used will always be equals to  $d = (0, 1)$ .

### 6.1.2 The GLCM textures results

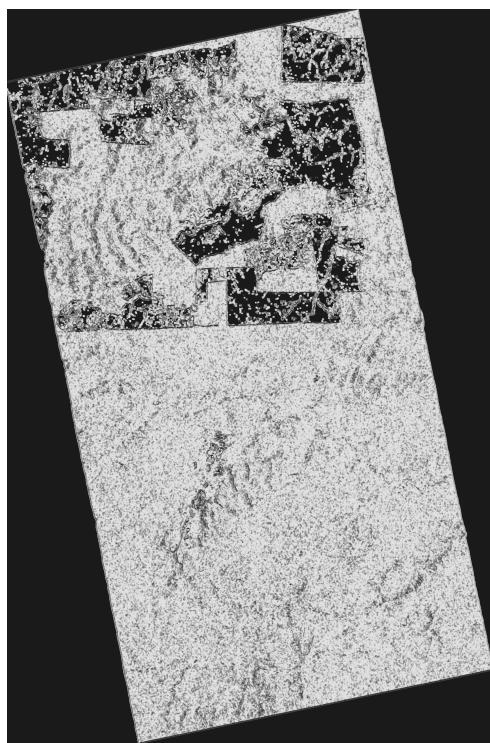
Let us visualize the texture results for the Coherence image above. It was computed 7 different textures for the coherence image above: the ASM, contrast, correlation, dissimilarity, energy, entropy and homogeneity. These textures can be visualized below



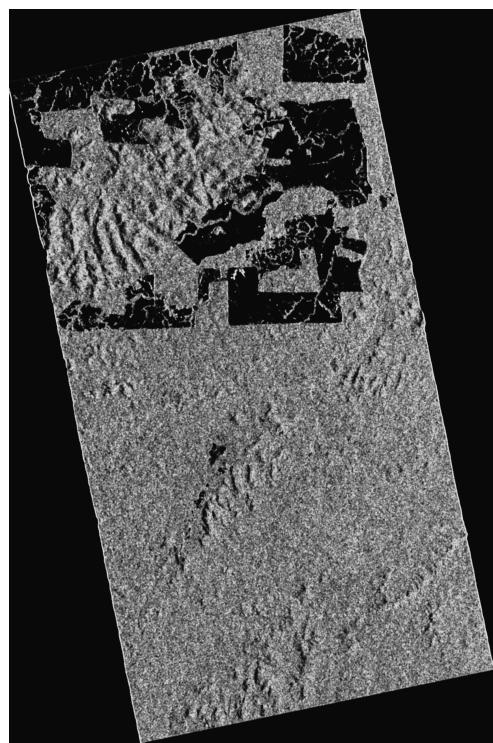
(a) ASM texture of Coherence Image



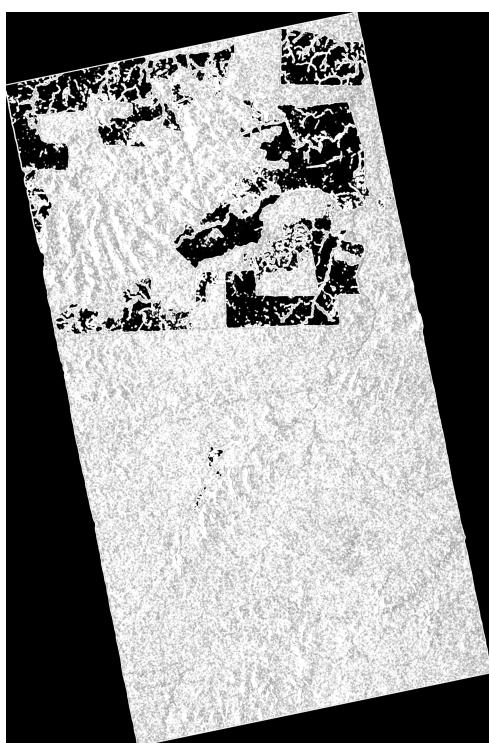
(b) Contrast texture of Coherence Image



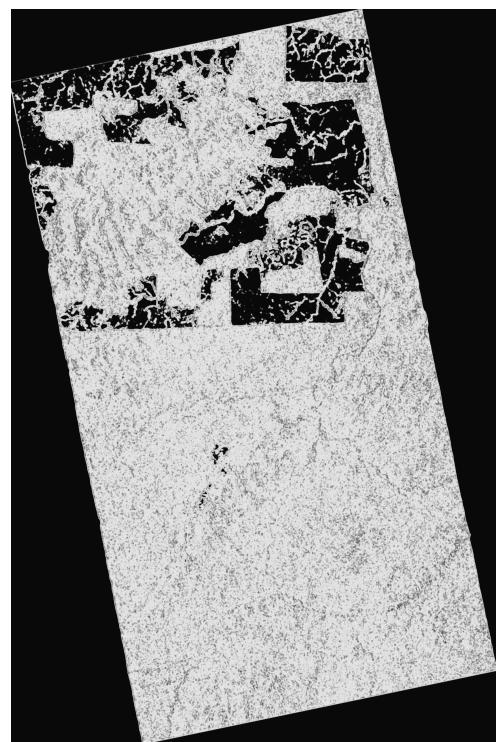
(c) Correlation texture of Coherence Image



(d) Dissimilarity texture of Coherence Image



(e) Energy texture of Coherence Image



(f) Entropy texture of Coherence Image



(g) Homogeneity texture of Coherence Image

Just by looking at the images above it might be hard to know if they are better suited for classification or not. One solution in order to be able to know quantitatively if they are better to classification is to look at the histogram of the pixels intensity and to compare

it with the histograms of the pixels intensity for forest and non forest area in the original coherence image. For this area it was provided the reference map for comparison, so it is known if a pixel is in a forest area or in a deforested area, and by using this reference map it is possible to make the histograms of the pixel intensity for each area and to compare it with the original histograms for the coherence.

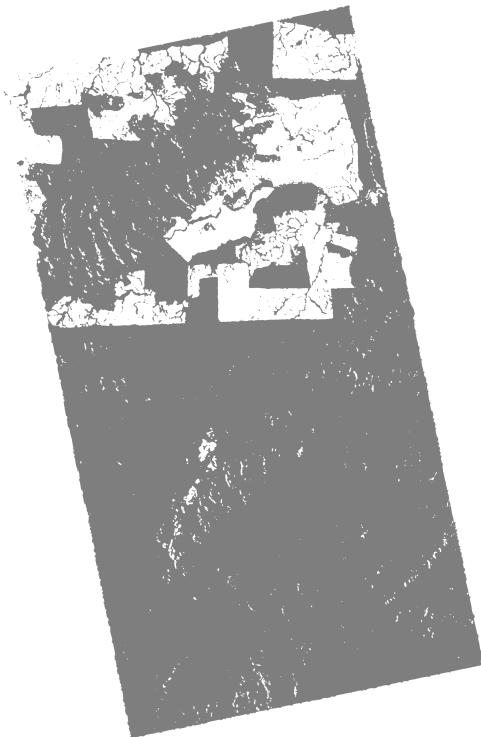
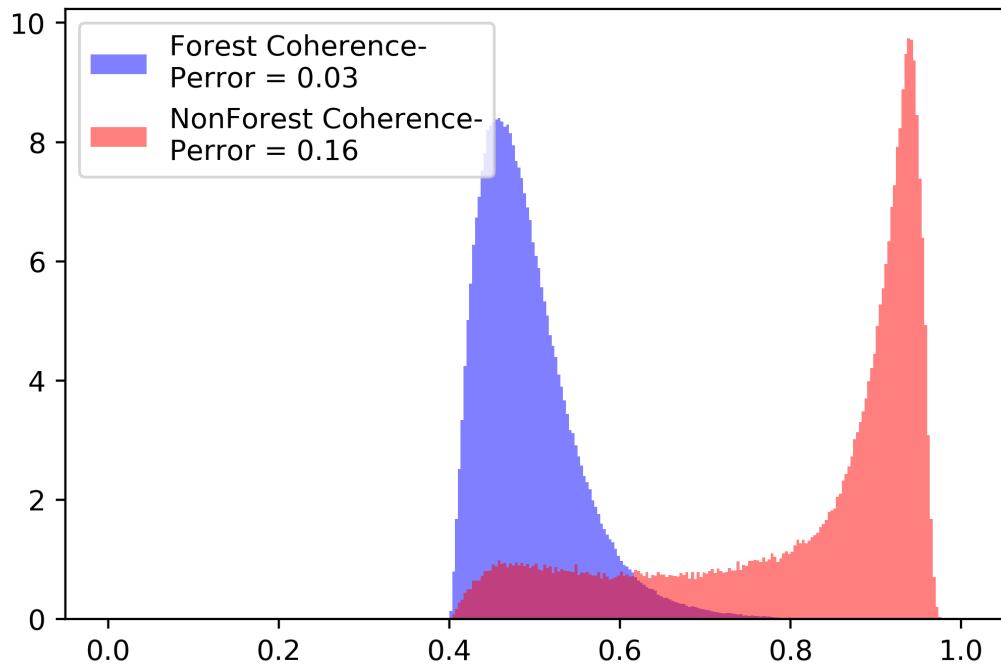


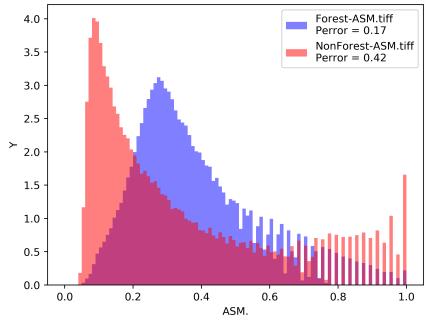
FIGURE 6.3 – Reference Map provided by DLR.

From the reference map above it was possible to extract the following probability density function from the histogram of the coherence.

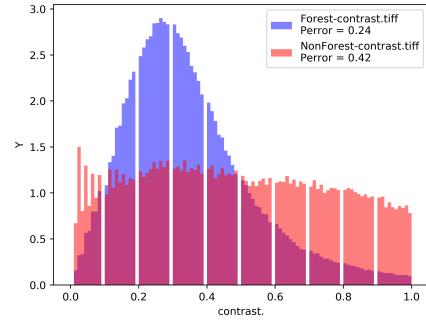


(a) Probability density Function for Coherence.

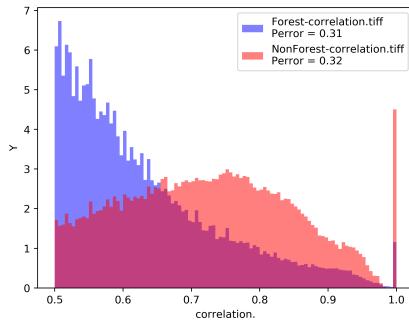
It is possible to see that the probability density functions (PDF) overlap at some point, which implies that there is always a probability of error when trying to classify the information of the pixel. On the image above it was also calculated which percentage of the area of one PDF is below the other PDF. For example, from the image above, 3 percent of the forest coherence PDF was below the Non Forest coherence PDF while 16 percent of the non forest coherence PDF was below the forest coherence PDF. We can see the PDFs obtained from the histograms of the texture features and see whether the PDFs are more separated compared to the PDFs of the coherence.



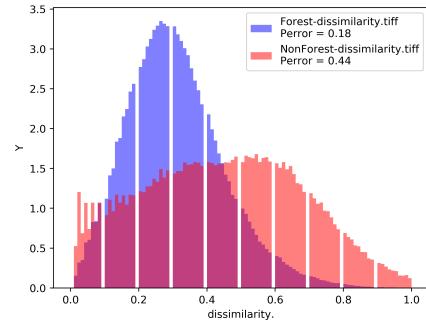
(a) Probability density Function for Angular Second Moment.



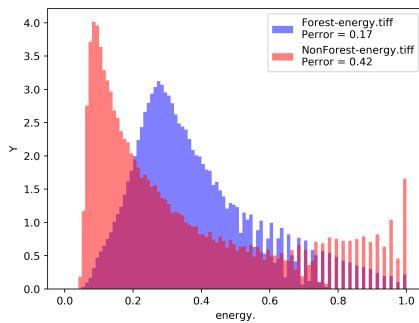
(b) Probability density Function for Contrast.



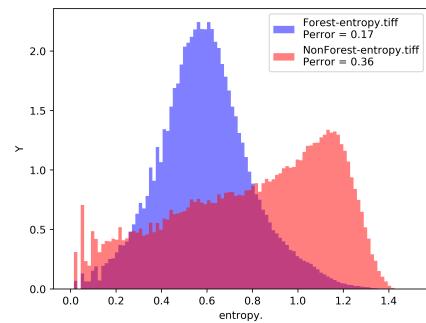
(c) Probability density Function for Correlation.



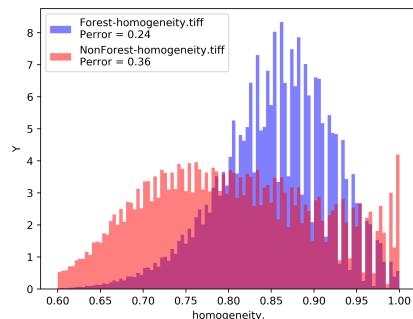
(d) Probability density Function for Dissimilarity.



(e) Probability density Function for Energy.



(f) Probability density Function for Entropy.

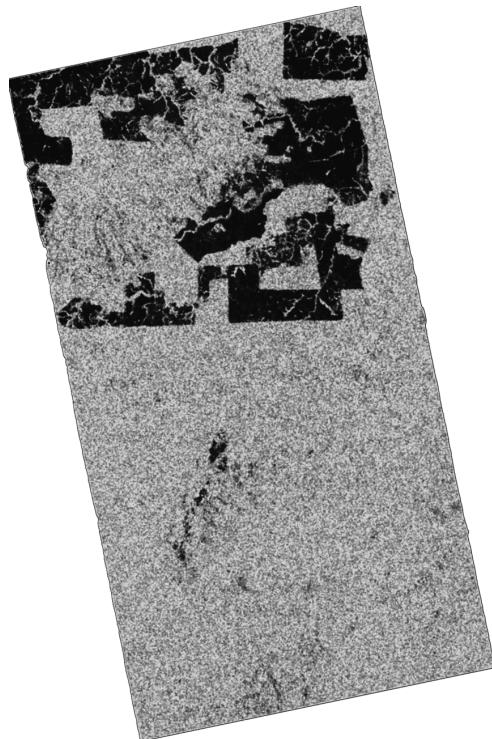


(g) Probability density Function for Homogeneity.

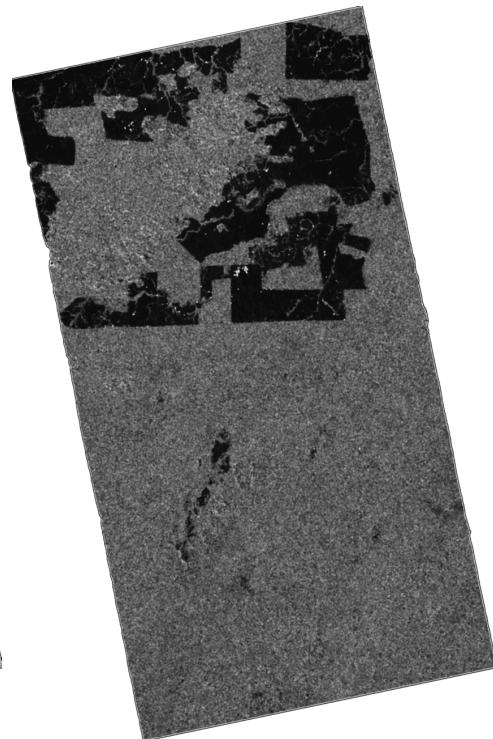
From the images above it is possible to see clearly that the PDFs of the textures are very different from the PDFs of the coherence, that means that the information contained in the textures is different from the information contained in the coherence, which implies that it is information that can be useful for image segmentation and classification, even if the intersection of PDFs is higher than in the coherence's PDF.

## 6.2 The Laws Textures results

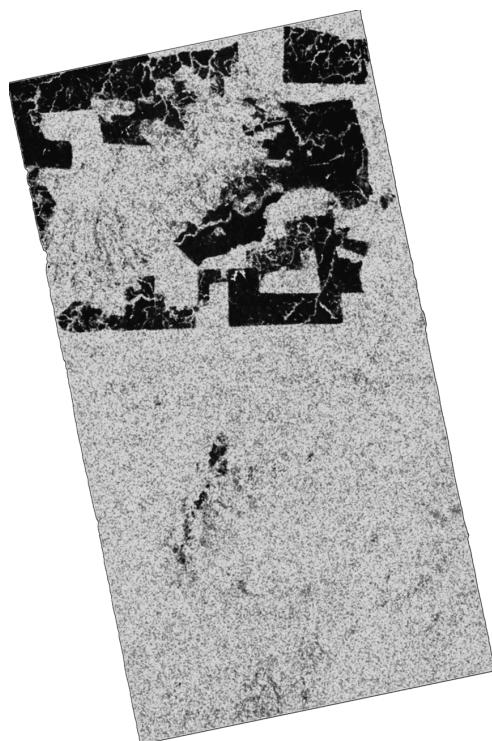
As mentioned in the previous chapter, the Laws textures are obtained by applying a series of linear filters to the image. Below there are images of the 9 possible texture results of these filters.



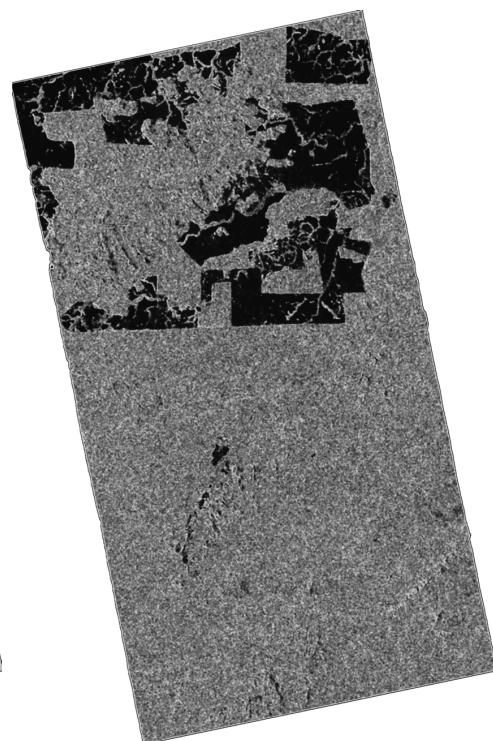
(a) Image of e5e5/e5e5 texture.



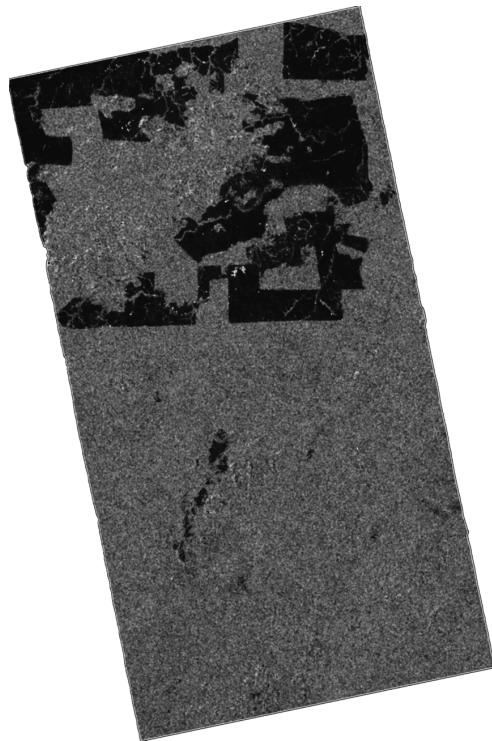
(b) Image of e5r5/r5e5 texture.



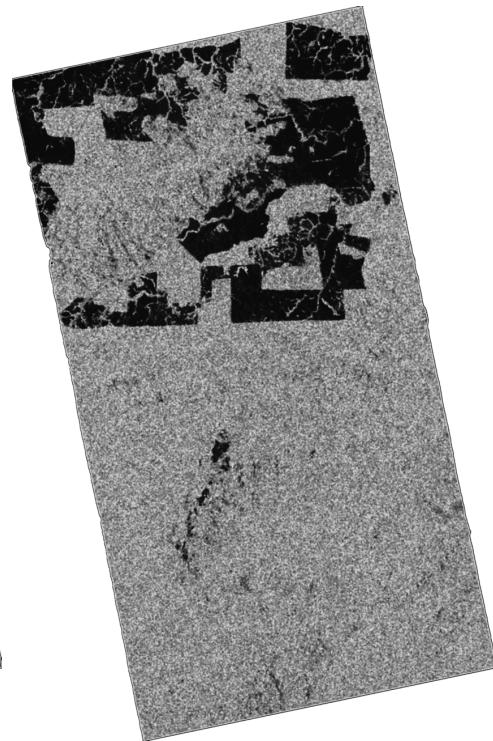
(c) Image of e5s5/s5e5 texture.



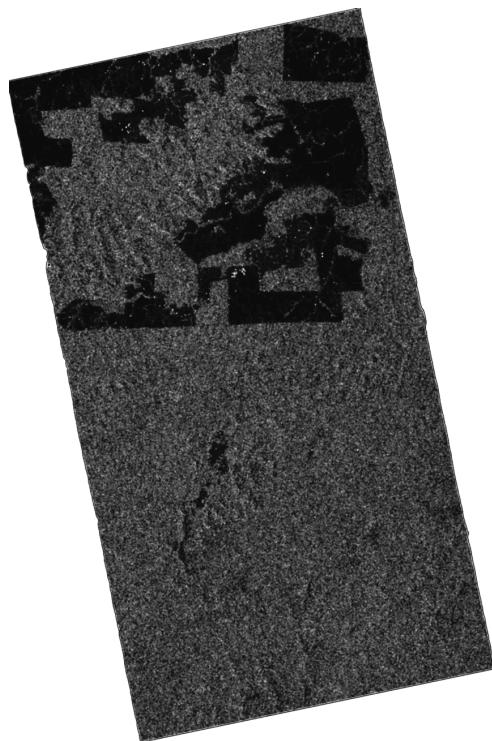
(d) Image of l5e5/e5l5 texture.



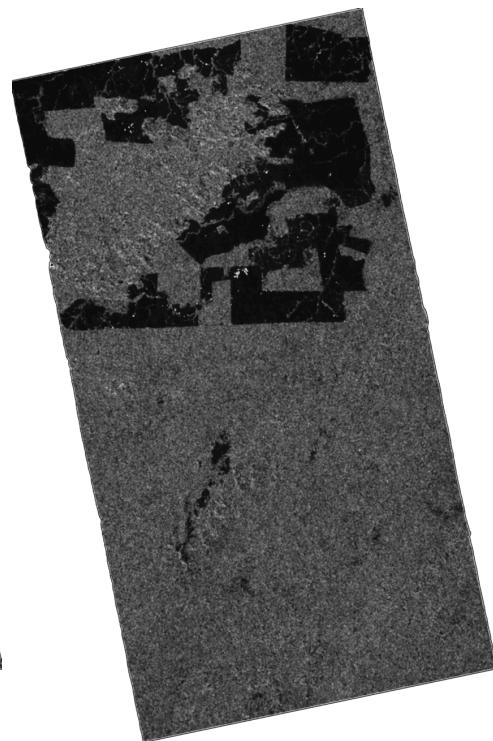
(e) Image of l5r5/r5l5 texture.



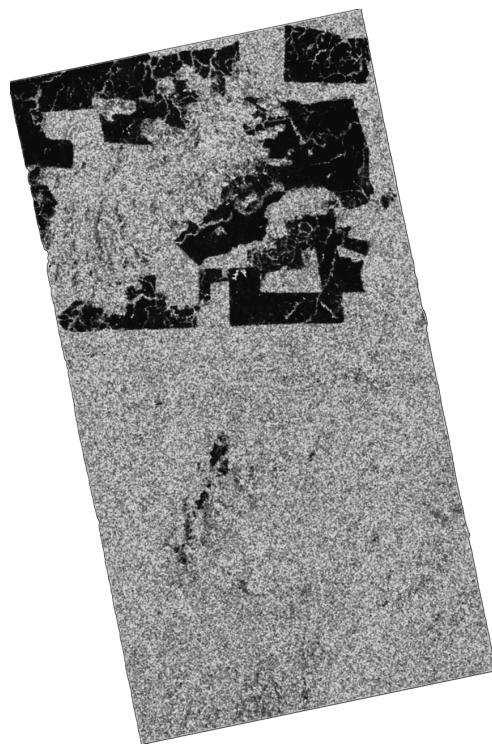
(f) Image of l5s5/s5l5 texture.



(g) Image of r5r5/r5r5 texture.

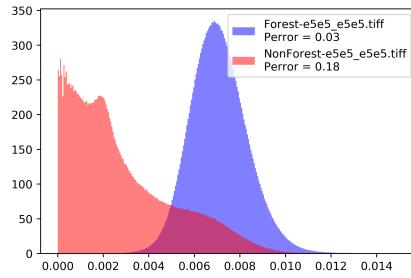
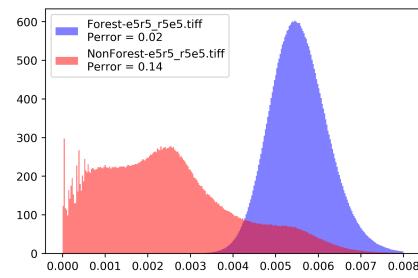
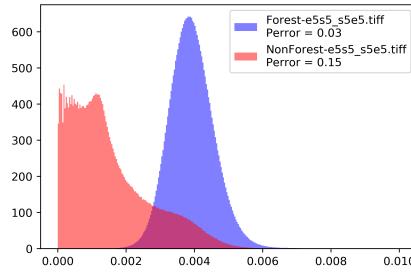
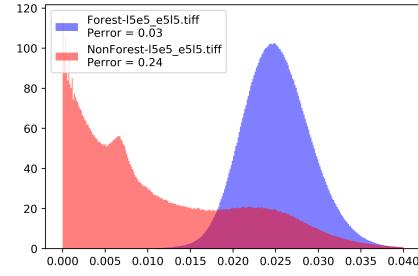
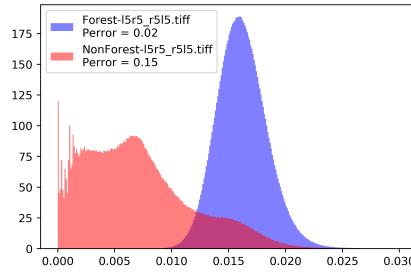
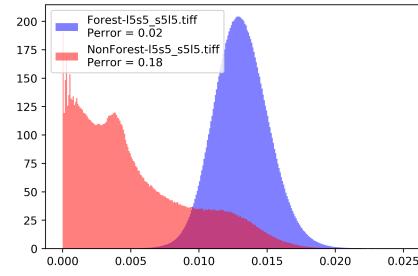
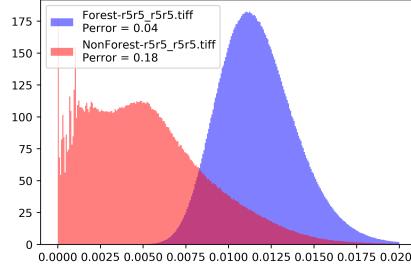
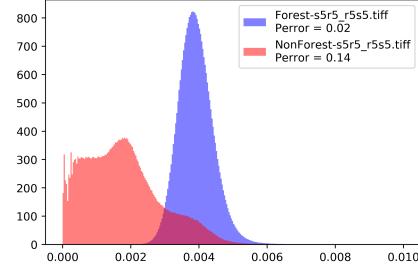


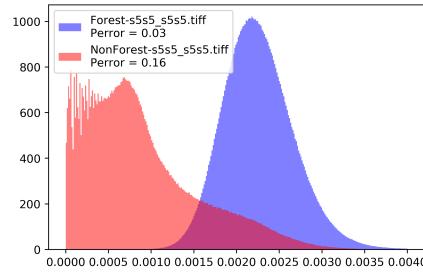
(h) Image of s5r5/r5s5 texture.



(i) Image of s5s5/s5s5 texture.

Again we must look at the PDFs for different classes and see whether they are different from the PDF of the coherence.

(a) Probability density Function for  $e5e5/e5e5$ .(b) Probability density Function for  $e5r5/r5e5$ .(c) Probability density Function for  $e5s5/s5e5$ .(d) Probability density Function for  $l5e5/e5l5$ .(e) Probability density Function for  $l5r5/r5l5$ .(f) Probability density Function for  $l5s5/s5l5$ .(g) Probability density Function for  $r5r5/r5r5$ .(h) Probability density Function for  $s5r5/r5s5$ .

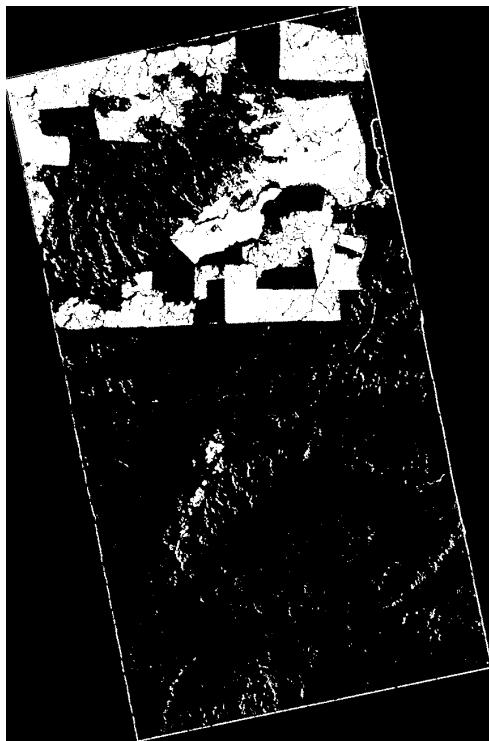


(a) Probability density Function for s5s5/s5s5 texture.

From the images above we can see clearly that the PDFs are more separated than the coherence PDF. An important thing that must be noticed is that all the different textures have very similar PDFs, which might be an indicator that the images are very similar between themselves, and that using more than one texture for classification means adding redundant information for the classification algorithms, which might degrade performance instead of increasing it.

### 6.3 Sum And Difference Histograms Textures

The last method analyzed for generating textures is the Sum and Difference Histogram Texture. Below there are the texture results for this method.



(a) Cluster Prominence Texture Image



(b) Cluster Shade Texture Image



(c) Contrast Texture Image



(d) Correlation Texture Image



(e) Energy.



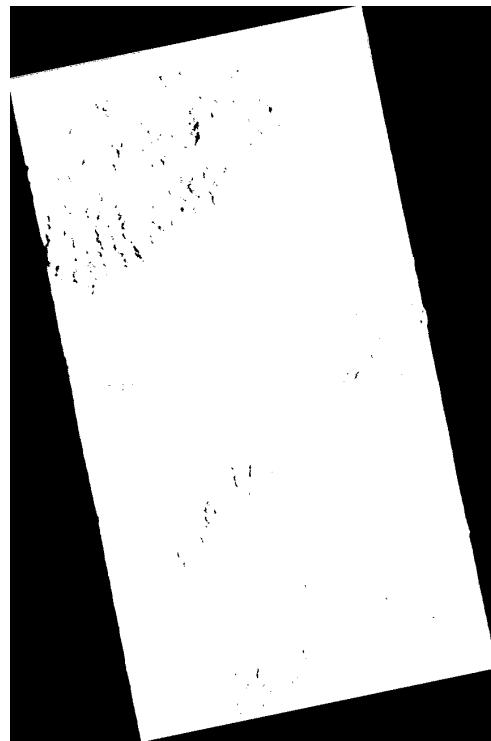
(f) Entropy Texture Image



(g) Homogeneity Texture Image

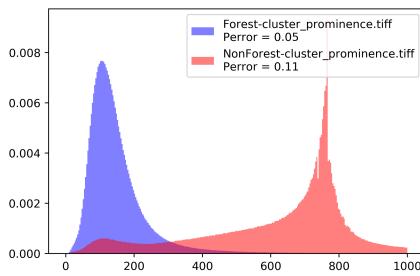


(h) Mean Texture Image

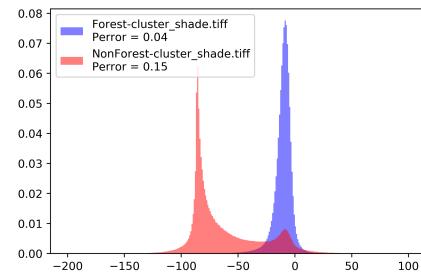


(i) Variance Texture Image

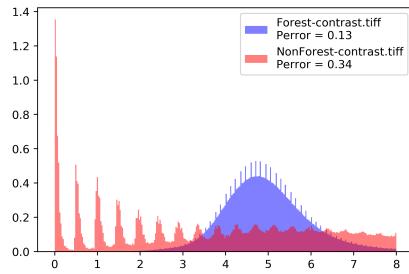
Again we must analyze the histograms to see if these features really are useful for classification and segmentation of a image.



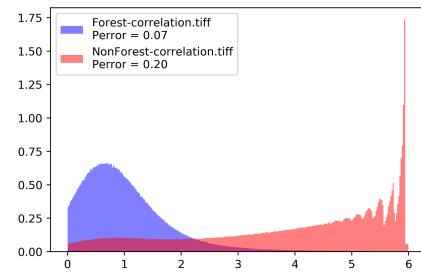
(a) Probability density Function for Cluster Prominence.



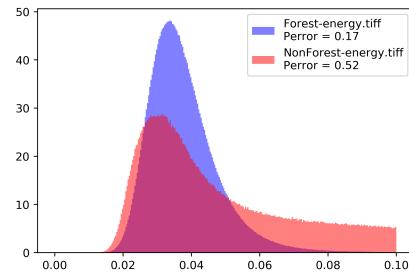
(b) Probability density Function for Cluster Shade.



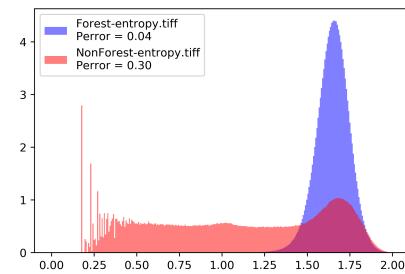
(c) Probability density Function for Contrast.



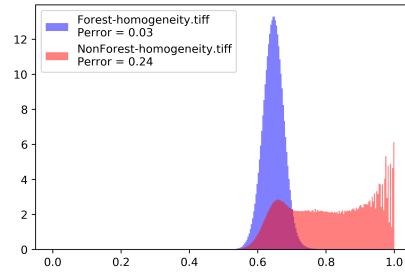
(d) Probability density Function for Correlation.



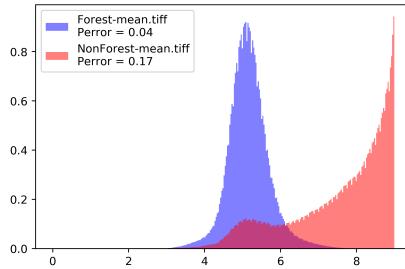
(e) Probability density Function for Energy.



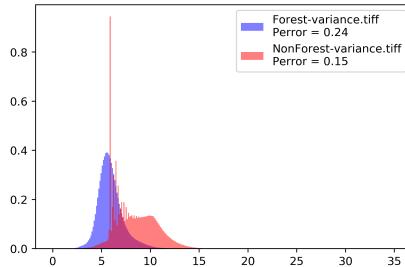
(f) Probability density Function for Entropy.



(g) Probability density Function for Homogeneity.



(h) Probability density Function for Mean.



(i) Probability density Function for Variance.

From the histograms and the texture images it is clear that some of the textures are excellently suited for classification, specially the Cluster Shade and Cluster Prominence textures which provide excellent separation between the probability density function. From those images it is clear that in the case of the TANDEM-X the sum and difference histogram methods for texture provided a significantly better result than the other textures, so through the rest of the work there will be a focus in using this method instead of the others, since due to computational limitations is not viable to use all texture methods for classification, since it takes a very long time to compute them.

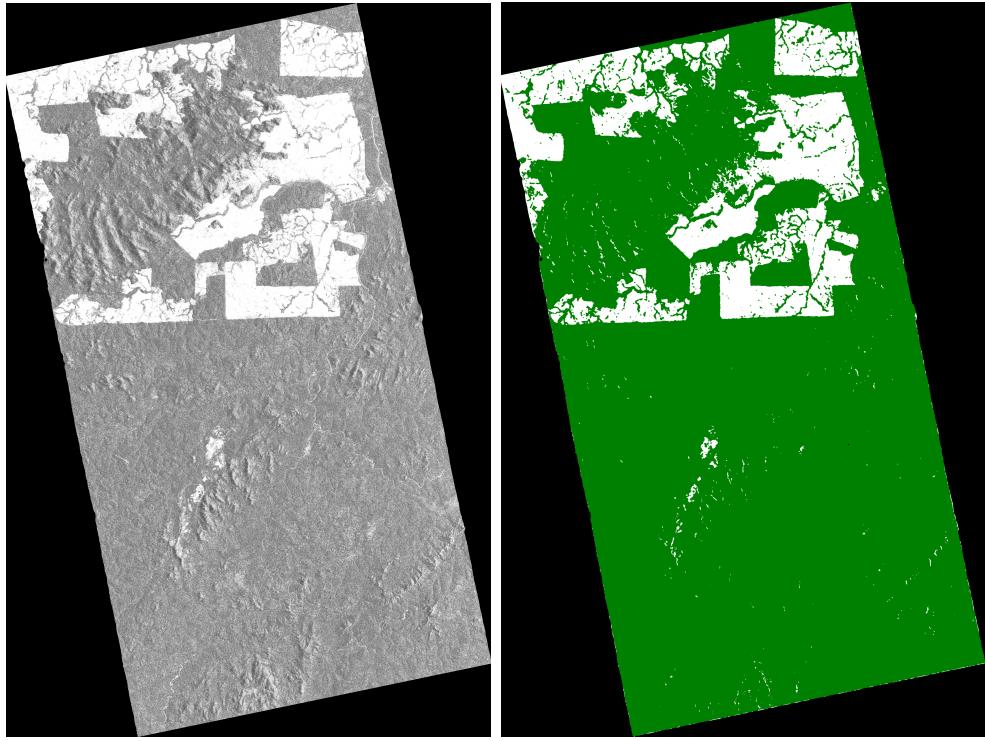
# 7 Classification Results

## Summary

In this chapter we will show the results of classification without the textures and compare it with the results of classification using the additional textures features. It will be shown the results for classification for forest and deforested areas in the Amazon Rain forest looking at acquisitions from the TANDEM-X satellite. Besides that, it will be also shown classification results using the acquisitions of the SENTINEL1 satellite for areas in the Amazon Rain forest, and compare the results with textures and without textures.

### 7.1 TANDEM-X Data Analysis

On this section it will be analyzed the result of textures for the volumetric correlation of a acquisition of the TANDEM-X satellite. The volumetric correlation image can be seen below. From this image, the textures were made using the sum and difference histogram method and it was given as a input to a Random Forest algorithm for classification. The random forest was trained using the reference map provided by DLR. Due to computational limitations not all textures were used, but just some of them. The textures choice was based on the PDFs of different classes. The textures chosen as input to the Random Forest algorithms were: Cluster Shade, Cluster Prominence, Contrast, Variance. Besides that, the coherence was also given as an input to the Random Forest Algorithm. The classification results with the textures can be seen below.



(a) Volumetric Correlation image of (b) Random Forest Results with Tex-  
TANDEM-X acquisition ture

FIGURE 7.1 – Volumetric Correlation and Classification

On the volumetric correlation image above the white pixels have higher correlation, which indicate the presence of a deforested area, while the dark pixels have a lower correlation, which is characteristic of a forested area. On the right it is possible to see the classification result given by the random forest. On the classification image on the right the white pixels indicate deforested area while the green pixels indicate the presence of a forested area.

It is clear from the figure 7.1 that the classification is accurate, but it is not possible to see how accurate the result is. Due to this it was also obtained the accuracy results when compared to the reference map provided by DLR . The accuracy results with different combinations of textures can be seen in the table below

Variance	46.78%
Coherence	4.70%
Variance/Coherence	4.70%
Cluster Prominence	3.82%
Cluster Prominence/Variance	4.17%
Cluster Prominence/Coherence	3.80%
Cluster Prominence/Variance/Coherence	4.70%
Cluster Shade	1.33%
Cluster Shade/Variance	1.30%
Cluster Shade/Coherence	1.30%
Cluster Shade/Variance/Coherence	1.30%
Cluster Shade/Cluster Prominence	1.27%
Cluster Shade/Cluster Prominence/Variance	1.26%
Cluster Shade/Cluster Prominence/Coherence	1.26%
Cluster Shade/Cluster Prominence/Variance/Coherence	1.28%
Contrast	21.13%
Contrast/Variance	21.07%
Contrast/Coherence	5.18%
Contrast/Variance/Coherence	4.70%
Contrast/Cluster Prominence	4.42%
Contrast/Cluster Prominence/Variance	3.81%
Contrast/Cluster Prominence/Coherence	3.47%
Contrast/Cluster Prominence/Variance/Coherence	3.47%
Contrast/Cluster Shade	1.30%
Contrast/Cluster Shade/Variance	1.30%
Contrast/Cluster Shade/Coherence	1.30%
Contrast/Cluster Shade/Variance/Coherence	1.40%
Contrast/Cluster Shade/Cluster Prominence	1.15%
Contrast/Cluster Shade/Cluster Prominence/Variance	1.52%
Contrast/Cluster Shade/Cluster Prominence/Coherence	1.47%
Contrast/Cluster Shade/Cluster Prominence/Variance/Coherence	1.29%

TABLE 7.1 – Error results with different textures combinations

From the table 7.1 we can quantitatively see how much was the improvement. Trying to run a classification using just the Coherence as an input gives a classification error of 4.70%, while if all textures are used then the error drops to 1.29%, in a way that the error 3 times smaller than the original error. Keep in mind that the classification map provided by DLR is not 100% accurate, so these numbers are not 100% accurate and

have a small error, which might explain why using just 2 textures, the cluster shade and cluster prominence, gives a smaller error (1.27%) than using the 4 textures together with the coherence (1.28%).

Even though this improvement might seem surprising it is important to notice that the coherence image used was a high quality image acquired with ideal conditions and with a height of ambiguity that yields a clear separation between the forest and deforested areas, but normally the SARs acquisitions are not always this good for classification, specially if it is not used a double satellite system like TANDEM-X (which has the advantage of not having temporal decorrelation between images since they are taken at the same time). On the next section this method will be used on acquisitions that are not so suited to make a classification in order to see the improvement that the textures can provide in non-ideal conditions.

## 7.2 Sentinel1 Data Analysis

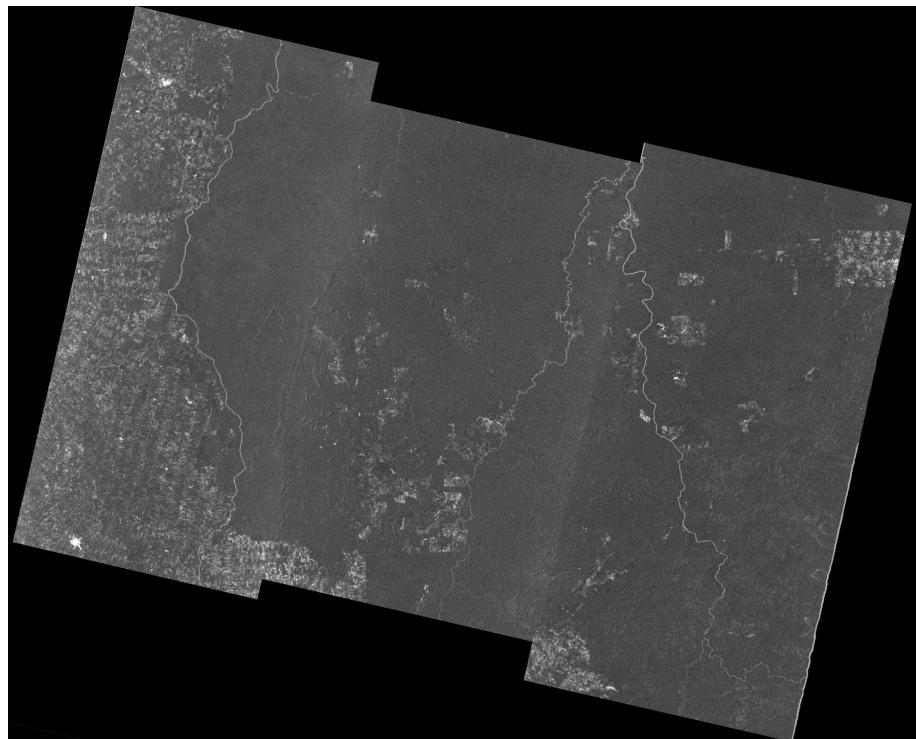
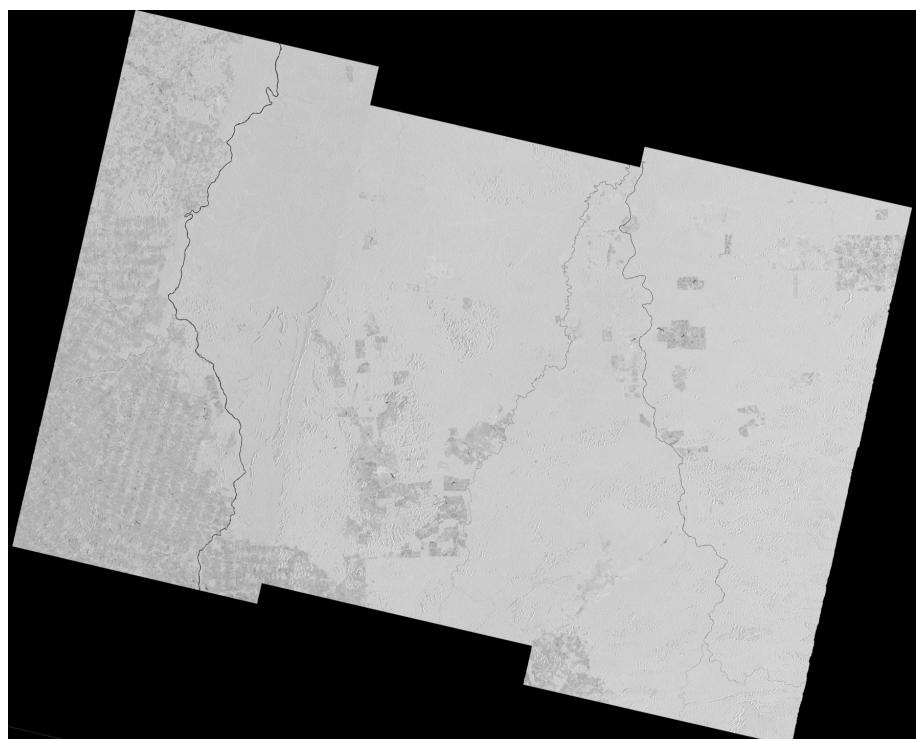
As said previously, the TANDEM-X images used for classification were taken in ideal conditions, in a way that making a classification on that image only is easy and even if no textures are used the classification results can provide a high accuracy.

The advantage of the TANDEM-X satellite is that it is twin satellite system that fly together taking pictures of the same area at the same time, in a way that the temporal decorrelation does not affect these images.

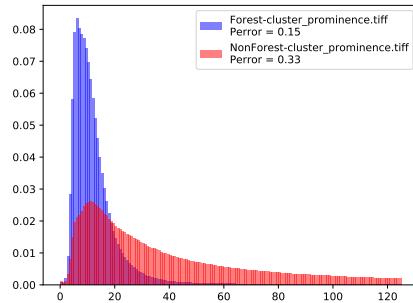
But normally, the pictures taken with SARs radars are taken at different days, in a way that the temporal changes on the scene strongly affects the image.

To illustrate this, on this section it will be shown coherence images of images of the satellite SENTINEL1, which have a 6 days temporal baseline between acquisitions. There were 5 images that were used to make the classification on this example, taken at the days 25/04/2019, 01/05/2019, 07/05/2019, 13/05/2019/ and 19/05/2019. Due to computational limitations it was chosen to make the mean between the images, since making the textures of each image might take a long time. The images have a 12m pixel resolution and cover an area of almost  $40000km^2$ , so making the textures of each image can take until 50 hours.

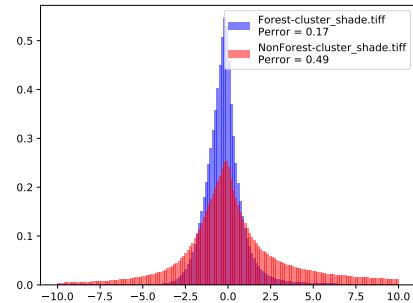
From the acquisitions it is possible to extract the brightness ( $\sigma^0$ ) of the scene and since there are 5 images it is possible to get 4 coherence images with a 6 days temporal baseline. Even though it is also possible to get the coherence at 12, 18 and 24 days it was chosen not to do so, since the images with such temporal baseline will provide little to no useful information for classification. Below there are images of the average  $\sigma^0$  and the average coherence  $\gamma$  of the acquisitions.

FIGURE 7.2 –  $\gamma$  acquisitionFIGURE 7.3 –  $\sigma^0$  acquisition

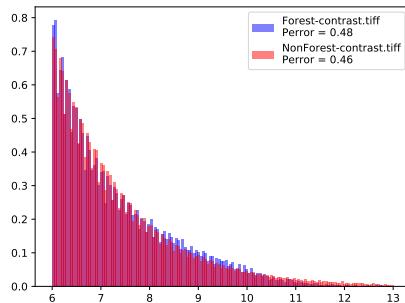
The respective histograms for the coherence and  $\sigma^0$  can be seen below.



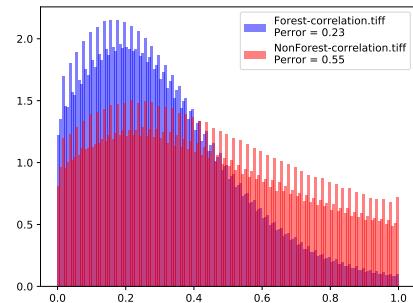
(a) Probability density Function for Cluster Prominence.



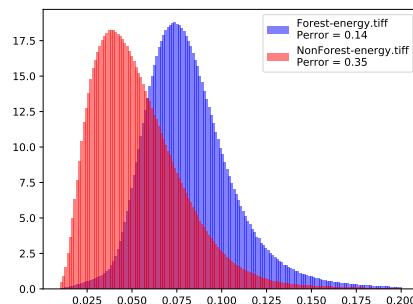
(b) Probability density Function for Cluster Shade.



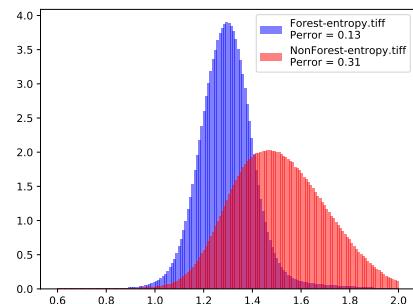
(c) Probability density Function for Contrast.



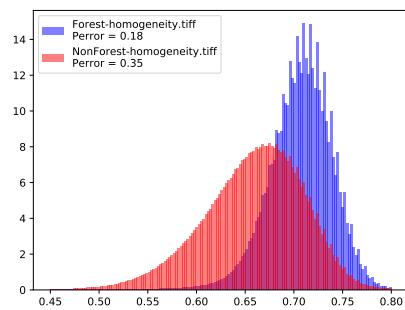
(d) Probability density Function for Correlation.



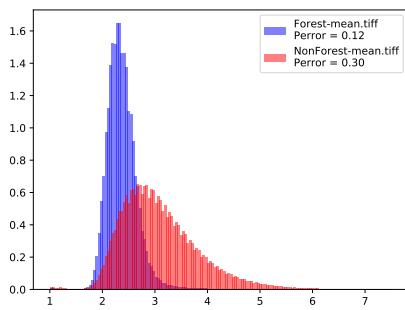
(e) Probability density Function for Energy.



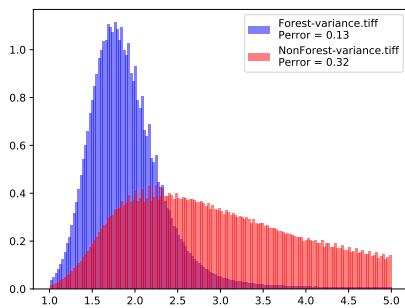
(f) Probability density Function for Entropy.



(g) Probability density Function for Homogeneity.

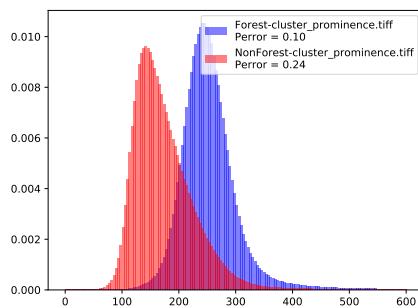


(h) Probability density Function for Mean.

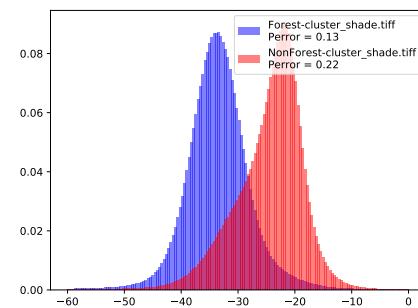


(i) Probability density Function for Variance.

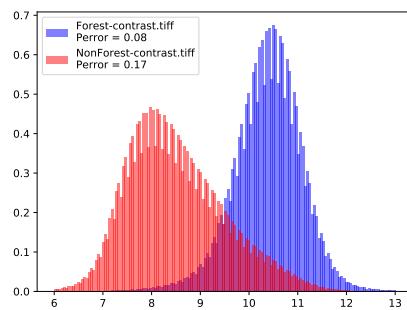
FIGURE 7.4 – Histograms of Coherence Textures



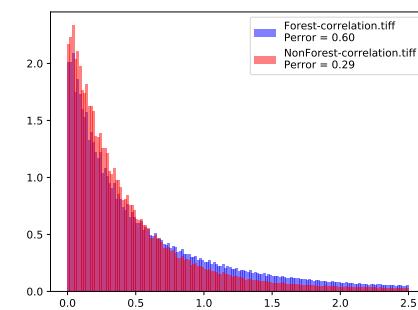
(a) Probability density Function for Cluster Prominence.



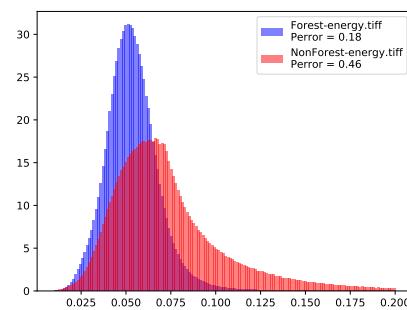
(b) Probability density Function for Cluster Shade.



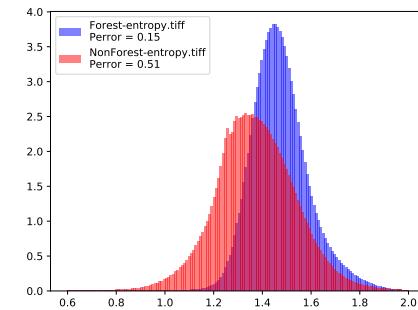
(c) Probability density Function for Contrast.



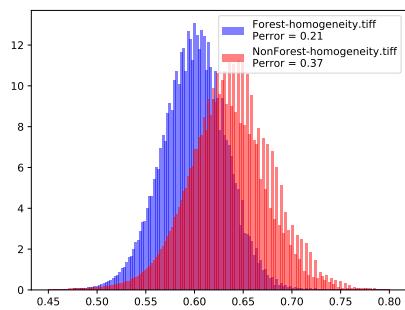
(d) Probability density Function for Correlation.



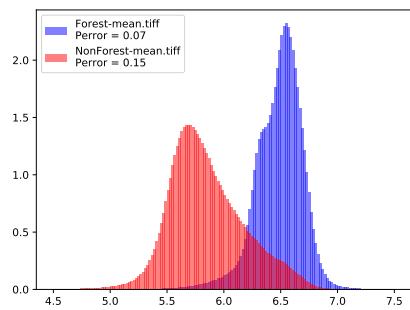
(e) Probability density Function for Energy.



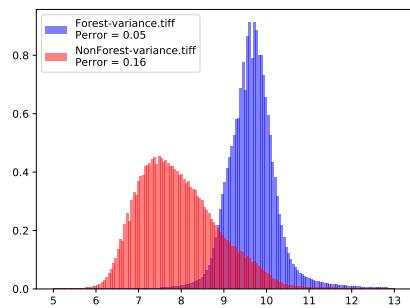
(f) Probability density Function for Entropy.



(g) Probability density Function for Homogeneity.



(h) Probability density Function for Mean.



(i) Probability density Function for Variance.

FIGURE 7.5 – Histograms of  $\sigma^0$  Textures

From the images and the histograms above it is clear that it is harder to make a classification, as the  $\sigma^0$  and  $\gamma$  are not as separated as TANDEM-X acquisition.

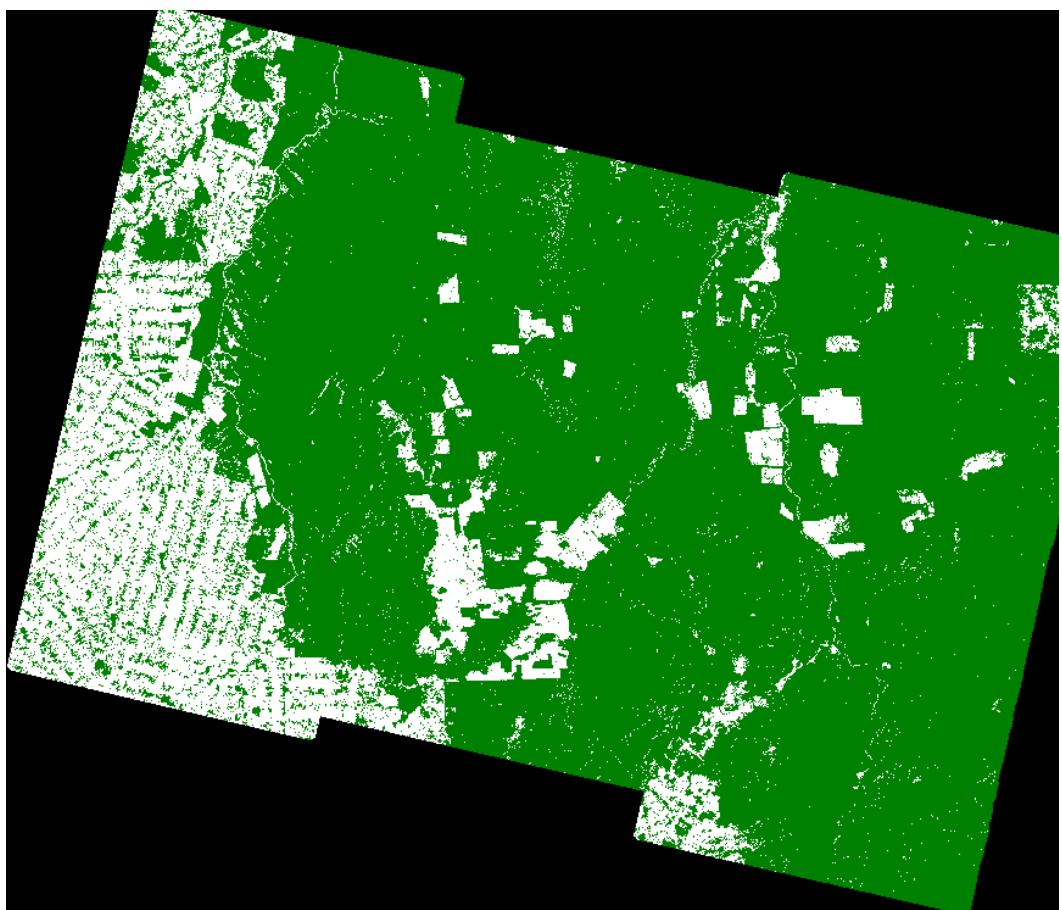


FIGURE 7.6 – Reference Map provided by DLR

Below it is possible to see the classification result using only the coherence and the  $\sigma^0$  as input to the random forest.

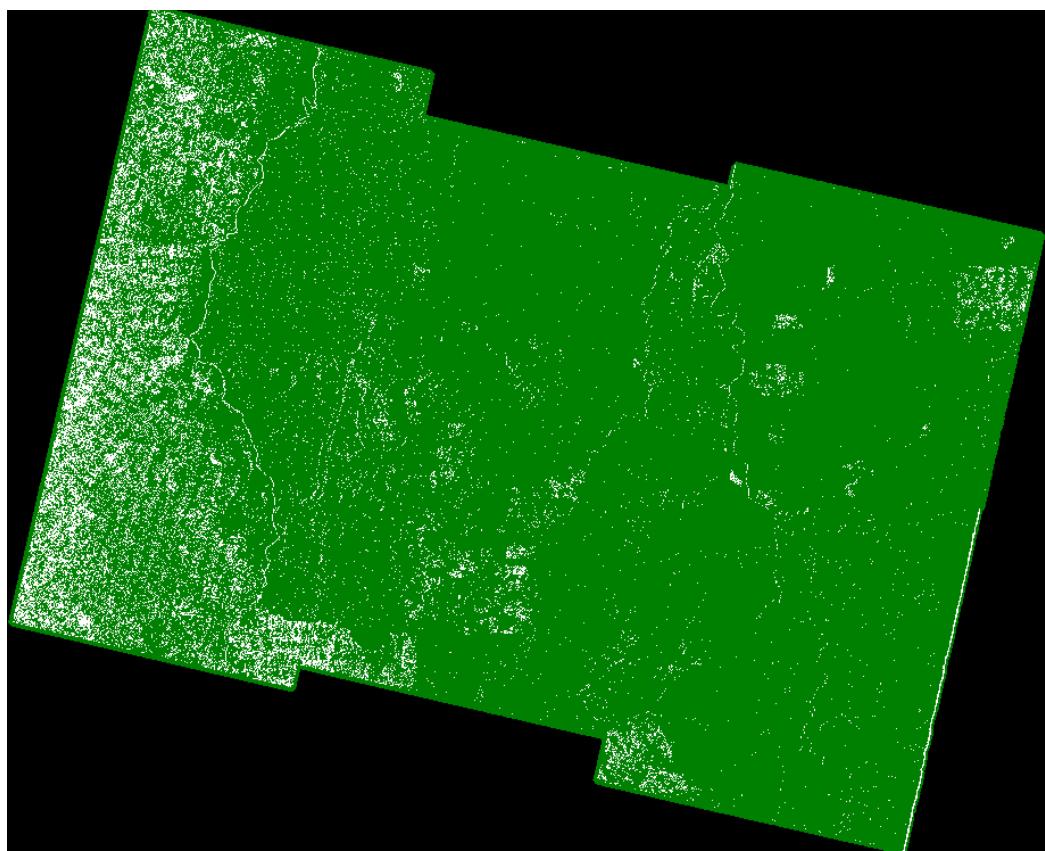


FIGURE 7.7 – Classification Results without textures

It is clear from that image the limitations of the machine learning algorithms for classification, since the classification clearly has a low accuracy. Below it is possible to see the random forest result using the textures.

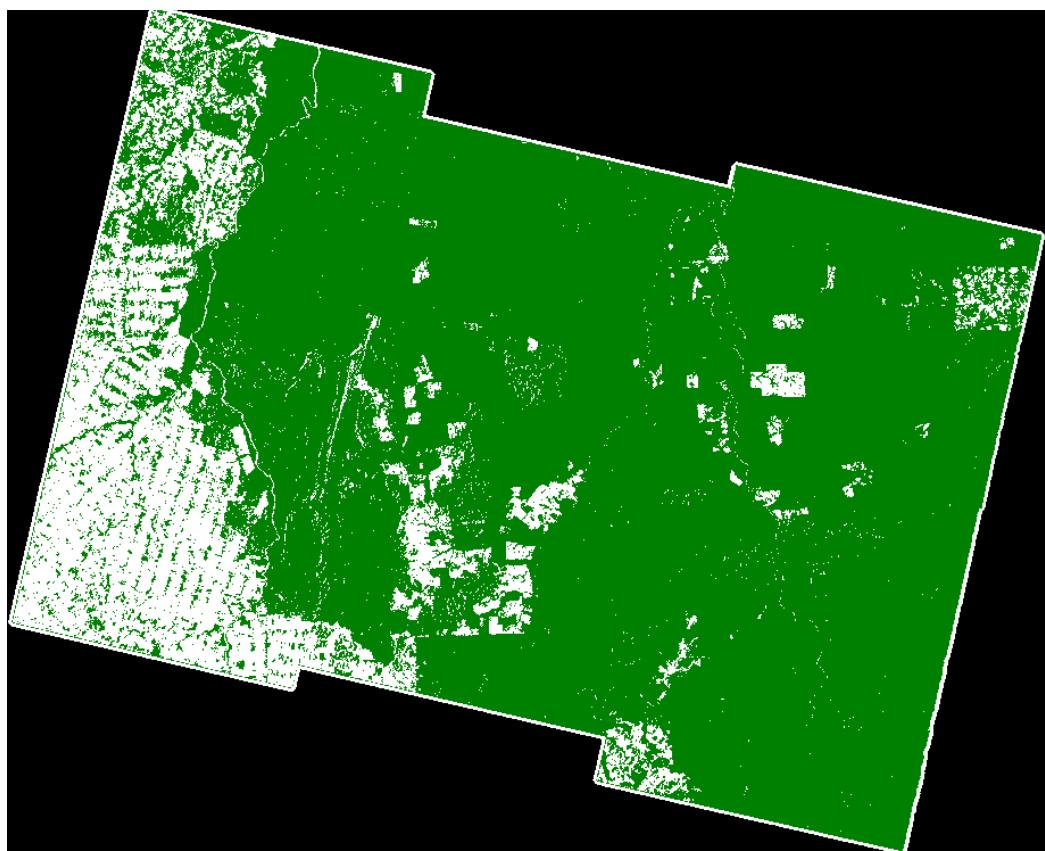


FIGURE 7.8 – Classification Results with textures

From that new image it is now clear how powerful the textures can be in aiding classification algorithms. For comparison, below it is also attached the optical image of the area acquired via Google Earth. The area acquired with Sentinel1 is show in the red square.

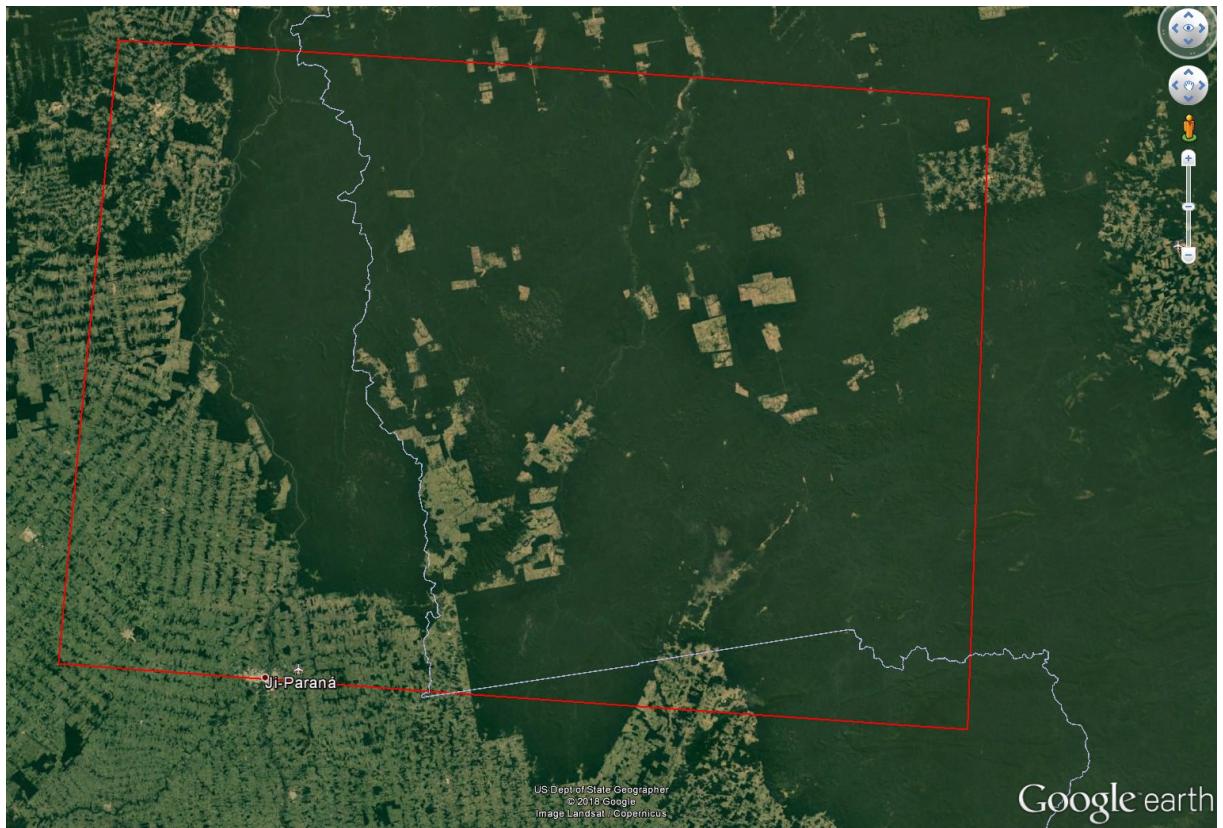


FIGURE 7.9 – Google Earth image of analysed area

### 7.3 Performance of Random forest and parameters tuning

There are 3 important parameters that have great impact on the performance of the Random Forest algorithm:

- Number of estimators: The number of trees in the forest.
- Maximum depth: The maximum depth of the tree
- Minimum samples for leaf: The minimum number of samples required to be at a leaf node. A split point at any depth will only be considered if it leaves at the minimum number of samples for leaf training samples in each of the left and right branches

Even though increasing those parameters can lead to greater accuracy, there are some breaking points in which the accuracy decreases, and there are some points in which the accuracy is stagnated and will not further increase, therefore only wasting computational time for no valuable return. For example, one might think that increasing the depth of the tree would yield a greater accuracy, but there is a limit in which the results will not improve, but it will take more time for the algorithm to make the training and the

classification. Below there area images that show how the accuracy of the classification algorithm changes with variations on the parameters:

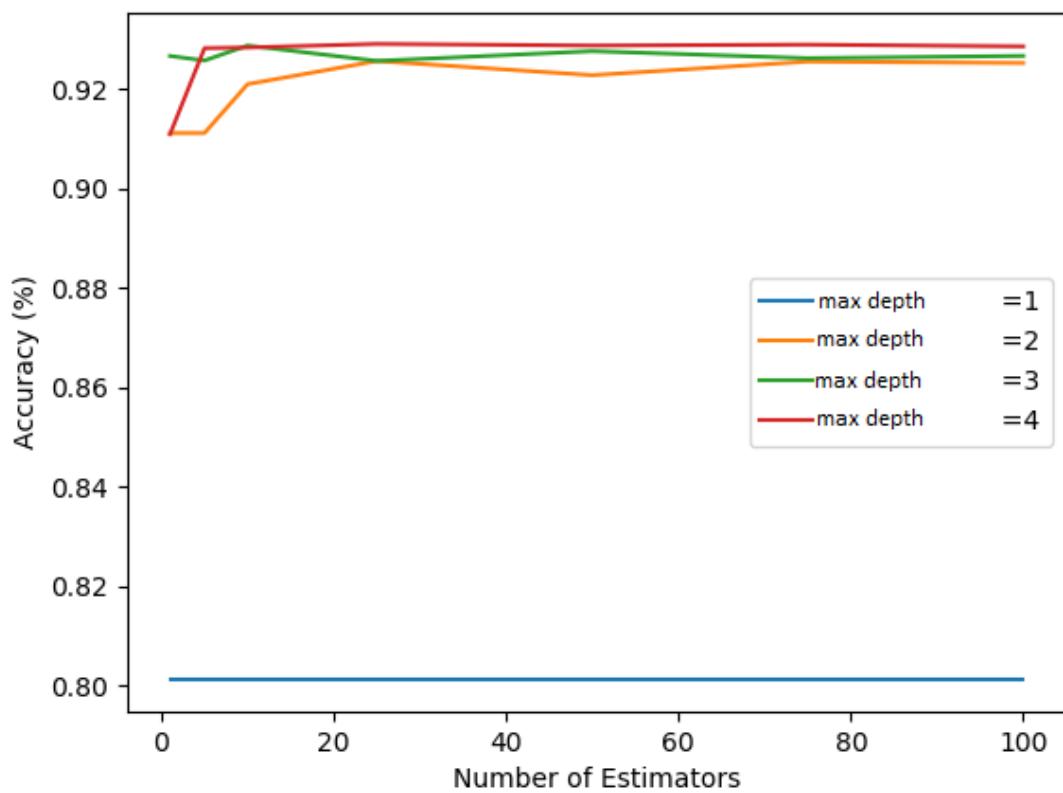


FIGURE 7.10 – Accuracy versus number of estimators for different tree depths

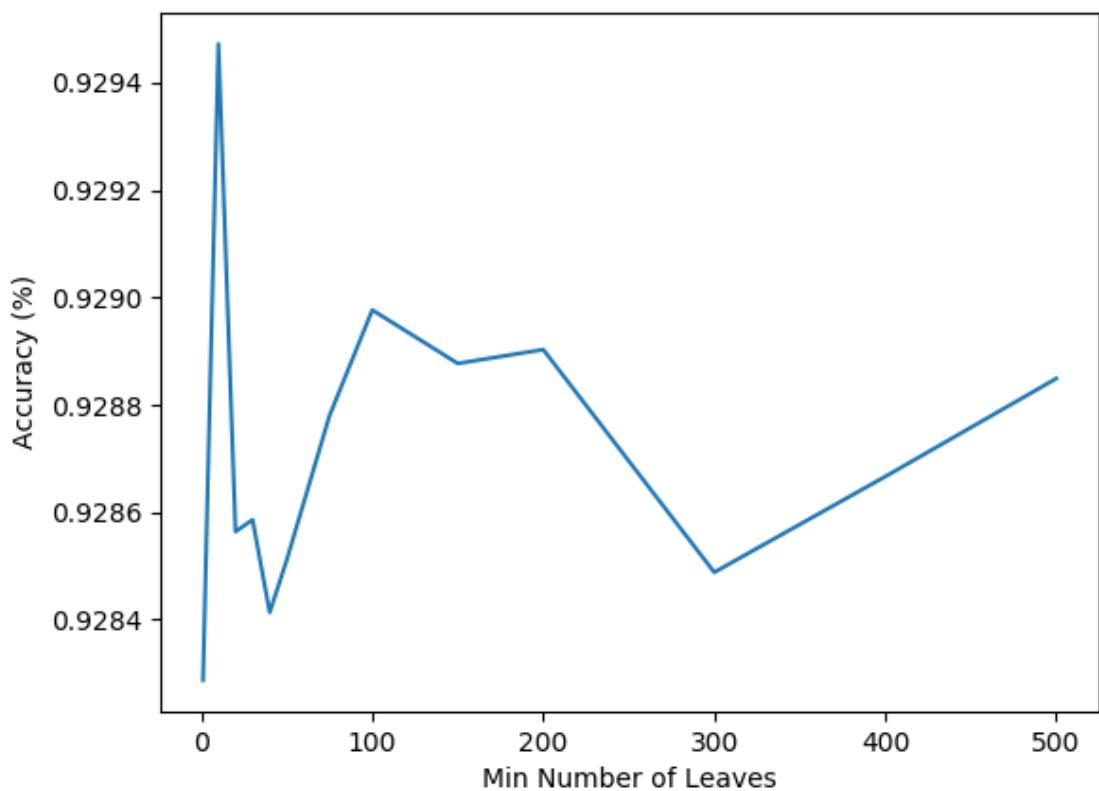


FIGURE 7.11 – Accuracy versus minimum number of samples for leaf parameter

## PART II

# 8 The CARABAS System and challenge problem

## 8.1 Summary

This chapter will present the change detection problem for targets concealed by a forest. It will also present the SAR system that was used to perform the data acquisition and the dataset that was created by this system. The dataset consists of SAR images collected that the Swedish low VHF-band SAR system CARABAS-II. The images were made publicly available by the Swedish Airforce Research center with the purpose of fostering research on the subject field of target detection under tree foliage.

## 8.2 Wavelength Resolution SAR Systems

Wavelength resolution SAR images are radar systems with resolution in the order of the radar system Wavelength. According to (ref), the resolution of a cell of a SAR image can be calculated by:

$$\delta = \frac{\lambda_c c}{4\theta_H B} \quad (8.1)$$

where  $\lambda_c$  is the wavelength corresponding to the radar central frequency,  $\theta_H$  is the aperture angle (in radians),  $c$  is the speed of light and  $B$  is the system bandwidth.

It is obvious that the backscatter characteristics of the SAR depends on the frequency band of the system. For wavelength-resolution SAR systems there is a significant difference between the backscatter of targets with size the order of the wavelength - which present the resonance scattering phenomena (ref63) - and targets small compared to the wavelength - which present the Rayleigh scattering phenomena (ref63).

According to (17ref), Wavelength SAR systems are not sensitive to small scatterers inside the resolution cell, thus the scattering process is mainly due to scatterers with

dimensions in the order of the system wavelength. Since the resolution of the cell is similar to the resolution of the scatterer, there can be only one single scatterer inside each cell, therefore the image will not be greatly affected by speckle noise (speckle noise is a major cause of problems in non wavelength-resolution SAR still). For example, according to (ref64) in frequencies below 100 MHz, in forest areas, the backscatter is dominated by backscattering from tree foliage.

Moreover, scatterers that are large tend to be static objects, hence a stack of images of the same area will present a high degree of similarity between multi-pass acquisitions. Therefore regular noise reduction techniques do not have to be applied when dealing with those set of images (ref 61).

It is also worth mentioning that according to (ref 66), signal attenuation in low-frequency wavelength-resolution SAR systems is not a major concern. Studies (ref63) have also shown that signal attenuation can be as little as 3dB in low-frequency SAR acquisitions in forest areas.

Due to all that, according to (63ref), low resolution SAR systems are better suited to Foliage Penetrating (FOPEN) applications. Also according to (63ref), VHF-band is the optimal radar system for FOPEN applications for vehicle-sized target detection.

### 8.3 The CARABAS-II Challenge Problem

There are many situations where there is a need for monitoring vehicles concealed by foliage over a large area. For example, for military applications, it might be necessary to monitor the positioning of enemy vehicles trespassing a restricted area. For environmental purposes it might be needed to pinpoint the position of vehicles that are being used for illegal deforestation over protected forests. If SARs are to be used for such purposes, it is ideal that low frequency systems are chosen since those provide a wide surveillance area and good foliage penetration capabilities. At low frequency (VHF-band) the main backscatter from the target area is due to large scatterers, e.g. tree canopy, houses, vehicles, and other man-made objects, which will appear as very bright objects in the image.

The challenge related to change detection is associated with the trade-off between having high accuracy detection and high false alarm rate. Normally what happens is by trying to increase the overall accuracy of detection algorithms will incur in having a high number of false alarms, therefore when designing an algorithm for detection it is necessary to have both good detection capabilities and a low enough false alarm rate to be used by the client. In foliage penetration applications, the main source of clutter comes from large tree trunks, and the more sparse the forest is the less will be the number of false targets.

It is also true that the larger the tree then the higher will be the number of false alarms (ref paper).

The objects that will be used to assess the quality of the change detection method will be a set of 25 military vehicles distributed among a testing area. Since those objects are large (compared to the bandwidth of the signal), and stationary, then their radar signature will be very stable between acquisitions. This will be used as an advantage by taking acquisitions with different flight passes to suppress clutter noise from tree canopy.

According to (ref 1/2 challenge) VHF-band has good performance for detecting targets under trees, but there are very few VHF SAR systems in the world. To overcome this problem researchers from the FOI released a VHF-band SAR image dataset to the public to foster the development of change detection methods for wavelength-resolution SAR systems in VHF band, which is the dataset that will be used in this work.

## 8.4 The CARABAS-II System

The CARABAS-II is the second generation SAR system designed by the Swedish Defense Research Agency (FOI) for FOPEN applications. The CARABAS-II has participated in numerous military campaigns and has been used for research purpose since the nineties. The radar is a VLF UHF SAR system that transmits HH-polarized radio waves in the frequency range of 20-90 MHz, therefore having resolution in the range of 3.3-15 m. The radar antenna is mounted on a Sabreliner aircraft as seen in figure ??.



FIGURE 8.1 – The CARABAS-II VHF SAR mounted in front of a Sabreliner airplane

In the table below the system parameters for the CARABAS-II system are presented for the flight campaigns that were used in this work.

System Parameters	Values
Nominal flight altitude	3 - 9 km
Nominal flight speed	127 m/s
Frequency band	20-86 MHz
Aperture angle	90 degrees
Transmitted power	500 W
Pulse modulation	Non-linear frequency modulation
Radio Frequency Interference (RFI) sniff	On
Frequency sub-bands	35 (36 with RFI-sniff)
Frequency step	1.875 Mhz
Center frequencies	21.25-85Mhz
Pulse repetition	frequency 5000Hz
Pulse length	15 $\mu$ s
Maximum range	26.4 km

TABLE 8.1 – CARABAS-II SAR system parameters. Source: (76 ref)

## 8.5 The CARABAS-II Dataset

By trying to promote research on change detection algorithms for wavelength-resolution images, FOI created a dataset of images acquired by CARABAS-II and made it publicly available. This dataset is the one used to test and assess the quality of the proposed CDA.

The dataset consists of 24 SAR images selected from over 150 images obtained during different flight campaigns. Each image covers the same ground area of  $6 \text{ km}^2$  (3 km vertically and 2 km horizontally) and is given in the form of a 3000 X 2000 matrix, where each pixel size is 1km x 1km. According to (ref 77,62,78) the images are already calibrated, pre-processed and geocoded.

The location of the image dataset is inside the military base station Missile Test Area North Vidsel in northern Sweden in 2002. The test site is a region near the village of Nausta (ref 75). The vegetation of the area is dominated by Scots Pine tree (ref 76), which consists of small and medium size trees. According to (ref75) the area also contains fields, roads and lakes.

With the objective of testing CDA quality, it was deployed 25 testing targets over the testing with different configuration and arrangements. The testing targets consists of ten TGB11 model military vehicles, eight TGB30 model military vehicles, and seven TGGB40 model military vehicles. The dimensions of each vehicle are presented in the table (numero). From the table (numero) it can be seen that the dimensions of the vehicles are similar to the wavelength of the CARABAS-II system, therefore all advantages of targets with similar dimension to the wavelength previously mentioned hold true for the dataset.

The dataset of 24 images were acquired using four different flight passes. Each flight

Military Vehicle	Lenght	Width	Height	Quantity
TGB11	4.4 m	1.9 m	2.2 m	10
TGB30	6.8m	2.5m	3.0 m	8
TGB40	7.8m	2.5m	3.0m	7

TABLE 8.2 – Target dimensions

pass has an incidence angle of 58 degrees, used the StripSAR mode and was acquired with the radar looking left (ref 75,76). The vehicles were positioned in two different forests, Forest 2 being in the northwest of the field, and forest 1 being in the southeast of the test area.



FIGURE 8.2 – Military vehicles used as target. Left:TGB11. Middle:TGB30. Right:TGB40. This picture also depicts the vegetation characteristics of the test area

The vehicles in mission 2 are positioned in forest 2 and have a heading angle of 225 degrees pointing southwest direction; vehicles in mission 3 are positioned in forest 2 and have a heading angle of 315 degrees pointing northwest direction; vehicles in mission 4 are located in forest 1 and have the same heading angle of mission 2; vehicles in mission 5 are located in forest 1 and have a heading angle of 270 degrees pointing west direction. Vehicles in mission 2 and 3 are separated approximately by 50m, as such for vehicles in mission 4 and 5. Figure 8.3 presents images of each mission with the vehicles area highlighted in red.

In table 8.3 it is presented the summary with the information of each image in the dataset

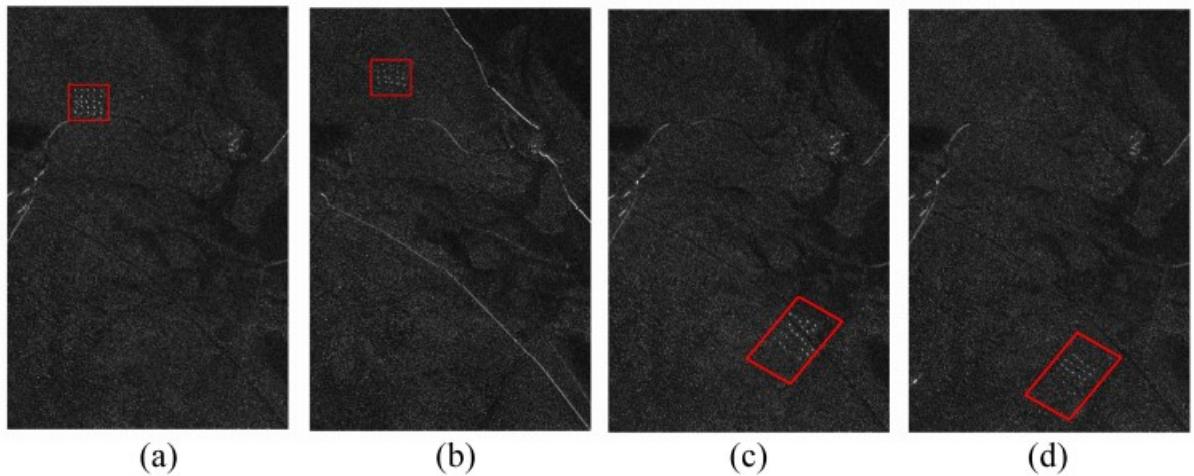


FIGURE 8.3 – Examples of CARABAS-II images where (a) is mission 2, (b) is mission 3, (c) is mission 4 and (d) is mission 5

Image Number	Mission	Pass	Flight Heading (degrees)	Target Heading
1	2	1	225	225
2	2	2	135	225
3	2	3	225	225
4	2	4	135	225
5	2	5	230	225
6	2	6	230	225
7	3	1	225	315
8	3	2	135	315
9	3	3	225	315
10	3	4	135	315
11	3	4	135	315
12	3	6	230	315
13	4	1	225	225
14	4	2	135	225
15	4	3	225	225
16	4	4	135	225
17	4	5	230	225
18	4	6	230	225
19	5	1	225	270
20	5	2	135	270
21	5	3	225	270
22	5	4	135	270
23	5	5	230	270
24	5	6	230	270

TABLE 8.3 – Measurements parameters for each image

## 8.6 Traditional Change Detection in Wavelength-Resolution SAR Images

The field of FOPEN target detection using wavelength-resolution SAR images (both in VHF and UHF bands) has been around of decades (ref 25,48,79,80). Among those studies, the CARABAS-II system dataset was used in many different FOPEN studies (ref).

As previously mentioned, the objective of the study is to create a change detection method that will compare different images of the same target field (in this case, the forests in Sweden) and try to identify the targets that have changed position in the image (in this case, the military vehicles TGB11, TGB30 and TGB40 ) and give the exact location of the vehicle positions.

There are several change detection methods that were already used in the CARABAS-II dataset. The first method that yielded accuracy high enough for real world applications was based in Bayes linear classification (ref81). After that, a myriad of new methods were proposed, such as: Likelihood-ratio test (LRT) (ref 61), combination of LRT and Space-Time adaptative processing (STAP) (ref 76) among others. Traditionally CDAs were also mainly based on traditional statistical decision theory, e.g., traditional hypothesis testing criterion methods, such as maximum a posteriori criterion (ref), likelihood ratio test(ref), generalized likelihood ratio test (ref), or Bayesian theory approaches (ref). Several of those CDAs have achieved high accuracy in terms of true positives, but most show unsatisfactory performance in terms of false positives percentage (ref).

# 9 Textures Applications for Change Detection Algorithms

## 9.1 Summary

In this chapter we will use the textural developed in previous chapters to create a change detection algorithm (CDA). The CDA will be assessed using the CARABAS-II image dataset, which is a set of images acquired by a Very High Frequency (VHF) Ultra Wideband (UWB) SAR system. The CDA is based on the UNET convolutional neural network (CNN) architecture and will use the textural information as additional inputs for image classification. It will be demonstrated that the use of textural information can improve the overall performance of the algorithm in terms of probability of detection and false alarm rate.

## 9.2 Introduction

As previously mentioned, SAR has several applications, such as, topology, oceanography, geology, environment monitoring, among others. The SAR can operate in two modes: single-pass acquisition and multi-pass acquisition. In the first it is captured only one picture of the monitored area, while in the latter there can be multiple acquisitions of the same area. When dealing with multi-temporal acquisitions, we can either assess the information contained in the entire time frame or focus in detecting changes between the acquisitions. When the latter is the focus, change detection algorithms (CDA) are traditionally used to identify the relevant changes (ref carabas, ricardo).

Generally, CDA's priority is the percentage of true positives with a lower regard for percentage of false positives. Those can however be relevant in many applications such as detecting hidden objects over a large area, where a small number of false positives is vital in the quality of the final data for operation application - which is the main concern of this work. While trying to develop a CDA for this work, it was kept in mind that the

CDA developed would have to provide a high percentage of true positives while trying to minimize the amount of false positives.

The traditional method for creating CDAs was to base the decisor on standard statistical theory, e.g., traditional hypothesis testing criterion methods, such as generalized likelihood ratio test refGLRT1,GLRT2,GLRT3, maximum a posteriori criterion ref Book May, Bayesian theory approaches refBayes1, Bayes2, or likelihood ratio test refLRT1,LRT2,LRT3. Even though these methods, most of the time, perform adequately in terms of probability of detection, but most of them show poor performance in terms of false alarm rate (refCarabas, Ricardo, LucasRamos, Chris). Those methods also have a few drawbacks, such as: depending on additional preprocessing of the images throughout morphological operations (erosion and dilation filters have to normally be used) and depend on calibration of the algorithm for the statistical analysis to work.

By trying to overcome the limitations of traditional statistical based CDAs, machine learning algorithms became popular in the last few years (ref Vinholi, Campos) since they have the capability to perform better than statistical based CDAs and don't have the usual drawbacks that those methods present. There are two kinds of machine learning algorithms: supervised and unsupervised algorithms. Supervised algorithms are the ones that have a classification reference to help guide the learning step, while unsupervised algorithms are the ones that do not have this classification reference - on this work we focused only in supervised algorithms. As it was mentioned, it is important to know what is the parameter of interest to quantify the quality of the algorithm, e.g., we can assess performance based on number of true positives, true negatives, false positives and false negatives (ref PefMe) - for this part of the work, the parameters of interest are number of true positives and number of false negatives.

Among supervised machine learning algorithms there are multiple algorithms that can be chosen to solve the problem, such as: Neural Networks, Naive Bayes, Linear Regression, Kernel Ridge Regression, Support Vector Machine, Decision Trees, Random Forest (ref PefMe). For this work it was chosen to use a neural network based algorithm for the CDA. There are multiple types of neural networks, e.g., perceptron, Feed Forward Neural Network, Multilayer Perceptron, Recurrent Neural Network, Long Short-Term Memory Neural Network, convolutional neural network, among others (ref PefMe). It was chosen to use a convolutional neural network, specifically a UNET convolutional neural network (Ref unet), for solving the target detection problem, since convolutional neural network have shown excellent results for computer vision problems. In (ref Kevin) it was used a similar approach to detect change in SAR images, although in (ref Kevin) the SAR images were acquired in a much higher frequency.

The UNET CNN will not only have the SAR acquisitions as input for the target detection, but will also have textural information that will be extracted according to the

methods explained in chapter 3. In (ref rodrigo) it was already shown that additional textural information can provide valuable information to improve classification results. It is important to mention that since CNN use convolutions, which are linear filters, to achieve better classification results, it would be redundant to use linear textural information as inputs to the UNET CNN. When comparing the proposed technique with other algorithms on the same dataset, we show that the results show similar detection performance but presenting a significant false alarm rate improvement.

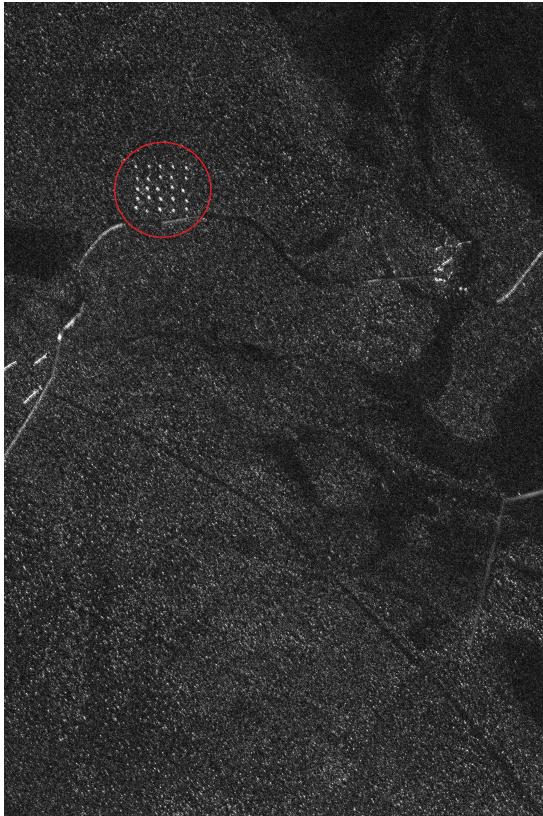


FIGURE 9.1 – Example of a SAR image acquired by the CARABAS-II system. The red circle shows the position of the cars hidden under a canopy of trees.

### 9.3 Relevant Textures for the Change Detection Algorithm

As it was done in (ref rodrigo) we will use the method of Sum and Difference Histograms to compute the textures (ref sarker), since they have been demonstrated to compute useful textures to be used as additional information for machine learning algorithms and it is a texture computation method that can run in linear time. We will again give a brief overview of the method.

For a given SAR image, we consider the random variable  $u_{[x,y]}$  to be the absolute value of the backscatter value received by the antenna. The first step of the algorithm is to quantize the random variable  $u$  to have  $N_g$  discrete integer possible values (on this

work  $N_g$  was chosen to be 20). The second step is to obtain an estimate of the joint probability density function of the random v  $[u_{[x,y]}, u_{[x+\delta x, y+\delta y]}]$  (where  $\delta = (\delta x, \delta y)$  is the displacement vector chosen to compute the texture direction and distance), which is a  $N_g \times N_g$  matrix. According to (ref Unser), it is possible to create an estimate of this Joint PDF in linear time by creating a random gaussian vector that has uncorrelated, therefore independent, random variables, and then computing the textures directly from the two linear independent variables. The new vector is  $[s_{[x,y]}, d_{[x,y]}]$  and is obtained as follows:

$$\begin{aligned} s_{[x,y]} &= u_{[x,y]} + u_{[x+\delta x, y+\delta y]} \\ d_{[x,y]} &= u_{[x,y]} - u_{[x+\delta x, y+\delta y]} \end{aligned} . \quad (9.1)$$

Since  $s_{[x,y]}$  and  $d_{[x,y]}$  are independent, it is possible to decompose the original Joint PDF as the product of two PDF as follows:

$$\begin{aligned} P(u_{[x,y]} = u_1, u_{[x+\delta x, y+\delta y]} = u_2) &= \\ P(s_{[x,y]} = u_1 + u_2, d_{[x,y]} = u_1 - u_2) &= \\ P(s_{[x,y]} = u_1 + u_2) \cdot P(d_{[x,y]} = u_1 - u_2) & \end{aligned} \quad (9.2)$$

By doing this it is now possible to compute the textures directly from the random variables  $P_s$  and  $P_d$ . Notice that by doing this, the time complexity of the texture calculation was reduced from  $O(N_g^2)$  to  $O(2 * N_g)$ .

After obtaining the probability density functions  $P(s_{[x,y]})$  and  $P(d_{[x,y]})$ , a set of textural information can be extracted as follows (for better readability, we define  $P_s(i) = P(s[x, y] = i)$  and  $P_d(j) = P(d[x, y] = j)$ ):

1. Mean( $\mu$ ): describes the mean value of the co-occurrences and is given by

$$\text{MEAN} = \frac{1}{2} \sum_{i=1}^{N_g} i \cdot P_s(i) \quad (9.3)$$

2. Cluster prominence: describes the fourth moment of the random variables and is given by

$$\text{CLP} = \sum_{i=1}^{N_g} (i - 2\mu)^4 \cdot P_s(i) \quad (9.4)$$

3. Cluster shade: describes the third moment of the random variables and is given by

$$\text{CLS} = \sum_{i=1}^{N_g} (i - 2\mu)^3 \cdot P_s(i) \quad (9.5)$$

4. Contrast: measure the difference of intensities of co-occurrences in the image and is

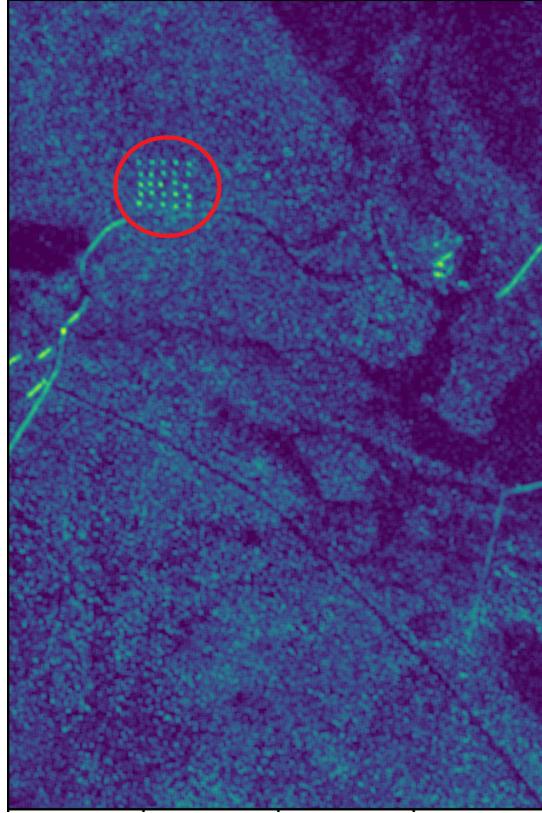


FIGURE 9.2 – Entropy image. The red circle shows the position of the vehicles.

given by

$$\text{CON} = \sum_{j=-\frac{N_g}{2}}^{\frac{N_g}{2}} (j) \cdot P_d(j) \quad (9.6)$$

5. Correlation: describes the linear correlation of intensities of co-occurrences in the image and is given by

$$\text{COR} = \frac{1}{2} \sum_{i=1}^{N_g} (i - 2\mu)^2 \cdot P_s(i) - \frac{1}{2} \sum_{j=-\frac{N_g}{2}}^{\frac{N_g}{2}} (j) \cdot P_d(j) \quad (9.7)$$

6. Energy: describes the energy of the received signal and is given by

$$\text{ENE} = \sum_{i=1}^{N_g} P_s(i)^2 \cdot \sum_{j=-\frac{N_g}{2}}^{\frac{N_g}{2}} P_d(j)^2 \quad (9.8)$$

7. Entropy: describes the uncertainty of the co-occurrence matrix and is given by

$$\text{ENT} = - \sum_{i=1}^{N_g} P_s(i) \log(P_s(i)) - \sum_{j=-\frac{N_g}{2}}^{\frac{N_g}{2}} P_d(j) \log(P_d(j)) \quad (9.9)$$

8. Homogeneity: describes the degree of similarity between images and is given by

$$\text{HOM} = - \sum_{j=-\frac{N_g}{2}}^{\frac{N_g}{2}} \frac{P_d(j)}{1+j^2} \quad (9.10)$$

9. Variance: describes the degree of dispersion of intensities between two images and is given by

$$\text{VAR} = \frac{1}{2} \sum_{i=1}^{N_g} (i - 2\mu)^2 P_s(i) + \frac{1}{2} \sum_{j=-\frac{N_g}{2}}^{\frac{N_g}{2}} j^2 P_d(j) \quad (9.11)$$

There are multiple textures that can be chosen as additional inputs for the UNET CNN. One have to keep in mind, that, for every texture that it is used as additional information, the time computation for the training and validation of the UNET will be also increased, so it is vital to keep in mind the trade-off between computation time and improvement of classification results.

For this work it was chosen to use the Entropy and Variance images textures as additional information. The original picture and the texture pictures can be seen in Fig. 9.1, Fig. 9.2 and Fig. 9.3.

## 9.4 UNET Based Algorithm for Target Change Detection

The UNET is a convolutional neural network that was developed in the Computer Science department of the University of Freiburg for the purpose of performing image segmentation for biomedical pictures (ref UNET). It became popular because it was able to perform precise segmentations with little time (segmentation of a 512x512 image takes less than a second using a modern GPU). Convolutional Neural networks normally work by performing a series of convolutions and pooling filters that decrease the size of the image, while the UNET works by also supplementing this usual operation pipeline with a series of successive layers in which the pooling operations are replaced by upsampling operators, therefore increasing the resolution of the output. The final goal is that a usual successive convolutional layer learn to create an output based on this information.

It is important to remind that an important aspect of the UNET is that there is a large

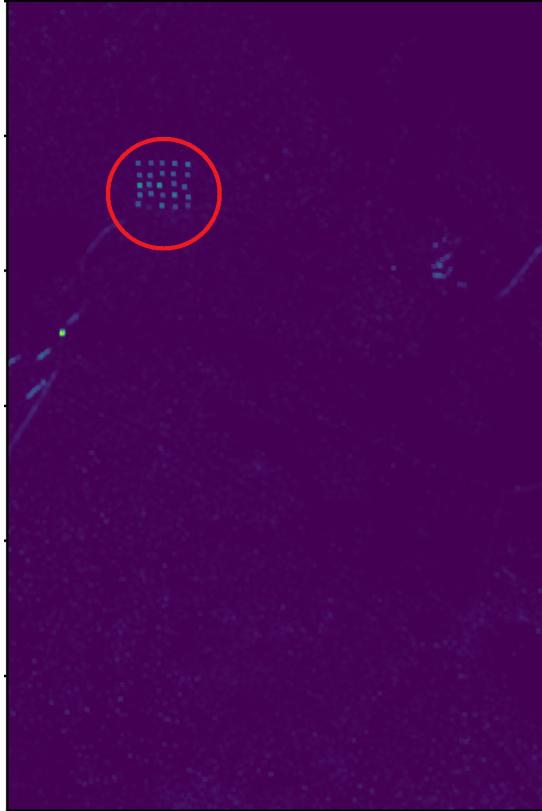


FIGURE 9.3 – Variance image. The red circle shows the position of the vehicles

number of feature channels responsible for the upsampling step, which ensures that the UNET propagates information to the higher resolution layers. After the expansive path is performed, it is combined with the usual pooling operations, creating a contractive path, which makes the UNET shows a u-shaped architecture. It is important also to mention that in the borders of the image, the UNET will extrapolate by mirroring the information to predict the pixels in the area. Fig 9.4 depicts the architecture of the UNET and it's u-shape.

The proposed CDA is based on this UNET architecture and has two phases: the training phase and the inference phase.

#### *A. Training Phase*

In this phase the model is treated like a regular supervised machine learning model. The given UNET architecture is trained in a supervised manner with the reference map to be able to perform semantic image segmentation.

#### *B. Inference Phase*

This phase is based on the Change Detection algorithm proposed by (ref Jong). This algorithm was based on modifying the original UNET two accept two images for the purpose of change detection. The original algorithm has 5 feature maps created by the convolutions that were generated and saved for the first image. After that, when the second image is used as input, a Difference Image (DI) is created at each of the 5 levels using

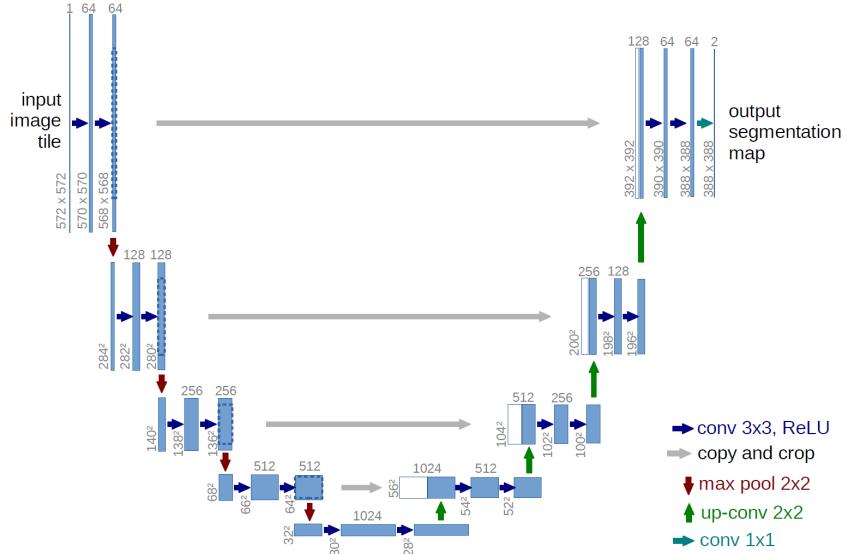


FIGURE 9.4 – UNET architecture

the feature maps of the first and second image. The DI is created by taking the absolute value of the difference between the feature maps, and setting the difference to be zero if the difference value does not surpass a threshold value (which was found empirically), and setting it to the value of the second image if it surpasses the threshold value. Intuitively this step works by setting it to zero all pixels that did not change sufficient (according to a threshold value). After that, the DI is semantically segmentated, rendering a visual interpretation of the collective changes that happened. A visual representation of this algorithm is show in Fig 9.5

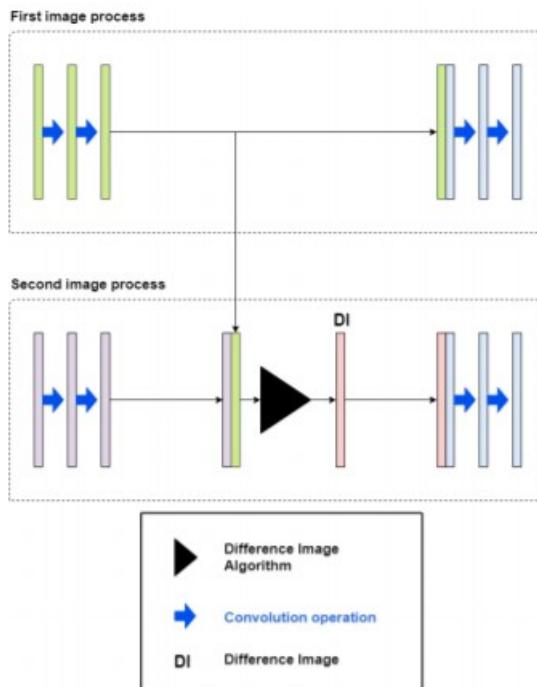


FIGURE 9.5 – Inference Phase for the corresponding DI

The CDA of this work is inspired by the algorithm explained above. Since the textures provide additional information, it would be very hard to empirically achieve the values for the threshold of each layer (in our case it would be a vector instead of a scalar), so it was decided to perform the difference image just for first input layer of the UNET. We used the textures of Entropy and Variance, which was combined with the original image to create a three-channel tensor which was fed, and trained, on the UNET.

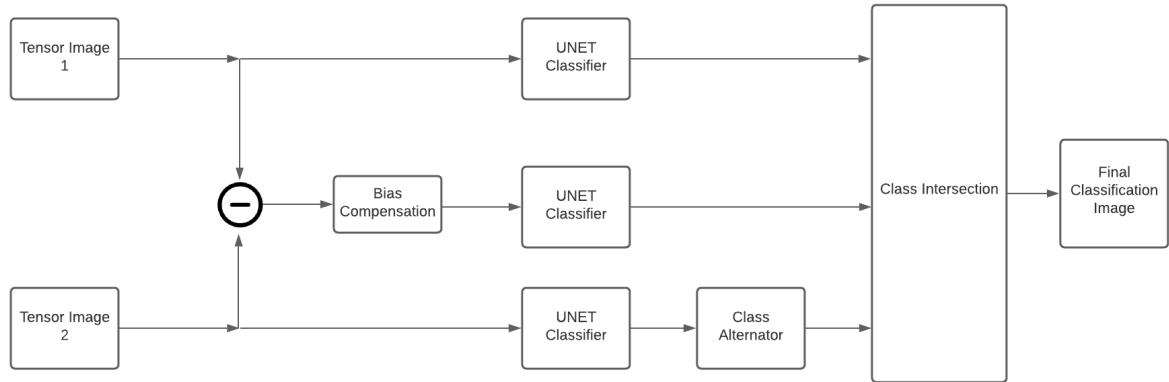


FIGURE 9.6 – Diagram of the proposed algorithm.

After creating the 24 image tensors, the dataset was divided into 12 tensors for training and 12 tensors for validation purposes. After the training was performed this trained UNET was used as part of the CDA. The proposed CDA has 4 steps, and works as follows:

1. Feed the first tensor as a input to the UNET and get the classification
2. Feed the second tensor as input to the UNET, and after that alternate the classes from the final image (switch all car classification pixels by forest classification pixels and vice-versa)
3. Take the absolute value of the difference of the tensors, perform a bias correction (the value for the bias correction was adjusted manually). After that feed this tensor as input to the UNET.
4. Get the three classifications obtained prior and take the intersection of all classes (in order to a pixel to be classified as a car in the final image, it has to be a car in all of the three images at the same time)

Item 2 was added to the algorithm simply to make it more robust against false errors. By adding another necessary step to identify the change as a target change difference we decrease the probability of false detection, but at the same time it will also decrease the overall probability of true detection for the algorithm.

After the final image is created by the CDA, a threshold filter selection was performed to eliminate noise classification from the image - we considered only sets of pixels which had 100 or more connected components.

In Fig 9.6 the proposed algorithm is represented visually.

Monitored Image	Reference Image	Detected Targets (with texture)	False Alarms (with texture)	Detected Targets (without texture)	False Alarms (without texture)
2_1	3_1	25	0	25	0
3_1	4_1	23	1	21	1
4_1	5_1	25	0	25	0
5_1	2_1	25	0	25	0
2_2	4_2	25	0	25	0
3_2	5_2	25	0	23	1
4_2	2_2	23	0	25	0
5_2	3_2	22	0	24	0
2_3	5_3	25	0	25	0
3_3	2_3	22	1	21	1
4_3	3_3	25	1	25	1
5_3	4_3	25	0	25	0
2_4	3_4	25	0	25	0
3_4	4_4	24	0	23	1
4_4	5_4	25	0	25	0
5_4	2_4	24	0	25	0
2_5	4_5	25	0	25	0
3_5	5_5	24	0	21	2
4_5	2_5	25	0	25	0
5_5	3_5	23	0	24	0
2_6	5_6	25	0	25	0
3_6	2_6	23	1	25	0
4_6	3_6	24	0	25	0
5_6	4_6	25	1	25	0

TABLE 9.1 – Results of CDA in terms of detected targets and false alarm rate for the same image pairs considered in ((??)). It is also presented the results when textural information is either used or not used as input of the proposed algorithm.

## 9.5 Results

As previously mentioned, the dataset of 24 images was divided into two parts: 12 images used for training the UNET and 12 images used for validation of the UNET training. After that 24 images pairs were selected to assess the quality of the change detection algorithm. The 24 pairs follow the standard selection that are used to validate

CDAs in the CARABAS dataset, as done in (ref carabas, ricardo, lucasramos).

The results are presented in Table 9.1, where each image M\_P was labelled in terms of Mission number (M) and flight pass (P). The compared images have different mission numbers, but the same flight pass heading. In Table 9.1 it is presented the results for the CDA algorithm using only the original image -in order to assess the quality of the CDA- and it is also presented the result using the textural information combined with the original image - in order to assess the quality improvement due to the additional textural information.

The quality of the algorithm can be measured in terms of probability of detection and False Alarm Rate (FAR). The proposed algorithm achieved a Probability of detection of 97% and a False Alarm Rate of  $0.034/km^2$ , while the original algorithm proposed in the CARABAS challenge paper achieved a probability of detection of 97% and a FAR of  $0.67/km^2$ . Other recent algorithms, such as the one presented in (ref ricardo) achieved a probability of detection of 97% and a FAR of  $0.28/km^2$  and (ref Vinholi) achieved a probability of detection of 99% and a FAR of  $0.0833/km^2$ .

## 9.6 Conclusion

The Change Detection Algorithm proposed yielded a probability of detection similar to the usual state-of-the-art algorithms existent - it had a probability of detection of 97% - but showed great improvements in terms of False Alarm Rate, reducing the probability of false alarm more than ten times, when compared to usual statistical change detection algorithms, and and reducing the probability of false alarm almost three times when compared to usual Convolutional Neural Network based machine learning algorithms.

# Bibliography

- BAMLER, R.; HARTL, P. Synthetic aperture radar interferometry. **Inverse Problems**, v. 14, n. 4, p. R1–R54, 1998. ix, 22, 28
- BRUZZONE, L.; MARCONCINI, M.; WERTMULLER, U.; WIESMANN, A. An advanced system for the automatic classification of multitemporal sar images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, p. 1321–1334, 2004. 34
- CUMMING, I.; WONG, F. **Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation**. first edition. [S.l.]: Artech House, 2005. 17
- GRAHAM, L. Synthetic interferometer radar for topographic mapping. **Proceedings of the IEEE**, v. 62, n. 6, p. 763–768, 1974. 22
- HAJNSEK, I.; MOREIRA, A. Tandem-x: Mission and science exploration. **Proceedings on the 6th European Conference on SAR. EUSAR '06, Dresden, Germany**, May 2006. 36
- KRIEGER, G.; MOREIRA, A.; FIEDLER, H.; HAJNSEK, I. Tandem-x: A satellite formation for high-resolution sar interferometry. **IEEE Transactions on Geoscience and Remote Sensing**, v. 45, p. 3317 – 3341, 2007. 30, 33
- MARTONE, M.; BRAUTIGAM, B.; RIZZOLI, P. Coherence evaluation of tandem-x interferometric data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 73, p. 21–29, 2012. 30
- MARTONE, M.; RIZZOLI, P. A method for generating forest/non-forest maps from tandem-x interferometric data. **2015 IEEE International Geoscience and Remote Sensing Symposium**, p. 2634–2637, 2015. 29, 30
- MARTONE, M.; RIZZOLI, P. The global forest/non-forest map from tandem-x interferometric sar data. **Remote Sensing of the Environment**, v. 205, p. 352–373, 2018. 29
- MOREIRA, A.; KRIEGER, G.; HAJNSEK, I.; HOUNAM, D.; WERNER, M.; RIEGGER, S.; SETTELMEYER, E. Tandem-x: a terrasar-x add-on satellite for single-pass sar interferometry. **Proceedings on Geoscience and Remote Sensing Symposium 2004. IGARSS '04, Anchorage, USA**, 20–24 Sept. 2004. 36
- MOREIRA, A.; PRATS-IRAOLA, P.; YOUNIS, M.; KRIEGER, G.; HAJNSEK, I. Sentinel-1, the gmes radar mission. **IEEE Geoscience and Remote Sensing Magazine**, v. 1, p. 6–43, 2013. ix, 18, 20

- PAPOULIS, A.; PILLAI, S. U. **Probability, Random Variables and Stochastic Processes**. fourth edition. [S.l.]: McGraw-Hill Europe, 2002. 48
- PULELLA, A. **An analysis of the potentials of TanDEM-X data to characterize tropical forest horizontal heterogeneity**. Tese (Doutorado) — University of Pisa, 2017. 29
- RANEY, K.; FREEMAN, A.; HAWKINS, B.; BAMLER, R. A plea for radar brightness. **Proc. IGARSS**, p. 1090–1092, 1994. 20
- RIZZOLI, P. **Radar Backscatter Modeling Based on Global TanDEM-X Mission Data**. Tese (Doutorado) — Karlsruhe Institute of Technology, 2018. ix, 23, 25, 26, 27, 33
- RIZZOLI, P.; MARTONE, M.; HOTT, H.; MOREIRA, A. Characterization of snow facies on the greenland ice sheet observed by tandem-x interferometric sar data. **Remote Sensing** 9, v. 4, p. 21–29, 2017. 29
- ROSEN, P.; HENSLEY, S.; JOUGHIN, I.; LI, F.; MADSEN, S.; RODRIGUEZ, E.; GOLDSTEIN, R. Synthetic aperture radar interferometry. **Proceedings of the IEEE**, v. 88, n. 3, p. 333 – 382, 2000. 22
- SICA, F.; PUDELLA, A.; NANNINI, M. P. M.; RIZZOLI, P. Repeat-pass sar interferometry for land cover classification: a methodology using sentinel-1 short-time-series. **Remote Sensing of Environment**, v. 232, 2019. 30
- SMALL, D. Flattening gamma: Radiometric terrain correction for sar imagery. **IEEE Transactions on Geoscience and Remote Sensing**, v. 49, n. 8, 2011. ix, 20, 21
- SNOEIJ, P.; ATTEMA, E.; DAVIDSON, M.; FLOURY, N.; LEVRINI, G.; ROSICH, B. Sentinel-1, the gmes radar mission. **2008 IEEE Radar Conference**, December 2008. 38
- UNSER, M. Sum and difference histograms for texture classification. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Pami-8, n. 1, p. 118–125, 1986. 50
- WERTMULLER, U.; WERNER, C. L. Sar interferometric signatures of forest. **IEEE Transactions on Geoscience and Remote Sensing**, v. 33, p. 1153– 1161, 1995. 33
- WOODCOCK, C.; STRAHLER, A. The factor of scale in remote sensing. **Remote Sens. Environ.**, v. 21, n. 3, p. 311–332, 1987. 43
- ZUCKER, S. W.; TERZOUPoulos, D. Finding structure in co-occurrence matrices for texture analysis. **Computer Graphics and image processing**, v. 12, n. 4, p. 286–308, 1980. 45, 52