This is the Title of my Thesis

Valdemar Edvard Sandal Rolfsen

Desember 2017

PROJECT

Department of Geomatics
Norwegian University of Science and Technology

Supervisor 1: The main supervisor

Supervisor 2: The co-supervisors (internal and external)

Preface

Here, you give a brief introduction to your work. What it is (e.g., a Master's thesis in RAMS at NTNU as part of the study program xxx and...), when it was carried out (e.g., during the autumn semester of 2021). If the project has been carried out for a company, you should mention this and also describe the cooperation with the company. You may also describe how the idea to the project was brought up.

You should also specify the assumed background of the readers of this report (who are you writing for).

Trondheim, 2012-12-16

(Your signature)

Ola Nordmann

Acknowledgment

I would like to thank the following persons for their great help during . . .

If the project has been carried out in cooperation with an external partner (e.g., a company), you should acknowledge the contribution and give thanks to the involved persons.

You should also acknowledge the contributions made by your supervisor(s).

O.N.

(Your initials)

Remark:

Given the opportunity here, the RAMS group would recognize Professor Emeritus Marvin Rausand for the work to prepare this template. Some minor modifications have been proposed by Professor Mary Ann Lundteigen, but these are minor compared to the contribution by Rausand.

Executive Summary

Here you give a summary of your your work and your results. This is like a management summary and should be written in a clear and easy language, without many difficult terms and without abbreviations. Everything you present here must be treated in more detail in the main report. You should not give any references to the report in the summary – just explain what you have done and what you have found out. The Summary and Conclusions should be no more than two pages.

You may assume that you have got three minutes to present to the Rector of NTNU what you have done and what you have found out as part of your thesis. (He is an intelligent person, but does not know much about your field of expertise.)

Contents

Chapter 1

Introduction

Pattern recognition and feature extraction in remote sensing, has been a field of research for many decades. The ability to extract geospatial information directly from satellite imagery has been crucial for fields such as environmental and demographic research. A vast development within the field of machine learning and artificial intelligence the last years, has enabled many researchers to find useful applications of such algorithms within their own scientific fields.

With new satellite technology providing frequently updated imagery, with resolutions less than 0.5 meters, the amount of data that can be extracted is almost incomprehensible. What part will pattern recognition using machine learning play in the future of geospatial analyses?

1.1 Background

Satellites has been collecting earth observation data for decades. Since the satellite Explorer 6 took the first picture of the earth in 1959, millions of satellite images has been taken, processed and stored Esa (2009). This information however, has been very difficult to access, and even harder to analyze when accessible. However, during the last decade, the development of machine learning based methods for Earth Science applications has experiences a considerable leap forward Lary (2010).

In 2014 the first satellite in the new family of earth observation satellites, called the Sentinels, was launched from Kourou, French Guiana. Since then 7 different constellations, each consisting of two satellites, have been launched, and are now orbiting and monitoring the earths surface. The goal of these satellites is to produce a continuous stream of timely data for Europe's Copernicus program, which will be used for environmental monitoring. The different constellations have different missions when it comes to providing datasets for the Copernicus Service.

While some provide very specific data, such as monitoring the earths atmosphere, other constellations provide more general data, such as multi-spectral, high-resolution imagery of the earths surface. In order to maximize the usage of these temporal datasets, they have all been provided free of charge to the public. The fact that all of these datasets are given free of charge

World View constellation

Cushing Oklahoma

Problem Formulation

You should define your problem in a clear an unambiguous way and explain why this is a problem, why it is of interest—and to whom. It is also important to delimit the problem area.

Related work

You should here present the main books and articles that treat problems that are similar to what you are studying, and give proper references to each of these as they are reported. If you, later in your thesis, describe the "state of the art" – with a detailed literature survey, you may just give a very brief survey here (approx. a quarter of a page). If this is the only literature survey, you need to go into more details. An objective of the literature survey is to show the reader that you are familiar with the main literature within your field of research – so that you do not "reinvent the wheel."

Sentinel-1 is a radar image mission, measuring both land and sea. It has two identical satellites orbiting 180 degrees apart.
Sentinel-2 provides multi-spectral, high resolution imagery of the earths surface.
Sentinel-3 is equipped with multiple instruments used to measure sea-surface topography, sea- and land-surface temperature, ocean color and land color.
Sentinel-4
Sentinel-5
Sentinel-5P
Sentinel-6

References to literature can be given in two different ways:

- As an *explicit* reference: It is shown by **?** and partly also by **?** that ....
- As an *implicit* reference: It is shown (e.g., see **?**, Chap. 4) that ....

In the example above, we have used "author-year" references, which is the preferred format.

Remark: Following agreement with your supervisor, you may also refer by numbers, for example, [1]. To do this, open the file `ramsstyle.sty` and comment out (by %) the command `\usepackage{natbib}` and un-comment the corresponding command `\usepackage[numbers]{natbib}`.[1]

You may include a link to the Internet in the text or in a footnote by using a command like: http://www.ntnu.edu/ross.

When you refer to the scientific literature, you should always write in *present* tense. Example: **?** show that ....

Remark: Hyperlinks are included by the command `\usepackage{hyperref}` in `ramsstyle.sty`. If you feel that the hyperlinks are disturbing when you enter the text, or want to avoid the hyperlinks in printed text, you may either comment out or edit this command in `ramsstyle.sty`.

What Remains to be Done?

After you have defined and delimited your problem – and presented the relevant results found in the literature within this field, you should sum up which parts of the problem that remain to be solved.

1.2 Objectives

The main objectives of this Master's project are

1. This is the first objective
2. This is the second objective
3. This is the third objective
4. More objectives

The objectives shall be written as *fundamental objectives* telling what to do and not *means objectives* telling how to do it.

---

[1] Notice the strange way we have to write the "backslash" in the text. This is because the "backslash" is a command in LaTeX.

All objectives shall be stated such that we, after having read the thesis, can see whether or not you have met the objective. "To become familiar with ..." is therefore not a suitable objective.

1.3 Approach

Here you should describe the (scientific) approach and experiments that you will use or have used to solve the problem and meet your objectives and tasks. Experiments may in this context relate analyses you need to carry out in order to investigate a specific hypothesis, task objective, or similar. You should specify the approach and experiments for each objective and/or task. It is preferred that you supplement your explanation of the approach with an illustration.

If there are any ethical problems related to your approach, these should be highlighted and discussed.

1.4 Contributions

Here you give a list of your main contributions in the project or master work.

1.5 Limitations

In this section you describe the limitations of your study. These may be related to the study object (physical limitations, operational limitations), to the environmental and operational conditions, to the thoroughness of the analysis, and so on.

1.6 Outline

Here, you give an overview of how the remaining part of the report is organized. A proposed structure of the main chapters in the report can be as follows (note that some chapters are not numbered):

- Preface: Contains practical information about what you have done, and where the work has been carried out. Any assumed background of the reader should be specified here.
- Acknowledgments: Here, you show the gratitude to who have been supporting your work, professionally and family as relevant.

- Summary: Contains the management summary, and should be a layman's explanation of what you have done and why it is important. This would be the talk you could give if you in an interview is asked about what you did in your thesis, or if some of your relatives ask the same question. This chapter should therefore include as few domain specific words as possible, so that no detailed background in the topic is required.
- Chapter 1. Introduction: Structure already discussed in this chapter.
- Chapter 2. Theoretical background: Here you identify and give the theoretical background needed in this report, with proper references to each literature reference used. The selection of what to include should be discussed and agreed with the supervisors. Theory may involve concepts, definitions, methods, regulations/key standards, theory to explain specific system behavior, and so on.
- Chapter 3..N-2: The naming of the following chapters relies entirely on the specific topic in question. Proposed structure should be discussed with supervisor.
- Chapter N-1 Results: This chapter should be the last chapter *before* "Conclusions, discussion, and ideas for further work"
- Chapter N. Conclusions, discussion, and ideas for further work.
- Bibliography
- Appendix A etc (as needed): Appendix A may for example be acronyms as shown here.

Remark: Notice that chapter and section headings shall be written in lowercase, but that all main words should start with a capital letter.

The report should be no longer than 60 pages in this format for the master as well as the specialization project, with the possible exception of appendixes (which may take up some space if including e.g. code from programming). This does not mean that the report must be at least 60 pages, and the effort should be directed to be as concise as possible throughout the report.

Chapter 2

Theoretical background

In this chapter we are going to investigate some of the key concepts related to pattern recognition and shape detection in imagery.

2.1 Shape detection

Identifying geometric shapes in computer vision has been a classical problem for decades. There are many theories related to what is the best way of detecting a particular shape in an image, with shapes defines as two dimensional features of an object that are invariant to scene factors, or whose variation can be modeled easily (Moon et al., 2002).

2.1.1 The standard approach

Most images are based on a raster format, meaning that pixels in the image are structured as an array or grid, where each pixel is associated with a position (row and column), and a numeric value. Raster images can represent a range of different shapes, where a point can be represented by a single pixel and a circle by a contiguous collection of pixels (Worboys, 2003). Even though rasters are easy to work with in most computer systems, since they are represented the way that they are, they do not contain any information about the topology of the objects in the image. For example, there is no way of knowing if a pixel is contained within a certain object or not. In order to identify objects, one approach is therefore to detect edge pixels.

There are many well-known methods that are used in order to find an objects external contour, such as the generalized Hough transform (GHT) (Ballard, 1981), fitting of polynomial curves and applying band matching after edge detection.

In recent years shape detection algorithms have come to increasingly rely on superpixel algorithms, which groups pixels into perceptually meaningful atomic regions (Achanta et al., 2012). Such regions replace the regular, rigid structure of the raster grid, as shown in Figure 2.1.

Figure 2.1: Visualization of image segmentation using SLIC (Achanta et al., 2012)

When constructing superpixels, there are some properties of the algorithm that are desirable, regardless of the problem that are being solved. These are, according to (Achanta et al., 2012), the following three points:

- Superpixels should adhere well to image boundaries
- When used to reduce computional complexity as a preprocessing step, superpixels should be fast to compute, memory efficient, and simple to use.
- When used for segmentation purposes, superpixels should both increase the speed and improve the quality of the results.

Achanta et al. (2012) proposes a superpixel-segmentation algorithm (SLIC), which in their opinion is best suited to meet these demands. They compare their algorithm to a variety of state-of-the-art superpixel methods, and conclude that none of the existing methods are satisfactory in regards to the points.

Simple Linear Iterative Clustering (SLIC)

The SLIC algorithm is fairly simple to understand. One of its key principles is that, by limiting the search space for each cluster center (points in the regular raster grid), it reduces the search speed significantly. This is achievable due to the fact that one of the primary goals of algorithm is to create a set of approximately equal sized superpixels. Thus, instead of searching the whole raster grid for each cluster center, the algorithm only has to search for edge pixels at a distance equal to D, as shown in Equation 2.1.

$$D' = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2} \tag{2.1}$$

In Equation 2.1 d_c is the euclidean distance between two pixels in terms of color and d_s is the pixels euclidean, spatial distance. Furthermore, S is the sampling interval of the cluster centers ($S = \sqrt{N/k}$, where N is the number of pixels in the grid and k is the desired number of superpixels) and m is a fixed constant based on the color diversity in the image.

Since the algorithm generates superpixels by clustering pixels based on their color and spatial proximity, creating a 5 dimensional, *labxy* space, one would think that the distance could be found by simply taking the 5D euclidean distance. However, it turns out that for large superpixels, spatial distance outweigh the color proximity. Which is why the two distances d_c and d_s are weighted.

2.1.2 The cognitive approach

In recent years, the development of machine learning, a branch of artificial intelligent systems, has become increasingly important in terms of pattern and object recognition in remote sensing and image analysis in general.

Lary et al. (2016) propose nine different branches of machine learning that are commonly used for data mining in remote sensing. These algorithms are artificial neural networks (ANN), support vector machines (SVM), self-organizing map (SOM), decision trees (DT), ensemble methods such as random forests, case-based reasoning, neuro-fuzzy (NF), generic algorithm (GA) and multivariate adaptive regression splines (MARS). In particular, three of these algorithms are of interest.

Artificial Neural Networks

The basic principle behind artificial neural networks (ANNs) is that it is built up by a network of many simple units that are working in parallel with no centralized control unit. The networks primary means of storage are the weights between the individual units, and the network learns by updating these weights in relation to being provided with a training example.

In order to understand the behavior of ANNs it is important to understand their structure. An ANN is build up by a given number of connected layers. A layer is a collection of simple units, called artificial neurons. All ANNs must consist of an input and output layer, but can also contain an optional number of hidden layers. A network which only consists of an input and output layer is called a perceptron.

Support vector machines

In machine learning, support vector machines (SVM) are supervised learning models used for classification and regression analysis.

In a SVM model training examples are represented as points in a n-dimensional space, mapped so that the examples belonging to separate categories are clustered together. One or more hyperplanes are then fitted in order to separate the different classes from each other. New data is then mapped into that same space and classified based on which side of the hyperplane they are mapped. SVMs are often called large margin classifiers because their goal is to separate the classes by fitting the hyperplanes between them in a way so that the planes has the highest possible margin between the classes as shown in Figure 2.3.

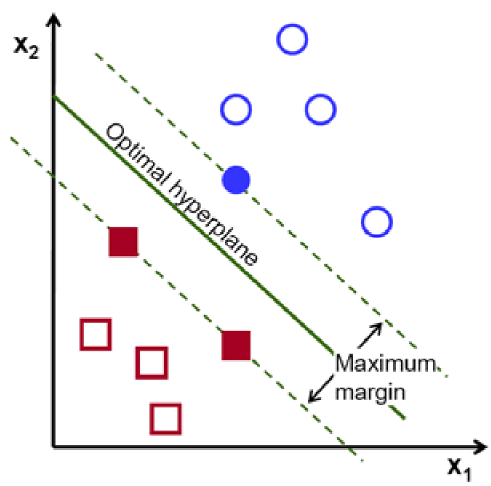

Figure 2.2: Example of optimal hyperplane for classification from Opencv (2017)

In order for support vector machines to work, the data points has to be separable by a plane. This is however not always possible in the dimension that the data is represented. In order to solve this issue SVM apply kernel functions.

Kernel functions enable SVM to operate in a high-dimensional, implicit feature space without ever computing the coordinates of the data in that dimension space. It simply computes the inner products between the images of all pairs of data in the feature space. This operation is often computationally cheaper than the explicit computation of the coordinates.

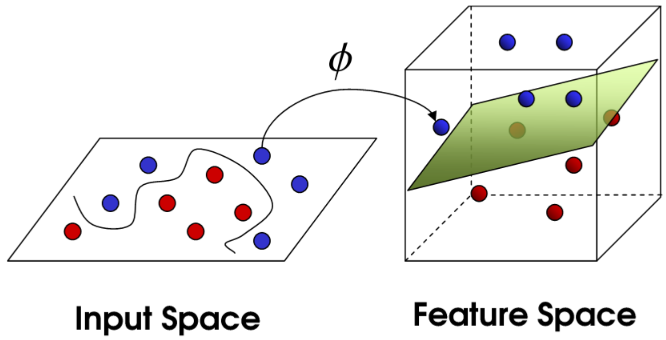

Figure 2.3: Classification using SVM with kernel function Opencv (2017)

Decision Trees

Decision tree learning is one of the simplest and yet most successful forms of machine learning. They are learned top-down, and they use a recursive learning technique.

A decision tree represents a function that takes as input a vector of attribute values and returns a "decision" - a single output value. Furthermore, it is worth mentioning that decision trees are only good for some kind of functions.

Links:

https://www.pyimagesearch.com/2014/10/20/finding-shapes-images-using-python-opencv/
https://www.pyimagesearch.com/2016/02/08/opencv-shape-detection/
http://melvincabatuan.github.io/SLIC-Superpixels/

Chapter 3

Equations, Figures, and Tables

The content of Chapter 2 will vary with the topic of your thesis. This chapter only gives guidance to some technical aspects of LaTeX.

Remark: If you want a shorter chapter or section title to appear in the Table of Contents and in the headings of the chapter, you just include the short title in square brackets before the title of the chapter/section. Example:

```
\section[Short Title]{Long Title}
```

.

3.1 Simple Equations

Mathematical symbols and equations can written in the text as λ, $F(t)$, or even $F(t) = \int_0^t \exp(-\lambda x)\,dx$, or as displayed equations

$$F(t) = \int_0^t \exp(-\lambda x)\,dx \tag{3.1}$$

The displayed equations are automatically given equation numbers – here (3.1) since this is the first equation in Chapter 2. Note that you can refer to the equation by referring to the "label" you specified as part of the equation environment.

You can also include equations without numbers:

$$F(t) = \sum_{i=1}^{n} \binom{n}{i} \sin(i \cdot t)$$

More Advanced Formulas

Long formulas that cannot fit into a single line can be written by using the environment `align` as

$$F(t) = \sum_{i=1}^{n} \sin(t^{n-1}) - \sum_{i=1}^{n} \binom{n}{i} \sin(i \cdot t) \tag{3.2}$$

$$+ \int_0^\infty n^{-x} e^{-\lambda x^t} \, dt \tag{3.3}$$

In some cases, you need to write ordinary letters inside the equations. You should then use the commands

`\textrm  and/or \mathrm`

The first command returns the normal text font and will be scaled automatically, while the second command will be scaled according to the use.

$$\mathrm{MTTF} = \int_0^\infty R_{\mathrm{avg}}(t) \, dt$$

Please consult the LaTeX documentation for further details about mathematics in LaTeX.

Definitions

If you want to include a definition of a term/concept in the text, I have made the following macro (see in `ramsstyle.sty`):

☞ **Reliability**: The ability of an item to perform a required function under stated environmental and operational conditions and for a stated period of time.

When text is following directly after the definition, it may sometimes be necessary to end the definition text by the command

`\newline`

I have not included this in the definition of the `defin` environment to avoid too much space when there is not a text-block following the definition.

Figure 3.1: This is the logo of NTNU (rotated 15 degrees).

3.2 Including Figures

If you use pdfLaTeX (as recommended), all the figures must be in pdf, png, or jpg format. We recommend you to use the pdf format. Please place the figure files in the directory **fig**. Figures are included by the command shown for Figure 3.1. Please notice the “path” to the figure file written by a *forward* slash (`/`). You should not include the format of the figure file (pdg, png, or jpg) – just write the “name” of the figure.

Each figure should include a unique *label* as shown in the command for Figure 3.1. You can then refer to the figure by the *ref* command. Notice that you can scale the size of the figure by the option `scale=k`. You may also define a specific width or height of the figure by replacing the `scale` options by `width=k` or `height=k`. The factor `k` can here be specified in mm, cm, pc, and many other length measures. You may also give `k` as a fraction of the width of the text or of the height of the text, for example, `width=0.45\textwidth`. If you later change the margins of the text, the figure width will change accordingly. As illustrated in Figure 3.1, you may also rotate the figure – and also do many other things (please check the documentation of the package `graphicx` – it is available on your computer, or you may find it on the Internet).

In LaTeX all figures are floating objects and will normally be placed at the top of a page. This is the standard option in all scientific reports. If you insist on placing the figure exactly where you declare the figure, you may include the command `[h]` (here) immediately after `\begin{figure}`. If you will force the figure to be located either at the top or bottom of the page, you may alternatively use `[t]` or `[b]`. For more options, check the documentation.

Large figures may be included as a *sidewaysfigure* as shown in Figure 3.2:[1]

3.3 Including Tables

LaTeX has a lot of different options to include tables. Only one of them is illustrated here.

[1] You can use a similar command for large tables.

NTNU – Trondheim
Norwegian University of
Science and Technology

Figure 3.2: This is the logo of NTNU.

Table 3.1: The degree of newness of technology.

Experience with the operating condition	Level of technology maturity: Proven	Limited field history or not used by company/user	New or unproven
Previous experience	1	2	3
No experience by company/user	2	3	4
No industry experience	3	4	4

Remark: Notice that figure captions (Figure text) shall be located *below* the figure – and that the caption of tables shall be *above* the table. This is done by placing the `\caption` command beneath the command `\includegraphics` for figures, and above the command `\begin{tabular*}` for tables.

3.4 Copying Figures and Tables

In some cases, it may be relevant to include figures and tables from from other publications in your report. This can be a direct copy or that you retype the table or redraw the figure. In both cases, you should include a reference to the source in the figure or table caption. The caption might then be written as: *Figure/Table xx: The caption text is coming here (**?**).*

In other cases, you get the idea from a figure or table in a publication, but modify the figure/table to fit your purpose. If the change is significant, your caption should have the following format: *Figure/Table xx: The caption text is coming here (adapted from **?**).*

3.5 References to Figures and Tables

Remember that all figures and tables shall be referred to and explained/discussed in the text. If a figure/table is not referred to in the text, it shall be deleted from the report.

3.6 A Word About Font-encoding

When you press a button (or a combination of buttons) on your keyboard, this is represented in your computer according to the *font-encoding* that has been set up. A wide range of font-encodings are available and it may be difficult to choose the "best" one. In the template, I have set up a font-encoding called UTF-8 which is a modern and very comprehensive encoding and is expected to be the standard encoding in the future. Before you start using this template, you should open the Preferences ->Editor dialogue in TeXworks (or TeXShop if you use a Mac) and check that encoding UTF-8 has been specified.

If you use only numbers and letters used in standard English text, it is not very important which encoding you are using, but if you write the Norwegian letters æ, ø, å and accented letters, such as é and ä, you may run into problems if you use different encodings. Please be careful if you cut and paste text from other word-processors or editors into your LaTeX file!

Warning

If you (accidentally) open your file in another editor and this editor is set up with another font-encoding, your non-standard letters will likely come out wrong. If you do this, and detect the error, be sure *not* to save your file in this editor!!

This is not a specific LaTeX problem. You will run into the same problem with all editors and word-processors – and it is of special importance if you use computers with different platforms (Windows, OSX, Linux).

3.7 Plagiarism

Plagiarism is defined as "use, without giving reasonable and appropriate credit to or acknowledging the author or source, of another person's original work, whether such work is made up of code, formulas, ideas, language, research, strategies, writing or other form", and is a very serious issue in all academic work. You should adhere to the following rules:

- Give proper references to all the sources you are using as a basis for your work. The references should be give to the original work and not to newer sources that mention the original sources.
- You may copy paragraphs up to 50 words when you include a proper reference. In doing so, you should place the copied text in inverted commas (i.e., "Copied text follows ..."). Another option is to write the copied text as a quotation, for example:

  > Birnbaum's measure of reliability importance of component i at time t is equal to the probability that the system is in such a state at time t that component i is critical for the system.
  >
  > **?**

Chapter 4

Conclusions, Discussion, and Recommendations for Further Work

In this final chapter you should sum up what you have done and which results you have got. You should also discuss your findings, and give recommendations for further work.

4.1 Summary and Conclusions

Here, you present a brief summary of your work and list the main results you have got. You should give comments to each of the objectives in Chapter 1 and state whether or not you have met the objective. If you have not met the objective, you should explain why (e.g., data not available, too difficult).

This section is similar to the Summary and Conclusions in the beginning of your report, but more detailed—referring to the the various sections in the report.

4.2 Discussion

Here, you may discuss your findings based on your results, their strengths and limitations. Note that this discussion is more high level than discussions made in relation to results you have achieved and presented in the previous chapter. The discussion here should put your work in larger context. You may address if you achieved what you had intended to do, why not (if you did not), if you got results in which you did not expect, why the results are important, why there are limitations in using the results, or if there are opportunities to transfer your results and findings into other domains, and so on.

4.3 Recommendations for Further Work

You should give recommendations to possible extensions to your work. The recommendations should be as specific as possible, preferably with an objective and an indication of a possible approach.

The recommendations may be classified as:

- Short-term
- Medium-term
- Long-term

Appendix A

Acronyms

FTA Fault tree analysis

MTTF Mean time to failure

RAMS Reliability, availability, maintainability, and safety

Appendix B

What to put in appendixes

This is an example of an Appendix. You can write an Appendix in the same way as a chapter, with sections, subsections, and so on. An appendix may include list of code (in case you are programming), more details about results that you have presented in the report (could be a more complete description of results, in case you decided to focus on the most important ones in the main report), supplementary information and descriptions you have found relating to the system you are analysing, such as drawings. You may discuss with your supervisor what are relevant information for appendixes.

B.1 Introduction

B.1.1 More Details

Bibliography

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2281.

Ballard, D. H. (1981). Generalizing the Hough Transform To Detect Arbitrary Shapes*. *Pattern Recoqnition*, 13(2):111–122.

Esa (2009). 50 years of Earth Observation.

Lary, D. J. (2010). Artificial Intelligence in Geosience and Remote Sensing. *RFID Technology, Security Vulnerabilities, and Countermeasures*, pages 75–100.

Lary, D. J., Alavi, A. H., Gandomi, A. H., and Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1):3–10.

Moon, H., Chellappa, R., and Rosenfeld, A. (2002). Optimal edge-based shape detection. *IEEE Transactions on Image Processing*, 11(11):1209–1227.

Opencv (2017). OpenCV 2.4.13.4 documentation.

Worboys, M. F. (2003). *GIS: A computing perspective*.