**VIETNAM NATIONAL UNIVERSITY – HO CHI MINH CITY**

**INTERNATIONAL UNIVERSITY**

**DEPARTMENT OF PHYSICS**

**SPACE ENGINEERING**



**PROJECT’S TOPIC**

Nguyen VAN A

A thesis submitted to

Department of Physics

in partial fulfillment of the Requirements

for the Degree of Engineer IN SPACE ENGINEERING

Ho chi minh city, viet nam  
2020

**PROJECT’S TOPIC**

**Nguyen VAN A (Full name)**

Under the guidance and approval of the committee, and approved by its members, this thesis has been accepted in partial fulfillment of the requirements for the degree.

Approved:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Chairperson

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Committee member

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Committee member

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Committee member

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Committee member

# HONESTY DECLARATION

My name is Nguyen Trong Nghia. I would like to declare that, apart from the acknowledged references, this thesis either does not use language, ideas, or other original material from anyone; or has not been previously submitted to any other educational and research programs or institutions. I fully understand that any writings in this thesis contradicted to the above statement will automatically lead to the rejection from the SE program at the International University – Vietnam National University Ho Chi Minh City.

Date:

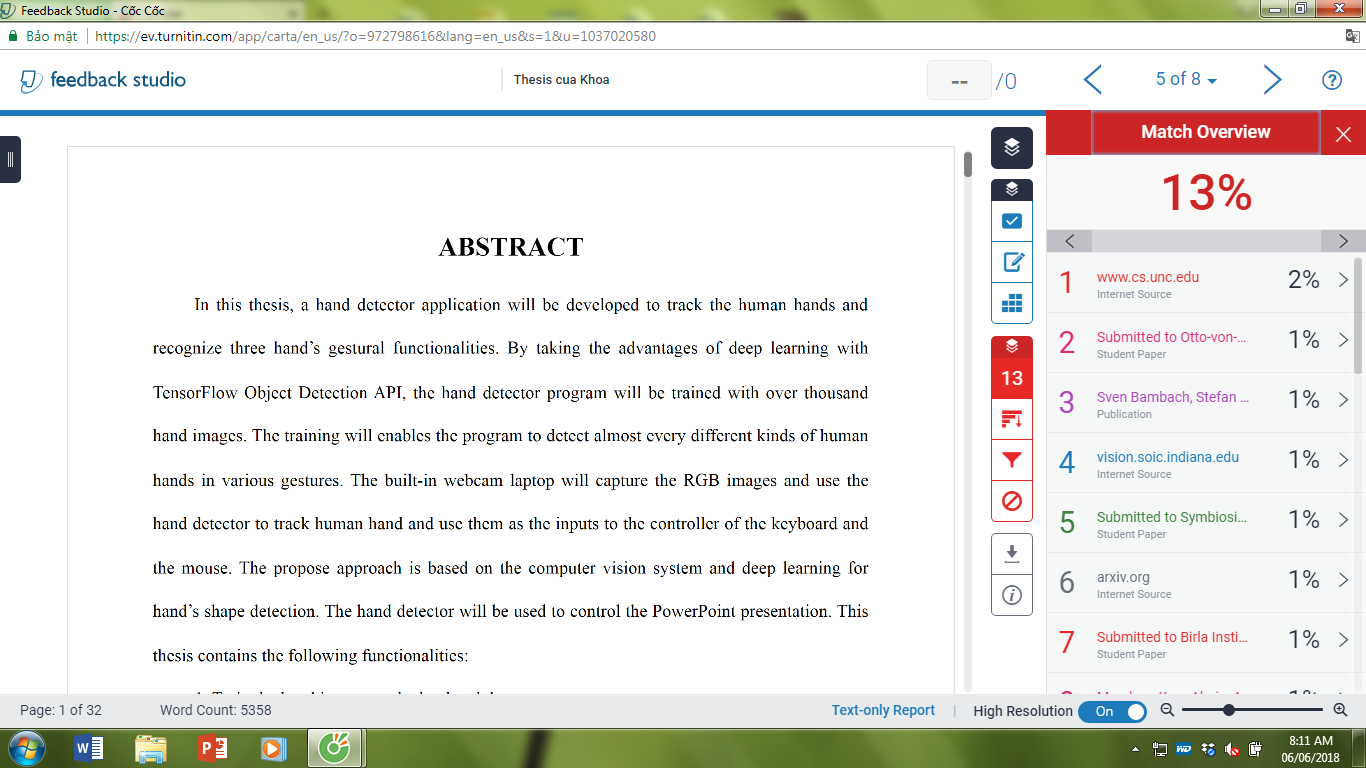
Student’s Signature

(Full name)

# TURNITIN DECLARATION

Name of Student: Nguyen Trong Nghia

Date:



Advisor Signature Student Signature

# ACKNOWLEDGMENT

It is with deep gratitude and appreciation that I acknowledge the professional guidance of Dr. Nguyen Van B. His constant encouragement and support helped me to achieve my goal.

………………………

# TABLE OF CONTENTS

[HONESTY DECLARATION 3](#_Toc21345696)

[TURNITIN DECLARATION 4](#_Toc21345697)

[ACKNOWLEDGMENT 5](#_Toc21345698)

[TABLE OF CONTENTS 6](#_Toc21345699)

[LIST OF TABLES 7](#_Toc21345700)

[LIST OF FIGURES 8](#_Toc21345701)

[ABBREVIATIONS AND NOTATIONS 9](#_Toc21345702)

[ABSTRACT 9](#_Toc21345703)

[CHAPTER I INTRODUCTION 10](#_Toc21345704)

[1.1. Level-2 Title 10](#_Toc21345705)

[1.1.1. Level-3 title 10](#_Toc21345706)

[CHAPTER II LITERATURE REVIEW 12](#_Toc21345716)

[CHAPTER III METHODOLOGY 12](#_Toc21345717)

[CHAPTER IV RESULTS 12](#_Toc21345718)

[CHAPTER V DISCUSSION AND IMPLEMENTATIONS 12](#_Toc21345719)

[CHAPTER VI CONCLUSION AND RECOMMENDATION 13](#_Toc21345720)

[REFERENCES 13](#_Toc21345721)

[APPENDICES 13](#_Toc21345722)

# LIST OF TABLES

Table 1.1 1

Table 1.2 2

Table 2.1 3

Table 2.2 5

Table 2.3 12

Table 2.4 22

# LIST OF FIGURES

Figure 1.1 1

Figure 1.2 1

Figure 2.1 6

Figure 2.2 13

Figure 2.3 22

Figure 3.1 26

# ABBREVIATIONS AND NOTATIONS

SE: Space Engineering

RS: Remote Sensing

# ABSTRACT

Remote sensing data are frequently used in extraction of many interesting physical characteristics such as land use, urbanization, vegetation index, water levels,… One of the most frequent problem in Remote Sensing is the lost of data due to spatial pixels being covered by clouds. We believe that these loss data can be recovered based on the spatial relationship between satellite data taken at different time. While there have been a lot of papers addressing the exact same problem – following the same assumptions as us. These papers usually interpolate continuous numerical data – meaning the interpolated values followed an implicit linear level in their characteristic. In this paper, we try to apply the same method for categorical satellite data, and explore the use of a Nature Language Processing neural network in interpolating discrete nominal satellite data.

# CHAPTER I INTRODUCTION

The Introduction presents the background of the thesis work. It reveals the state of the art of the topic by quoting previously published works in the fields and important questions that your work relates to them. It should contain parts such as Rationale, Problem statement, Objectives, Scope, Limitation, Research framework, Structure of research and the summary of the contains of the following parts of the report. It is important that this section be unique and specific to the report. As a common practice, it is conceived before you do your work but is written when you finish writing all other chapters. Most of its contents can be used in your thesis report.

After decades of continuous innovations, satellite-based sensing technologies have achieved remarkable results in their ability to offer high spatial resolution data on a global scale, giving us the necessary tools to make wide and continuous observations of the Earth’s surface. One such outstanding example is the LandSat missions, which have been capturing high quality imagery of the Earth’s surface since July 1972 [1].

While LandSat offer high spatial resolution data of the Earth’s surface every 14 days. Many of these valuable data has to be discarded due to atmospheric conditions, mainly clouds, which force many studies to use data with high temporal gap. For example, a study regarding Mangrove Forest change detection in Ca Mau, Viet Nam used LandSat data from 1979 to 2013 [2]. Such limitation limits the amount of data researcher can use to form and validate their hypothesis, creating unnecessary challenges in earth surface analysis.

Several methods for interpolating cloud-covered pixels have been developed. The main assumption of these methods is that covered pixels at a specific point in time can be interpolated using an image in which this pixel is not covered by clouds. Meaning pixels obtained on a cloudy day can be interpolated using information obtained in more favorable conditions. Most of the methods developed for cloud-filling are based on matrix decomposition, most noticeably, DINEOF (Data Interpolation Empirical Orthogonal Functions) [3].

DINEOF assumes that data matrix reconstruction using leading orthogonal functions will capture low-frequency large-scale data, while minimizing the noise and high frequency components resulted from missing values. The aforementioned matrix is a collection of flatten satellites images – meaning it is a 2D matrix, which each row being a satellite image flatten into 1D vector. This representation of satellite data capture both time and spatial information – for which DINEOF will try to find the latent factors represented by high-ranking orthogonal functions. The most recent implementation of DINEOF is called TIEOF (Tensor Interpolation using Empirical Orthogonal Functions) [4], which make use of unflatten satellite images – introducing the distinction between horizontal and vertical into the scheme.

Another scheme based on matrix decomposition were explore by Ruo-Qian Wang (2021) [5]. In which a back propagation implementation of matrix decomposition called funk-svd were used [6]. While originally inspired by Recommender System used in Netflix film recommendation, the main ideal behind funk-svd is using back propagation to minimize the loss of the reconstructed matrix.

Matrix decomposition-based schemes achieved remarkable results in many cloud-filling tasks [3, 4, 5, 7, 8]. However, due to the inherent underlying assumption of matrix decomposition, these methods only work for continuous numerical data – where there exists a linear relationship in the reconstructed between values in the reconstructed matrix. In another word, a value of in the reconstructed matrix will have an inherent relationship with a value of in the matrix – and the value of is closer to the value of , compared to the value of . Practically speaking, this inherent relationship makes matrix decomposition-based scheme ill-suited for categorical data – where there exists no such relationship between the value of and . This was proven true in our experiment, which will be discussed in Chapter 3.

After such experiment with matrix decomposition-based scheme, Dr. Van suggest the usage of an Attention based NLP model (Nature Language Processing) [9] – whose use of tokenization and query pair eliminate the associated linear relationship found in matrix decomposition. Specifically, BERT (Bidirectional Encoder Representation from Transformer) [10] were suggested

## **1.1. Level-2 Title**

Your text should begin from here. This is the illustration for the use of the level-2 title style

* 1. **1. Level-3 title**

Your text should begin from here. This is the illustration for the use of the level-3 title style

The format of equations, formula:







Table 1.1.Table’s Name

|  |  |  |  |
| --- | --- | --- | --- |
| **Circuit** | **Vref (V)** | **Iref (mA)** | **Iref2 (mA)** |
| Circuit 1 | 0.01 | 1 | 2 |
| Cirucit 2 | 0.1 | 2 | 3 |
| Etc. |  |  |  |

Table 1.2.Table’s Name

|  |  |  |  |
| --- | --- | --- | --- |
| **Circuit** | **Vref (V)** | **Iref (mA)** | **Iref2 (mA)** |
| Circuit 3 | 0.01 | 1 | 2 |
| Cirucit 4 | 0.1 | 2 | 3 |
| Etc. |  |  |  |



Figure 1.1. Figure’s Name



Figure 1.2. Figure’s Name

# CHAPTER II LITERATURE REVIEW

The Literature Review Chapter should be briefly summarized problems which relate to your topic. In particular, you can collect reliable information from journals, conference papers and other materials.

# CHAPTER III METHODOLOGY

The Methodology Chapter should show the methods and materials used in our work. Explain why you chose them. Describe in such a way that others can replicate exactly your work.

# CHAPTER IV RESULTS

The Results Chapter presents what you obtained or failed to obtain

# CHAPTER V DISCUSSION AND IMPLEMENTATIONS

The Discussion and Implementations Chapter presents all critical analysis of merits and shortcomings of your results, comparison with reported results in literature, and how you will carry on the incomplete or planned works in your thesis work.

# CHAPTER VI CONCLUSION AND RECOMMENDATION

The Conclusion Chapter summarizes your results, remarks, comments, suggestions and directions to be developed or improved in the future.

# REFERENCES

The References part indicates the sources of your information: books, brochures, catalogs and names and coordinates of people you consulted, if applied. List them in alphabetical order. Use scarcely a website as a reference because the information may not be scientific approved or be erroneous and the sites may disappear in the future. In the chapters refer an article or book by using author's name and date of publication, and put them in parentheses such as (Smith, 1990) or (Smith et al., 1990) if there are more than 2 authors or (Smith et al., 1990a) if there are different cited articles started with the same name. Use appropriate format for the bibliography as advised in the Endnote software.

|  |  |
| --- | --- |
| [1] | C. N. E. Anagnostopoulos, "License plate recognition: A brief tutorial," *IEEE Intelligent transportation systems magazine,* vol. 6, no. 1, pp. 59–67, Jan. 2014. |
| [2] | W. Jackson, *Learn Android App Development*, Apress, 2013. |
| [3] | K. Aida-zade, E. Mustafaev, and J. Hasanov, "Intelligent reading system based on mobile platform," in *International Conference on Problems of Cybernetics and Informatics*, Baku, Azerbaijan, 2012, pp. 1-4. |
| [4] | A. Mutholib, T. S. Gunawan, and M. Kartiwi, "Design and implementation of automatic number plate recognition on android platform," in *International Conference on Computer and Communication Engineering*, Kuala Lumpur, Malaysia, 2012, pp. 540–543. |
|  |  |

# APPENDICES

The Appendices part contains all related documents, materials and samples you obtained. Use this part wisely to alleviate the containers of the chapters. In other words, give a brief description or results in the chapters to make the reading flows well and use this part to describe in detail the issues for readers who desire to get in depth information. Organize the specific information in different appendices with a title such as Appendix 3: Specifications Sheet of integrated circuit 744. Refer them appropriately in the chapters using the brackets such as [Appendix 1].