AUTOMATED EGGPLANT SORTING AND GRADING SYSTEM USING COMPUTER VISION AND MACHINE LEARNING FOR QUALITY CLASSIFICATION

An Undergraduate Thesis

Presented to the
Faculty of Bachelor of Science in Computer Engineering
University of Science and Technology of Southern Philippines
Cagayan de Oro City

In Partial Fulfillment
of the Requirements for the Degree of
BACHELOR OF SCIENCE IN COMPUTER ENGINEERING

ALEXANDRA E. JAGONOS
MARK MADORABLE
RALPH B. GALANIDO
JIHARA MAE B. APUT
JONAS RON FRANCO P. MALAZARTE
November 2025

APPROVAL SHEET

This Thesis entitled: "AUTOMATED EGGPLANT SORTING AND GRADING SYSTEM USING COMPUTER VISION AND MACHINE LEARNING FOR QUALITY CLASSIFICATION", prepared and submitted by ALEXANDRA E. JAGONOS, MARK MADORABLE, RALPH B. GALANIDO, JIHARA MAE B. APUT, and JONAS RON FRANCO P. MALAZARTE in partial fulfillment of the requirements for the degree of **BACHELOR OF SCIENCE IN COMPUTER ENGINEERING** has been examined and approved.

Engr. Darwin Jone H. Jupiter
Adviser

PANEL OF EXAMINERS

Approved by the committee on Oral Examination with a grade of **PASSED**.

Engr. Darwin Jone H. Jupiter
Chair

Engr. Jasper Jay A. Jementiza
Member

Mrs. Mirasol D. Rizon
Member

Approved and accepted in partial fulfillment of the requirements for the degree **BACHELOR OF SCIENCE IN COMPUTER ENGINEERING**.

Approved:

Dr. Israel A. Baguhin
Dean, College of Engineering and Architecture

ABSTRACT

This study...

Keywords: *computer vision, grading system, eggplant classification*

This piece of work is wholeheartedly dedicated

to my parents

Papang

and

Nanay

ACKNOWLEDGMENT

I would like to express my gratitude to the following people and institutions which, in one way or another, greatly contributed towards the completion of this study.

To my adviser ***Engr. Jodie Rey D. Fernandez*** for painstakingly checking every detail of this paper. I am extremely thankful and indebted to you Sir for sharing your expertise and valuable guidance. Had not of your encouragement and moral support, this endeavor would surely fall to a complete obscurity.

To my panel members, ***Engr. Mark Lister V. Nalupa*** and ***Engr. Miriam M. Bergado*** for sharing your knowledge and diligently spending your time in giving valuable insights, corrections and suggestions for the betterment of this study. I have high regards to both of you.

To the faculty and staff of the ***Department of Computer Engineering*** for your encouragement and support.

To my classmates, ***Methos***, ***Shirley***, ***Lovely***, ***Lady Lee***, ***Cris***, and ***Ma'am Bern*** for sharing your laughter and wonderful moments with me. Finally, sleepless nights and tiring days are over but they are all worth it.

To my friends, ***Ruemar***, ***Chyn***, ***Vine***, ***Abigail***, ***Rachel***, ***Long***, ***Elaine***, ***Claire***, ***Allan***, ***Xander***, ***Laiza***, ***Ate Leslie***, ***Sir Greig***, ***Jean***, ***Sir Harold***, ***Ate Josh***, ***Ma'am Amelia***, ***Jovy***, ***Leonel***, ***Jerson***, ***Sir Raymund***, ***Sir Nur P.***, ***Mama Marge***, ***Mama Belle***, ***Mama Lady***, ***Mama Hazel***, ***Brother Mark***, ***Sir Nur***, ***Sir Elmer*** and to my ***AHK Family*** for your prayers and moral support. You all made this journey worthy of remembering.

To my family, ***Nanay Angga***, ***Papang Willy***, ***Nanay Toto***, ***Ina***, ***Ate Love***, ***Ate Ging***, ***Kuya Boboy***, ***Andi***, ***Marlo*** and ***Ate Lyn*** for your unceasing love, moral and financial support. All of you are my inspirations which constantly remind me of not giving up amidst difficulties and uncertainties.

And to those people I failed to mention, who directly or indirectly, lent their hands, extended their prayers and support for the completion of this study.

Above all, to our ***Almighty God***, whose unconditional love infinitely transcends all human comprehensions, provides us with good health, divine protection and bountiful blessings.

Jodie

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Equations

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Eggplant (*Solanum melongena* L.) has been cultivated for centuries and is now a staple in cuisines worldwide. Also known as aubergine, brinjal, or *talong* in the Philippines, eggplant holds a prominent position in the vegetable sector, with global production of approximately 60.8 million metric tons from an area of over 1.9 million hectares (Food and Agriculture Organization of the United Nations (FAO), 2025). In April to June of 2023, the Philippines produced 102.98 thousand metric tons of eggplant (Philippine Statistics Authority, 2023), making it one of the most important and widely grown vegetable crops in the country.

But before being sold, eggplant is manually categorized by visual inspection to determine their quality (Sun et al., 2025). This includes assessing the fruit's diameter, length, volume, curvature, color homogeneity, calyx area (the green, leaf-like "cap" at the fruit's stem end), calyx color, and surface defects (Lyu et al., 2025; Lalam et al., 2025). Sorting only takes a few hours for small-scale farmers, but it can take a little longer and more workers on large plantations, since the quantity of harvested eggplants is directly proportional to labor and time needed (Khan et al., 2025).

Because most people just use their eyes to judge the quality of eggplant, there is a significant rate of inaccuracy in addition to the time-consuming and expensive nature of manual eggplant sorting (Waghmare et al., 2025). As a result, a number of cutting-edge technologies were combined to increase classification accuracy and decrease inefficiencies, particularly deep learning and computer vision. For instance, the study of

1.2 Statement of the Problem

This study seeks to investigate some properties of decomposable hyper KS-semigroups in the context of strong, weak, quasi- and bi-hyper KS-ideals.

1.3 Objectives of the Study

In view of the above stated problem, we have the following objectives:

1. To introduce the concept of strong, weak, quasi- and bi-hyper KS-ideals;
2. To provide characterizations of strong, weak, quasi- and bi-hyper KS-ideals and investigate their relationships;
3. To introduce the idea of decomposable hyper KS-semigroups and give some characterizations.

1.4 Significance of the Study

The concept of hyperstructures is itself, a powerful mathematical tool since algebraic hyperstructures seem to occur very naturally in many areas of mathematics and even in other disciplines.

1.5 Scope and Limitations

The primary motivation of this study lies within the structural properties of hyper

1.6 Definition of Terms

Data Logger An electronic device that records data over time or in relation to location either with a built-in instrument or sensor or via external instruments and sensors.

GPS Tracking Using the Global Positioning System to determine and track the precise location of a person, vehicle, or other asset.

Real-time Monitoring The process of continuously observing a system or process and immediately reporting any changes or anomalies.

Sensor A device that detects or measures a physical property and records, indicates, or otherwise responds to it.

Telemetry The process of recording and transmitting the readings of an instrument.

Wireless Communication The transfer of information between two or more points that are not connected by an electrical conductor.

CHAPTER 2

REVIEW OF RELATED LITERATURE

This chapter focuses on related studies or projects that have provided additional relevant information to the proponent. Section 2.1 presents the theoretical background related to computer vision, cyber-physical systems, and machine learning. Section 2.2 specifies the concept of how the system will classify the eggplants by class based on surface defects, color homogeneity, and size. Section 2.3 presents a review of literature related to the design of employing computer vision, machine learning, and mechatronic integration in agricultural automation for image-based inspection and mechanical sorting applications. Collectively, these reviewed literatures support the concept presented in this study.

2.1 Theoretical Background

The development of an automated eggplant grading system is fundamentally grounded in the integrated theoretical paradigms of computer vision, deep learning, and mechatronics. The entire process can be conceptualized through a sequential computer vision pipeline, which moves from image acquisition to physical actuation (Szeliski, 2022). This pipeline begins with the digital image formation theory, where a camera sensor captures light reflectance from objects on a conveyor belt, a common setup in food process engineering (Dougherty, 2020). The stability and quality of this initial stage are paramount, as controlled, diffuse illumination is critical to minimize specular reflections and shadows that can obscure critical features like color and surface defects, thereby ensuring consistent input for subsequent algorithmic processing.

Following acquisition, image pre-processing theories are applied to enhance data quality and standardize inputs. This involves digital signal processing techniques such as noise reduction using Gaussian or median filters (Kumar and Sodhi, 2020) and, crucially, color space transformation. Converting images from the default Red-Green-Blue (RGB) space to Hue-Saturation-Value (HSV) or CIEL*a*b* is a well-established step in agricultural product inspection (Khan and AlGhamdi, 2024). The theoretical underpinning for this conversion lies in the decoupling of color information (chrominance) from intensity (luminance) in these spaces, making the extracted color features more robust to minor variations in lighting conditions, which is essential for accurate color-based grading.

The core of the system resides in the theoretical distinction between traditional machine learning and deep learning for feature extraction. Traditional approaches are based on manual feature engineering, drawing from image processing and pattern recognition theory to hand-craft descriptors for color (e.g., histograms, mean values), shape (e.g., aspect ratio, roundness, area), and texture (e.g., using Gray-Level Co-occurrence Matrix (GLCM) or Local Binary Patterns (LBP)) (Haralick et al., 2007; Ojala et al., 2002). In contrast, deep learning, specifically Convolutional Neural Network (CNN) theory, posits that models can automatically learn a hierarchical representation of features directly from raw pixel data (LeCun et al., 2015; Goodfellow et al., 2016). The lower layers of a CNN learn generic features like edges and corners, while deeper layers synthesize these into complex, task-specific features relevant to defect identification, thereby eliminating the need for manual feature design.

For the classification task itself, the theoretical models diverge. Machine learning classifiers operate on the principles of statistical learning theory, finding optimal hyperplanes to separate classes or constructing ensembles of decision trees, respectively, based on the handcrafted features (Hearst et al., 1998; Breiman, 2001). Conversely, an end-to-end deep learning model uses a fully connected output layer with a normalized exponential function (also known as *Softmax*) to perform classification based on the high-level features it learned itself. The theoretical advantage of CNNs is their superior ability to model complex, non-linear relationships in visual data, a capability that has been demonstrated to outperform traditional methods in numerous agricultural applications (Bharman et al., 2022).

The integrity of any supervised learning system is dependent upon the quality of its labeled dataset. The establishment of a reliable ground truth is a theoretical problem rooted in psychometrics and expert systems. The grading labels for the eggplant images must be derived from standardized agricultural protocols, such as those provided by the United States Department of Agriculture (USDA) or equivalent bodies, which define quality classes based on size, color uniformity, and defect tolerances (USDA, 2013). Locally, the project will utilize the criteria defined by the Philippine National Standard (PNS), and Bureau of Agricultural and Fisheries Product Standards (BAFPS) (specifically the PNS/BAFPS 52:2007), which classifies eggplants into Extra Class, Class I, and Class II based on visual quality, size, and defect tolerances (BAFPS, 2007; PNS, 2019). To ensure scientific rigor, the concept of inter-rater reliability, often measured by Cohen's Kappa statistic (κ), must be applied to quantify the agreement between human experts who label the dataset, thereby validating the consistency of the training labels (McHugh, 2012; He et al., 2022).

The closed-loop system theory integrates the digital classification with physical actuation. The decision output from the classification model (e.g., a specific grade) is transmitted via a software-to-hardware interface to a Programmable Logic Controller (PLC) or microcontroller (Bolton, 2015). This triggers a mechatronic actuator, such as a pneumatic pusher or a servo-controlled diverter, which physically sorts the eggplant into its designated bin on the conveyor line. This final stage embodies the principle of cyber-physical systems, where computational intelligence directly controls a physical process to achieve full automation, replicating and potentially surpassing human grading efficiency (Lee, 2008; Zhang et al., 2022).

CHAPTER 3

METHODOLOGY

In this chapter, we detail the methodology employed to conduct the study, providing a comprehensive overview of the research design, data collection, and analytical procedures.

3.1 Research Design

Your research design.

3.2 Formula

3.3 Tables

Table 1: A sample long table.

First column	**Second column**	**Third column**
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778

Continued on next page

Table 1 – continued from previous page

First column	Second column	Third column
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778

Continued on next page

Table 1 – continued from previous page

First column	Second column	Third column
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778
One	abcdef ghjijklmn	123.456778

3.4 Images

Table 2: Sample Data Table

Item	**Quantity**	**Price ($)**
Apples	10	0.50
Bananas	5	0.30
Cherries	20	1.20
Dates	50	2.50

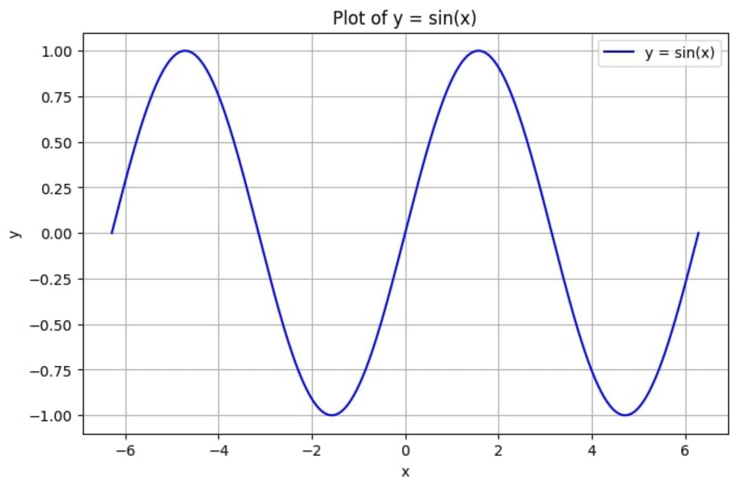

Figure 1

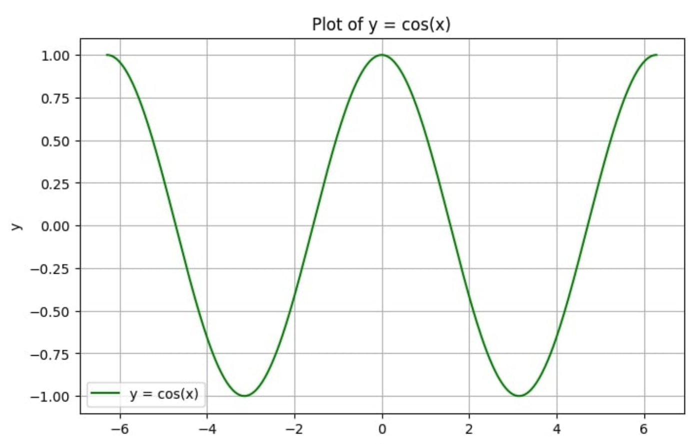

Figure 2: Cosine Graph

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents the findings from the research conducted and provides a thorough analysis and interpretation of these results.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This chapter provides the summary of the results obtained in this study and gives some recommendations for further investigation.

5.1 Summary of Findings

The study's findings address the initial research questions by confirming the effectiveness, reliability, and diverse applications of telemetry systems. The "Summary of Findings" section provides a concise overview of the key results from your research. This section should be factual and focus on presenting the data without interpretation. It should include:

Key Results:

Briefly summarize the most significant findings. Use bullet points or numbered lists for clarity if appropriate. Present the data as it was found, highlighting major patterns, relationships, or trends. Data Presentation:

Include tables, graphs, or charts that succinctly summarize the data. Make sure each visual aid is clearly labeled and includes a brief description. Coverage of Research Questions:

Address each of the research questions or hypotheses posed at the beginning of the study. Summarize the results relevant to each question.

5.2 Conclusion

The "Conclusions" section interprets the findings and discusses their implications. This section should:

Interpret Findings:

Provide an interpretation of the data summarized in the previous section. Discuss what the results mean in the context of the research questions or hypotheses. Implications:

Explain the significance of the findings. Discuss how the results contribute to the field of study or practical applications. Limitations:

Acknowledge any limitations in the study that may affect the results or their interpretation.

5.3 Recommendations

The ”Recommendations” section provides actionable suggestions based on the study’s findings and conclusions. This section should:

Practical Applications:

Offer specific recommendations for practitioners, policymakers, or other stakeholders based on the findings. Future Research:

Suggest areas for further investigation that could address the study’s limitations or build on its findings. Implementation:

Provide guidance on how the recommendations can be implemented effectively.

APPENDICES

Type your appendix here.

References

BAFPS (2007). PNS/BAFPS 52:2007 Fresh vegetables - Eggplant - Grading and Classification. As cited in Pinakbet Adventures ni Pina: Komik Serye.

Bharman, P., Saad, S. A., Khan, S., Jahan, I., Ray, M., and Biswas, M. (2022). Deep learning in agriculture: a review. *Asian Journal of Research in Computer Science*, 13(2):28–47.

Bolton, W. (2015). *Programmable logic controllers*. Newnes.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Dougherty, E. R. (2020). *Digital image processing methods*. CRC Press.

Food and Agriculture Organization of the United Nations (FAO) (2025). FAOSTAT: Production (Crops and livestock products) - Eggplants (aubergines). Data retrieved for item 'Eggplants (aubergines)'.

Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. (2016). *Deep learning*, volume 1. MIT press Cambridge.

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (2007). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.

He, M., Baker, R., Hutt, S., and Zhang, J. (2022). A less overconservative method for reliability estimation for cohen's kappa.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28.

Khan, M. A. and AlGhamdi, M. A. (2024). An intelligent and fast system for detection of grape diseases in rgb, grayscale, ycbcr, hsv and l* a* b* color spaces. *Multimedia Tools and Applications*, 83(17):50381–50399.

Khan, Z., Rawat, S., Kumar, S., and Sharma, A. (2025). Using computer vision for automated grading and sorting of fruits and vegetables. *International Journal of Agriculture and Food Science*, 7:1056–1065.

Kumar, A. and Sodhi, S. S. (2020). Comparative analysis of gaussian filter, median filter and denoise autoenocoder. In *2020 7th international conference on computing for sustainable global development (INDIACom)*, pages 45–51. IEEE.

Lalam, R., Lavanya, K., Nadella, V., and Kiran, B. R. (2025). Automatic sorting and grading of fruits based on maturity and size using machine vision and artificial intelligence. *Journal of Scientific Research and Reports*, 31(1):153–163.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.

Lee, E. A. (2008). Cyber physical systems: Design challenges. In *2008 11th IEEE international symposium on object and component-oriented real-time distributed computing (ISORC)*, pages 363–369. IEEE.

Lyu, J., Jin, L., Ma, X., Li, Y., Sun, M., Jin, N., Wang, S., Hu, L., and Yu, J. (2025). Agronomic and quality traits of 30 eggplant germplasm resources from china. *Plants*, 14(12):1838.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.

Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987.

Philippine Statistics Authority (2023). Major vegetables and root crops quarterly bulletin, April-June 2023: Eggplant. Bulletin published in September 2023.

PNS (2019). Philippine National Standard on Good Agricultural Practices (PNS-GAP): "Gabay para sa ligtas na pagkain". Diliman, Quezon City, Philippines. Story by Mandigma, M. G. R., and Tecson, J. V. D. R.; Arts and Graphics by Reyes, A. V.

Sun, H., Fu, R., and Kang, D.-K. (2025). A novel efficient eggplant disease detection method with multi-scale learning and edge feature enhancement. *Frontiers in Plant Science*, 16:1666955.

Szeliski, R. (2022). *Computer vision: algorithms and applications*. Springer Nature.

USDA (2013). United States Standards for Grades of Eggplant, §51.2190 to §51.2207. Supersedes third issue effective October 29, 1953.

Waghmare, D., Mulani, A., Takale, S. R., Godase, V., and Mulani, A. (2025). A Comprehensive Review on Automatic Fruit Sorting and Grading Techniques with Emphasis on Weight-based Classification. *Research & Review: Electronics and Communication Engineering*, 2:1–10.

Zhang, K., Shi, Y., Karnouskos, S., Sauter, T., Fang, H., and Colombo, A. W. (2022). Advancements in industrial cyber-physical systems: An overview and perspectives. *IEEE Transactions on Industrial Informatics*, 19(1):716–729.

CURRICULUM VITAE

Name:	**Jodie Rey D. Fernandez**
Date of Birth:	July 27, 1992
Place of Birth:	Alae, Manolo Fortich, Bukidnon
Home Address:	Zone 5, Alae, Manolo Fortich, Bukidnon
E-mail Address:	`jodierey.fernandez@ustp.edu.ph`

EDUCATIONAL BACKGROUND:

Master's Degree:	MSU-Iligan Institute of Technology Tibanga, Iligan City Master of Science in Computer Applications June 2020
Undergraduate Degree:	Mindanao University of Science and Technology Lapasan, Cagayan de Oro City Bachelor of Science in Computer Engineering April 2013
Secondary:	Alae National High School Alae, Manolo Fortich, Bukidnon March 2008
Elementary:	Alae Elementary School Alae, Manolo Fortich, Bukidnon March 2004

CERTIFICATE OF AUTHENTIC AUTHORSHIP

I hereby declare that this submission is my own work and, to the best of my knowledge, it contains no materials previously published or written by another person, nor material which, to a substantial extent, has been accepted for the award of any other degree or diploma at USTP or any other educational institution, except where due acknowledgement is made in the manuscript. Any contribution made to the research by others, with whom I have worked at USTP or elsewhere, is explicitly acknowledged in the manuscript.

I also declare that the intellectual content of this manuscript is the product of my own work, except to the extent that assistance from others in the project design and conception or in style, presentation and linguistic expression is acknowledged.

JODIE REY D. FERNANDEZ

JODIE REY D. FERNANDEZ

JODIE REY D. FERNANDEZ

Date: ”insert date signed