Development of a Baybayin Words Recognition System Using Support Vector Machine

Kyeong Nan C. Gim [1.1], Fiona S. Nabua [1.2], Meo Vincent C. Caya [1.3]

[1]School of Electrical, Electronics, and Computer Engineering
Mapúa University
Manila 1002, Philippines

Email: kncgim@mymail.mapua.edu.ph, fsnabua@mymail.mapua.edu.ph, mvccaya@.mapua.edu.ph

2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET) | 979-8-3503-8969-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/IICAIET62352.2024.10730538

***Abstract*— Ancient scripts like Baybayin hold immense cultural heritage yet remain inaccessible in modern digital applications. This is due to optical character recognition (OCR) systems struggling to recognize Baybayin characters accurately due to their rather complex structure. This research develops a Support Vector Machine (SVM) system that recognizes and converts Baybayin words into their respective Tagalog representation. The linear and RBF kernels enhanced character prediction, improving accuracy and robustness. An overall accuracy score of 96.88% was obtained for the RBF kernel on the more straightforward Baybayin characters and 93.59% for the Linear Kernel on the more complex Baybayin characters. The processed images are classified into central character bodies and accent signs; the characters are predicted and combined before changing into their respective character equivalents from the accents. Characters are recombined after accent modifications to form the predicted Baybayin word. Testing on a dataset of 120 sample words of varying words written in Baybayin achieved 90.83% accuracy in word recognition. This indicates SVM's capability to distinguish the intricate strokes and variations of the Baybayin script. The optical recognition system provides an advancement towards digitally archiving and revitalizing the Baybayin cultural heritage**

Keywords*— Support Vector Machine, Baybayin Recognition, Word Recognition Model, Machine Learning, Kernel, Character Recognition, Optical Character Recognition

I. Introduction

During the pre-colonial era, Baybayin was widely used as a writing system in the Northern Philippines. Recently, it has become popular due to efforts by various groups and the desire to preserve it for future generations [1]. Research shows that Baybayin helps people feel a sense of cultural identity, pride, and belonging [2]. The Baybayin alphabet has 17 characters: 14 are syllabic consonants and 3 are vowels. Even with advances in digital text storage, there are still challenges in preserving ancient languages like Baybayin digitally, mainly because traditional Optical Character Recognition (OCR) tools struggle with its unique features [3].

Baybayin character recognition requires a robust image processing algorithm to handle complex character strokes. The Support Vector Machine (SVM) algorithm demonstrates significant potential in this area, with its successful applications across diverse fields. SVM's ability to discern subtle features and identify intricate patterns is essential in tasks such as coffee bean classification, medical diagnosis such as red blood cell identification, handwriting analysis for forgery, and rice quality assessment [4], [5], [6], [7]. SVM's effectiveness is further demonstrated in tasks requiring precise feature analysis, such as facial recognition for class attendance, BMI estimation, and pineapple ripeness determination [8], [9], [10], and in its adaptability to diverse applications like food quality assessment and disease detection [11], [12], [13]. Similarly, SVM's precision in feature-based classification tasks is shown in its use for assessing fish freshness and species with high accuracy, demonstrating the method's effectiveness in detailed visual inspections [14].

Alternately, several datasets of texts and images are put into an experimental setup which uses multiple algorithms and neural networks for feature extraction. This resulted in varying degrees of accuracy on SVM classification, with the algorithm feature extraction giving better accuracy [15]. In contrast, one study found that SVM could recognize tiny or capital handwritten Latin characters with 90% precision. Another study, on the other hand, used printed characters and was able to identify Khmer characters with 98.62% accuracy. Furthermore, SVM produced a 96.29% accuracy rate in identifying Chinese letters when combined with directional histograms [16], [17], [18]. Another OCR application for Tamil and English is presented in an article that uses SVM and Gabor filters to achieve 97% accuracy in Tamil and 84% accuracy in English across various font sizes. Despite SVM's initial design for binary classification, combining two-class SVMs facilitates multiclass classification [19] [20]. Moreover, SVM has demonstrated its capability in classifying complex and informal data sets, as evidenced in a study where it successfully categorized informal academic-related messages from Facebook with an accuracy of 81.6% [21]. SVM's broad applicability extends beyond character recognition into fields such as bioinformatics, data mining, and audio recognition [22]. There also have been studies where a neural network like YOLO has been used to identify Baybayin, showing a relatively good accuracy of 98.92% [23]. While there has been limited study on the use of image processing in identifying Baybayin, there is research that focuses on using SVM in identifying Baybayin. The multiclass SVM classification has an accuracy of 96.51% and 95.8% in accuracy for character classification for both Baybayin and Latin. The binary SVM shows a 95.8% accuracy and 100% accuracy for script recognition of Baybayin and Latin, as well as diacritical marks classification of Baybayin.

Although Baybayin has been declared as a national writing language, only a few people know how to write and read the characters due to the lack of digital storing for the Baybayin characters. This study will look further into the gaps of other research, such as taking account of image preprocessing techniques that could be explored to enhance the accuracy of Baybayin character recognition. SVM performance in identifying various handwritten

979-8-3503-8969-2/24/$31.00 ©2024 IEEE

combinations of Baybayin characters with accompanying accents.

The main objective of this study is to develop a system that would recognize Baybayin words using the Support Vector Machine. The study's specific objectives are: (1) Use the Raspberry Pi 4 connected to a camera module to capture and detect Baybayin words. (2) Training and developing the software model using Support Vector Machine. (3) Use a confusion matrix to assess the system's performance.

The scope of the study focuses on analyzing the efficiency of an OCR that uses SVM for identifying Baybayin words. This will be done by separating each Baybayin character and their corresponding accents, doing a prediction for each character, and recombining the characters to form a word. The accents are considered as a separate class by the system from the main body allowing for better prediction by the SVM algorithm. The study is only limited to the 17 Baybayin characters trained for recognition and classification of Baybayin words. It is not trained with images of Baybayin characters with accents and primarily geared towards detecting Tagalog words. The image dataset is also constrained to clear photographs and scans depicting individual Baybayin words against clean backgrounds, under adequate lighting conditions, to enable precise preprocessing and character segmentation.

II. Methodology

The system's goal in constructing the algorithm model would be to detect and recognize Baybayin. It develops from existing technologies like image character recognition and image processing using an SVM classifier. The algorithm model is expected to detect Baybayin and convert each character to their respective Latin character classification. After conversion, all the converted characters should be combined with their respective equivalent words. This system is created to recognize Baybayin words by detecting the accents and Baybayin characters, which will solve the issue of the difficulty in understanding Baybayin. The primary function of this model is to detect and recognize Baybayin words and translate them into the equivalent Filipino word.

A. Research Methodology

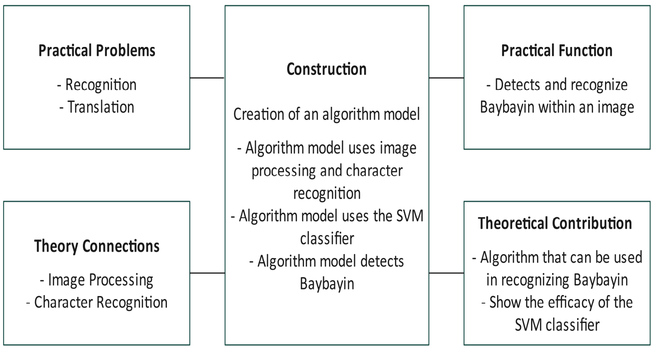


Fig. 1. Constructive Research Methodology

B. Conceptual Framework

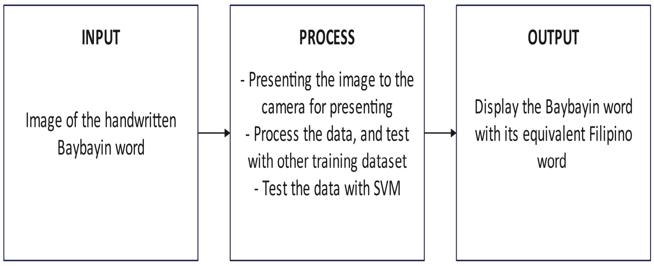


Fig. 2. Conceptual Framework

Figure 2 shows the conceptual framework of the proposed hardware prototype. An image of the handwritten Baybayin word is presented to the camera for processing. In this case, the camera will be connected to a microprocessor, a Raspberry Pi, to save the captured image locally. It will be tested on the trained SVM model and other stored testing datasets. The output would show the Baybayin word and its corresponding character in the Filipino word.

C. Hardware Development

The system comprises the following hardware components: a 5 MP CSI camera module, a 7-inch touchscreen display, a Raspberry Pi 4 Model B, and a USB Type-C power adapter, as seen in Fig 3. Given the model's need for image inputs, the camera module will be connected to the Raspberry Pi. This camera module captures good-quality images and ensures compatibility with the Raspberry Pi. Lastly, after the Raspberry Pi processes the data, the 7-inch touchscreen display will display the results produced by the model.

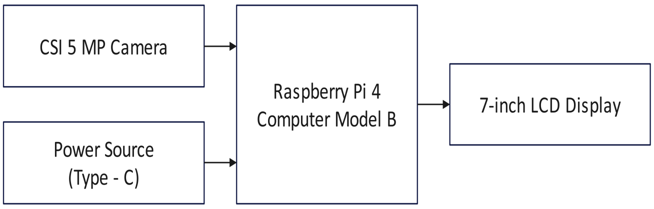


Fig. 3. Proposed Prototype Block Diagram

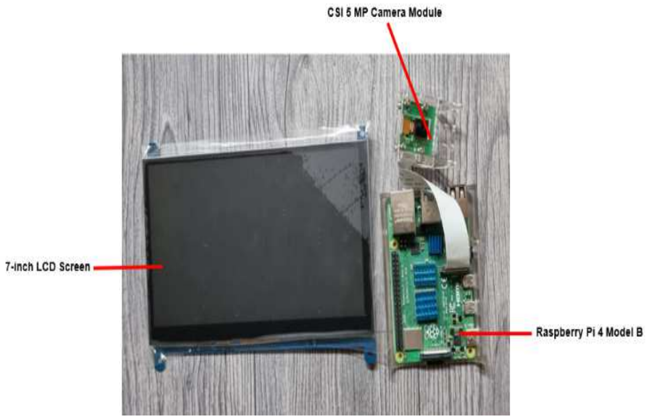


Fig. 4. Physical Hardware Materials

D. System Flowchart

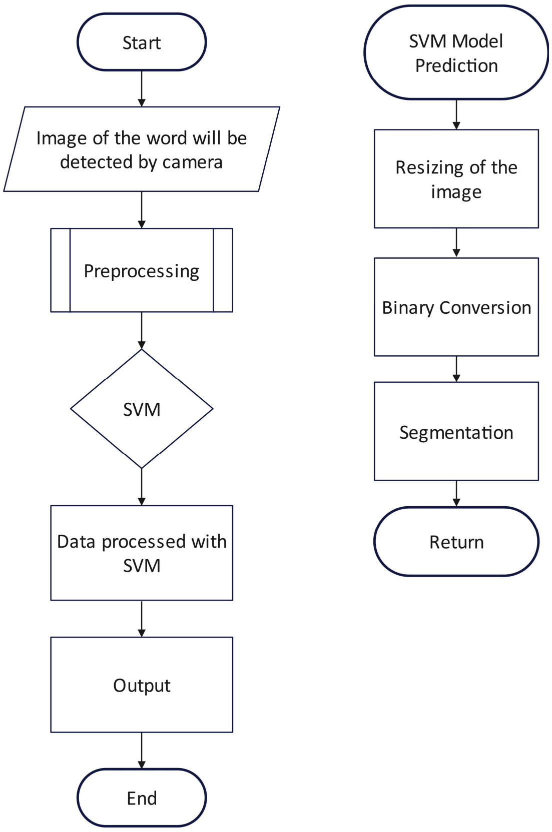


Fig. 5. System Flowchart

The flowchart depicted in Fig 5 illustrates the system's process. First, the camera will capture the image of the word. The image will undergo preprocessing and be fed into the SVM algorithm, and then its data will be processed with SVM. In the preprocessing stage, the image will be first cleaned and then resized to match the matrix of the training dataset, which would be 50x50. From there, it will undergo binary conversion to allow the data to be uniform and then undergo segmentation. Segmentation separates each character from each other for individual translation. After preprocessing, it will now undergo the prediction from the SVM models. After the image undergoes prediction, the system will combine the characters, and the combined output will be shown on the monitor.

III. Results and Discussions

A. Data Gathering and Preprocessing

Data collection involves gathering information from diverse sources, including online databases like Kaggle and GitHub, and personally captured images, creating a comprehensive dataset. The Baybayin Character dataset comprises 38,390 images, while the Baybayin Word dataset for testing contains 120 images. After data collection, preprocessing is performed. Images are loaded in grayscale, followed by Gaussian blur to reduce noise. Otsu's thresholding method is applied for binarization; random image rotation enhances perspective diversity. All images are resized to 50x50 pixels for consistency. Flattening 2D image arrays into 1D vectors facilitates classification, and data is prepared for model fitting through feature-label pairing, shuffling, and splitting into test and train sets.

B. Prototype Setup

After completing the development of the hardware and software components, the device was tested several times using handwritten Baybayin words. The camera module was positioned to face the paper with the Baybayin word from the back of the screen, ensuring a clear view. It was set at a high angle to ensure a direct and clear focus on the text, capturing detailed close-up images of the written Baybayin word.

C. Model Training

The training phase utilizes a Support Vector Machine (SVM) to analyze the labeled data. This process begins with a grid search to determine the optimal hyperparameters, finding C=25 and gamma=1e-08 as the best settings for the Radial Basis Function (RBF) kernel. In contrast, the linear kernel performs best with C=25 and gamma=0.01. The parameter C is crucial as it determines the compromise between minimizing the training error and simplifying the model complexity to ensure generalization. A higher value of C may lead to a model that fits the training data very well but at the risk of overfitting. While the decision boundary's curvature is influenced by gamma, a lower gamma produces a more linear decision boundary. A higher gamma allows the boundary to adapt more closely to the data. Initially, training with just the RBF kernel for all characters leads to a relatively low accuracy. This results in two kernels being utilized to handle the Baybayin characters more effectively: the linear kernel and the RBF kernel. The linear kernel is straightforward and effective for more linearly separable characters like "Ha", making it ideal for complex but structurally distinct characters. In contrast, the RBF kernel, which can handle varied forms and separations in the data due to its flexibility in shaping the decision boundary, is better suited for more straightforward characters with specific strokes.

D. Model Evaluation

After training the SVM models, precision, F1 scores, recall, and accuracy were evaluated for each Baybayin character. The linear and RBF kernels were compared regarding their performance with complex characters, as shown in Tables 1 and 2. The linear kernel achieved higher overall accuracy at 93.59% compared to 78.74% for the RBF kernel, indicating its better performance on characters with simpler strokes.

The RBF kernel demonstrated high recall in most classes, indicating few missed positive cases, but it had lower precision, resulting in more false positives than the linear model. The F1-score, which balances precision and recall, is higher on the linear kernel, except for the character Ma, which is equal in value to the RBF kernel. While the linear kernel excelled in accuracy and related metrics in this context, the RBF kernel tends to overfit the training dataset.

TABLE I. RBF Kernel on the Second Model

Character	Precision	Recall	F1-score
E/I	53.25%	98.83%	69.21%
Ha	94.16%	90.89%	92.49%
Ma	75.96%	95.39%	84.58%
Ta	99.06%	50.12%	66.56%
Ya	87%	64.13%	73.83%
Average	81.89%	79.87%	77.34%
Accuracy		78.74%	

TABLE II. LINEAR KERNEL ON THE SECOND MODEL

Character	Precision	Recall	F1-score
E/I	96.20%	98.06%	97.12%
Ha	97.14%	97.14%	97.14%
Ma	89.76%	90.42%	84.58%
Ta	94.95%	94.09%	94.52%
Ya	90.38%	89.52%	89.95%
Average	93.69%	93.85%	93.76%
Accuracy		93.59%	

TABLE III. RBF KERNEL ON THE FIRST MODEL

Character	Precision	Recall	F1-score
A	96.70%	82.66%	89.13%
Ba	99.45%	97.32%	98.37%
Da/Ra	96.77%	97.74%	97.25%
Ga	94.84%	94.84%	96.98%
Ka	98.74%	98%	98.36%
La	98.26%	98.50%	98.38%
Na	99.25%	99.25%	98.88%
Nga	97.09%	95.70%	96.39%
O/U	100%	95.74%	97.83%
Pa	91.16%	96.08%	93.56%
Sa	93.56%	98.78%	96.09%
Wa	98.99%	98.49%	98.73%
Average	97.07%	96.40%	96.66%
Accuracy		96.88%	

Table 3 summarizes the SVM model's classification performance using an RBF kernel for more complex Baybayin characters. The table demonstrates that the model achieves high accuracy, with an overall accuracy of 96.9%. It consistently achieves precision and recall scores above 95% for 8 of 12 character classes, indicating reliable recognition. However, precision is higher than recall for characters like "A," primarily due to similarities with the character "Pa." The F1-scores, which measure consistency in capturing accurate labels, consistently reach over 89%, affirming reliable classification by the SVM's nonlinear decision boundaries.

Some classes with more variance, such as Pa and Sa, have precision dropping below 95%, potentially due to the similarity of strokes with the other characters. Despite this, recall rates remain consistent, with characters like Ga displaying high recall but lower precision, indicating broader classification at the cost of some false labels. On the other hand, O/U exhibits 100% precision but lower recall, showing returns of false labels.

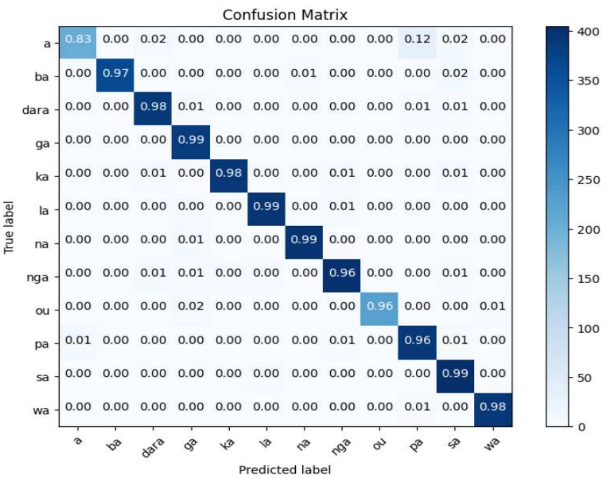


Fig. 6. RBF Kernel SVM Model Confusion Matrix

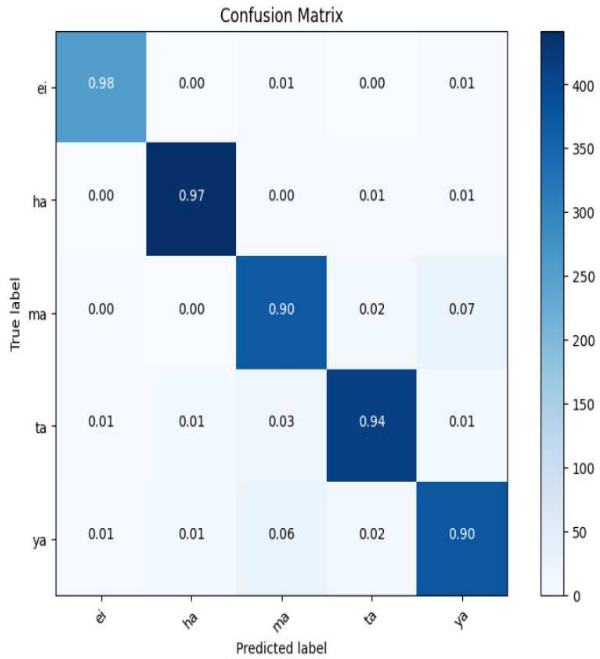


Fig. 7. Linear Kernel SVM Model Confusion Matrix

The confusion matrix for the first SVM model is displayed in Fig 6; it provides a detailed breakdown of the performance of the SVM model with an RBF kernel for each Baybayin character. The values along the diagonal show the true label prediction for the characters. For instance, 99% of the actual Ga samples get correctly labeled, with only a small amount of misclassification from other characters. However, characters like "A" display 83% recall, meaning nearly 17% got misclassified. Checking the off-diagonal cells reveals some ambiguity in distinguishing "A" from Pa, likely due to almost identical strokes. Lower scores for Nga, O/U, and Wa indicate more errors, though still under 5%.

While Fig 7 presents the confusion matrix of the linear SVM model. The diagonal rows have the most values at over 0.90, showing generally accurate label prediction. For instance, 98% of actual E/I cases were correct. However, 2% of those were still mixed up as Ma or Ya. Likewise, around 3% of true Ma labels get mistaken for Ta, Ya, or E/I - pointing to the confusion between the characters.

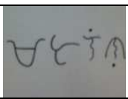
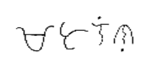
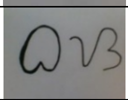
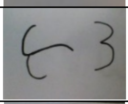
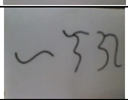
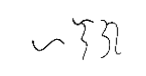
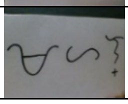
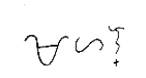
E. Statistical Treatment of Data

OpenCV was used to acquire Baybayin word images for word prediction. Frames from the video stream were captured and saved locally. Image enhancement involved grayscale conversion before Otsu's binarization. The processed images were resized to 50x50 to match the expected input of the SVM models. The prediction process processes the segmented Baybayin word, identifies the characters, dots, and signs, and then undergoes prediction under the two SVM models. Predictions go through weighted voting. The system handles confusion with challenging character pairs and recombines modified characters based on accent placement. The closest matching word from a Filipino dictionary is selected based on similarity, with additional possible words considering variations with Da/Ra, E/I, and O/U characters.

TABLE IV. GUIDELINE TO THE CHARACTER CHANGE BASED ON THE ACCENTS

	Position	**Change in the character**
Dot	Above	Removes the A in the character and replaces it with E/I. Does not change the character "A".
	Below	Removes the A in the character and replaces it with O/U. Does not change the character "A".
Plus Sign	Below	Removes the A in the character. Does not affect the character "A".

TABLE V. SAMPLE ACCURACY TESTING OUTPUT

Input image	**Processed image**	**Predicted word**	**True Label**
		Matalino	Matalino
		Basa	Basa
		Tao	Tao
		Halaga	Halaga
		Mahal	Mahal

A Table 5 provides 5 sample input images, the processed state of the input image, and the predicted word. The input image should meet the following requirements to ensure accurate recognition: The Baybayin word should be written against a white background in thick strokes to facilitate clear segmentation. Additionally, ample space should be left between each character to prevent merging during the segmentation process. The dot accent should only be written below or above the character, while the plus sign accent should only be below a character. Lastly, only the word and paper should be seen, objects and backgrounds are converted into noise if caught in the input image, thus leading to inaccurate predictions. The recognition accuracy was computed using the formula below:

$$Accuracy = \frac{Correctly\ Predicted\ Images}{Total\ Images} * 100\% \qquad (1)$$

A Baybayin word prediction is considered accurate if all characters are correctly labeled, proper modifications based on the detected accents, and the combined translated characters match the expected word. The recognition system was tested on 120 Baybayin word images. Out of 120 images, 109 came out as correctly predicted; this results in a 90.83% recognition accuracy. This high recognition rate indicates that the SVM models can reliably handle various Baybayin word images, including those with accents.

IV. CONCLUSION AND RECOMMENDATION

This research focused on developing SVM models to create an OCR system to identify Baybayin words. The study demonstrated the effectiveness of SVM, notably the linear kernel, in accurately recognizing modern and traditional Baybayin characters. The hardware prototype, featuring a Raspberry Pi and camera module, successfully captured Baybayin word images for processing, showcasing the practical application of this technology. With a recognition accuracy of 90.83%, the system has the potential to make Baybayin more accessible and promote its preservation, contributing to conserving Filipino cultural heritage.

To build on these findings, future researchers can focus on expanding the dataset to include Baybayin characters with accents, various handwriting styles, and different accents. Developing a mobile app would enhance accessibility and engagement, making Baybayin more attractive to a broader audience. An easy-to-use app would enable users to interact with the recognition system, helping them build a deeper connection to the Baybayin script and its cultural significance.

Lastly, future researchers should explore methods to improve the accuracy of character and accent recognition and enhance the overall performance and usability of the system.

ACKNOWLEDGMENT

We would like to sincerely thank all the groups and individuals who have contributed to the completion of this thesis. First and foremost, we are grateful to God for His wisdom and grace, which have been the foundation of our strength during this academic journey.

We owe a great deal to Engr. Meo Vincent C. Caya, our thesis advisor for his invaluable advice, insightful feedback, and commitment to excellence. His knowledge and support have significantly shaped our research.

A special appreciation goes to our panel members, Engr. Analyn N. Yumang, Dr. Jocelyn F. Villaverde, and Engr. Noel B. Linsangan. Your detailed reviews, insightful critiques, and helpful suggestions have improved this thesis's quality.

We are also grateful to everyone who has helped with this project. Your support has been extremely valuable. Thank you all for being an essential part of this academic journey.

REFERENCES

[1] R. Andrei Vilvar, D. C. Shawn Hammond, F. Mark Santos, and H. S. Alar, "BAYBAYIN SCRIPT WORD RECOGNITION AND TRANSLITERATION USING A CONVOLUTIONAL NEURAL NETWORK." [Online]. Available: https://ssrn.com/abstract=4004853.

[2] A. T. Camba, "Baybayin: The Role of a Written Language in the Cultural Identity and Socio-Psychological Well-Being of Filipinos," Mar. 2021.

[3] M. J. A. Daday, "Recognition of Baybayin Symbols (Ancient Pre-Colonial Philippine Writing System) using Image Processing," International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 1, pp. 594–598, Feb. 2020, doi: 10.30534/ijatcse/2020/83912020.

[4] M. V. C. Caya, R. G. Maramba, J. S. D. Mendoza, and P. S. Suman, "Characterization and Classification of Coffee Bean Types using Support Vector Machine," in 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology,

Communication and Control, Environment, and Management (HNICEM), IEEE, Dec. 2020, pp. 1–6. doi: 10.1109/HNICEM51456.2020.9400144.

[5] C. C. Hortinela, J. R. Balbin, J. C. Fausto, P. Daniel C. Divina, and J. P. T. Felices, "Identification of Abnormal Red Blood Cells and Diagnosing Specific Types of Anemia Using Image Processing and Support Vector Machine," in 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/HNICEM48295.2019.9072904.

[6] Ma. C. Q. Jasmin, M. J. F. Dela Cruz, and A. N. Yumang, "Detection of Forged Handwriting Through Analyzation of Handwritten Characters Using Support Vector Machine," in 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), IEEE, Sep. 2022, pp. 1–5. doi: 10.1109/IICAIET55139.2022.9936769.

[7] K. M. A. Bejerano, C. C. Hortinela IV, and J. J. R. Balbin, "Rice (Oryza Sativa) Grading classification using Hybrid Model Deep Convolutional Neural Networks - Support Vector Machine Classifier," in 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), IEEE, Sep. 2022, pp. 1–6. doi: 10.1109/IICAIET55139.2022.9936869.

[8] A. J. C. Arceo, R. Y. N. Borejon, M. C. R. Hortinela, A. H. Ballado, and A. C. Paglinawan, "Design of an E-Attendance Checker through Facial Recognition using Histogram of Oriented Gradients with Support Vector Machine," in 2020 IEEE 8th International Conference on Smart City and Informatization (iSCI), IEEE, Dec. 2020, pp. 1–5. doi: 10.1109/iSCI50694.2020.00008.

[9] J. C. Gonzales, J. R. G. Garcia, and J. F. Villaverde, "BMI Estimation from 2D Face Images Using Support Vector Machine," in 2022 6th International Conference on Communication and Information Systems (ICCIS), IEEE, Oct. 2022, pp. 118–123. doi: 10.1109/ICCIS56375.2022.9998152.

[10] E. J. L. Aguilar, G. K. P. Borromeo, and J. Flores Villaverde, "Determination of Pineapple Ripeness Using Support Vector Machine for Philippine Standards," in 2021 IEEE 7th International Conference on Control Science and Systems Engineering (ICCSSE), IEEE, Jul. 2021, pp. 283–287. doi: 10.1109/ICCSSE52761.2021.9545163.

[11] C. C. Hortinela, J. R. Balbin, J. C. Fausto, J. E. C. Espanillo, and J. K. P. Padilla, "Detection of Staleness in Raw Chicken Meat Due to Salmonella spp. and Escherichia Coli Bacteria Using Metal Oxide Gas Sensor With Support Vector Machine," in 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE, Dec. 2020, pp. 1–5. doi: 10.1109/HNICEM51456.2020.9399997.

[12] Y. B. L. Macatol, S. V. S. Geronimo, C. C. Hortinela, M. C. Gomez, R. P. Tongol, and J. C. Fausto, "Detection of Chronic Myelogenous Leukemia Applying Support Vector Machine and Feature Extraction," in 2023 Seventh International Conference on Advances in Biomedical Engineering (ICABME), IEEE, Oct. 2023, pp. 44–49. doi: 10.1109/ICABME59496.2023.10292943.

[13] G. L. H. Tuazon, H. M. Duran, and J. F. Villaverde, "Portable Sigatoka Spot Disease Identifier on Banana Leaves Using Support Vector Machine," in 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE, Nov. 2021, pp. 1–6. doi: 10.1109/HNICEM54116.2021.9731808.

[14] M. R. Abalos and J. F. Villaverde, "Fresh Fish Classification Using HOG Feature Extraction and SVM," in 2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE, Dec. 2022, pp. 1–5. doi: 10.1109/HNICEM57413.2022.10109442.

[15] B. H. Belay, T. A. Habtegebrial, and D. Stricker, "Amharic Character Image Recognition," in 2018 IEEE 18th International Conference on Communication Technology (ICCT), IEEE, Oct. 2018, pp. 1179–1182. doi: 10.1109/ICCT.2018.8599888.

[16] R. Ramanathan, S. Ponmathavan, N. Valliappan, L. Thaneshwaran, A. S. Nair, and K. P. Soman, "Optical character recognition for English and Tamil using support vector machines," in ACT 2009 - International Conference on Advances in Computing, Control and Telecommunication Technologies, 2009, doi: 10.1109/ACT.2009.155.

[17] G. Katiyar, "Off-Line Handwritten Character Recognition System Using Support Vector Machine," American Journal of Neural Networks and Applications, 2017, doi: 10.11648/J.AJNNA.20170302.12.

[18] H. Byun and S. W. Lee, "A SURVEY ON PATTERN RECOGNITION APPLICATIONS OF SUPPORT VECTOR MACHINES," *http://dx.doi.org/10.1142/S0218001403002460*, 2011, doi: 10.1142/S0218001403002460.

[19] R. Pino, R. Mendoza, and R. Sambayan, "Optical character recognition system for Baybayin scripts using support vector machine," PeerJ Computer Science, 2021, doi: 10.7717/PEERJ-CS.360.

[20] P. Sok and N. Taing, "Support Vector Machine (SVM) based classifier for Khmer Printed Character-set Recognition," 2014 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2014, 2014, doi: 10.1109/APSIPA.2014.7041823.

[21] J. G. O. Ebora, J. C. N. Espanol, and D. A. Padilla, "Text Classification of Facebook Messages Using Multiclass Support Vector Machine," in 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, Oct. 2022, pp. 1–6. doi: 10.1109/ICCCNT54827.2022.9984554.

[22] J. X. Dong, A. Krzyzak, and C. Y. Suen, "An improved handwritten Chinese character recognition system using support vector machine," Pattern Recognition Letters, 2005, doi: 10.1016/J.PATREC.2005.03.006.

[23] A. M. P. Ligsay, J. B. Rivera, and J. F. Villaverde, "Optical Character Recognition of Baybayin Writing System using YOLOv3 Algorithm," in 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), IEEE, Sep. 2022, pp. 1–5. doi: 10.1109/IICAIET55139.2022.9936792.