Kernel Function Exploration in Support Vector Machine for Digit Handwritten Recognition

Original article

Received: Sept. 6,2025
Revised: Jan. 13, 2026
Accepted: Mar. 10,2026
Published: Apr. 20,2026

Natthapat Markpradit, Prasong Kitidamrongsuk, Jutatip Sillabutra,
Pichitpong Soontornpipit
Department of Biostatistics, Faculty of Public Health, Mahidol University, Thailand

Abstract

Handwriting classification plays a vital role in biomedical informatics, particularly for digitizing handwritten records and automating data entry. This technology can significantly benefit Thai public health by improving the efficiency of medical record digitization enhancing disease surveillance and outbreak response, and supporting research and development. These applications can lead to better healthcare access, improved patient outcomes, and more impactful health promotion and disease prevention efforts. This study aims to compare the performance of the difference kernel functions in Support Vector Machine, including linear kernel, polynomial kernel and Radial Basis Function (RBF) apply to the MNIST dataset, and the benchmark dataset EMNIST Digits, the extended size of MNIST. This study investigates the efficacy of three kernel function in Support Vector Machine (SVM) classification models including linear kernel, polynomial kernel and Radial Basis Kernel (RBF) to recognize digit handwriting. These classification methods initially to fitting and validate in "MNIST" as the starting dataset. The performance metrices resulting from these methods will be calculated together with the loss function to obtain the result of error analysis. These methods also will be implemented in "EMNIST Digits", the extended dataset of MNIST as a generalization indicator. The tools for running model of training part and test part evaluate by Python programming (Google Colab Pro), supported by high-RAM NVIDIA T4 and/or GPU hardware accelerators to enhance processing efficiency. The result of model performance metrices analysis from R programming for statistical computation.

The results from three repeated measurements revealed that the average test accuracy of SVM by using linear kernel, polynomial kernel and RBF on the MNIST and EMNIST Digits datasets achieved as 94.28% and 91.90%, 98.42% and 99.14%, 98.67% and 99.25%, respectively, whereas hyperparameter tunning execution time on the MNIST and EMNIST Digits datasets are 50.61,2084.86, 556.99 and 62.93, 654.94,3936.94 seconds respectively.The results revealed that the linear kernel of SVM achieved the shortest execution time during hyperparameter tuning, making it suitable for applications with limited computational resources. In contrast, the RBF kernel of SVM demonstrated the highest classification accuracy but required significantly longer processing time. Notably, the polynomial kernel offered a balanced trade-off, yielding competitive performance with moderate computational demands. This study highlights the importance of selecting an appropriate kernel function based on the specific requirements of the task, such as accuracy, interpretability, and computational efficiency.

Keywords: Handwriting digit recognition, Support Vector Machine, Linear kernel, Polynomial kernel, Radial Basis Function kernel

Corresponding author: Pichitpong Soontornpipit
Email: pichitpong.soo@mahidol.ac.th

1. Introduction

Handwriting character recognition (HCR) is the extension of optical character recognition (OCR) technology, the innovation that converts printed text to machine-editable text. HCR is classified as two types: offline character recognition and online character recognition. Offline character recognition involves processing handwriting data from physical documents. This means that the system analyzes an image or a scanned document containing the handwritten text but is not limited to any situation where the handwriting data is already in a static form.

Online character recognition also involves capturing a sequence of coordinates representing the pen's movement in real time. This system analyzes the trajectory, speed, and pressure of the pen to recognize characters. While it requires specialized hardware for real-time data capture, such as tablet PCs, smartphones, or digital pens, it generally offers higher accuracy due to the additional temporal information in real-time as it is being written. However, the cursive handwriting is challenging, as stated in Sayre's paradox in segmentation task[1]. In cursive, or connected handwriting, it's often unclear where one-character ends and another begins.

This makes it difficult to isolate individual characters for recognition, as the system cannot determine character boundaries without first understanding what the characters are. In 1976, Shelia Guberman successfully applied the pattern recognition knowledge to recognize a cursive script by analyzing metastrokes, which refer to larger, composite units of handwriting that encompass multiple individual strokes forming a meaningful shape or pattern, such as parts of letters or entire characters[2-3]. This also contributed to the holistic-analytical recognition of handwritten text as a character feature identification[4], which reported that it is a viable alternative to the popular analytical (character segmentation-based) approach to handwriting recognition, as the error rates achieved are comparable to those of multi-writer recognition systems for high-quality input pages.[5].

Nowadays, artificial intelligence and powerful computers are leading to intensive handwriting recognition research, which could receive the high-resolution input and deliver the high accuracy of recognition together with modern sophisticated algorithms such as neural networks (NN), hidden Markov chain models (HMMs), fuzzy algorithms, and natural language processing(NLP).[6]These advancements have further boosted OCR technologies, enabling more efficient and accurate text recognition across various languages and scripts[5]. Despite advancements in OCR and successful applications for commercial use, Handwritten Character Recognition (HCR) still faces significant limitations such as variability in handwriting styles, connected or cursive writing, limited datasets, and multilingual challenges. These limitations can impact the accuracy of HCR systems, potentially affecting their effectiveness in applications like medical record digitization.

Therefore, it is essential to explore approaches that are interpretable and balanced between model accuracy and computational cost. Support Vector Machines (SVM) are well known for their ability to manage overlapping features and perform reliably across datasets with complex and diverse characteristics. Moreover, choice of kernel is a critical factor in the performance of SVM, as it determines the way of input data are transformed into a higher-dimensional feature space where linear separation becomes possible. The selection of an appropriate kernel, such as linear, polynomial, or radial basis function (RBF) kernel, directly influences the model's ability to capture underlying patterns, generalize across datasets, and manage computational complexity.

Medical documents such as medical prescriptions, surgical notes, and patient charts, especially in rural areas of Thailand, are still handwritten, requiring time and labor for entering those data into public health databases, compromising data integrity for Thai public health policymakers.

This study focuses on selecting appropriate kernel functions to improve handwritten digit recognition contribute to innovations that convert handwritten records into

electronic form. This conversion can significantly benefit Thailand's public health by improving healthcare efficiency and enabling population-level data analysis for initiatives and research.

Moreover, the future goal of handwritten digit recognition is to develop a preliminary assessment tool for identifying Parkinson's disease symptoms in patients with shaky hands on a global scale, which is a significant part of building sustainable health and well-being for all ages, as digits serve as an international language.

This study aims to compare the efficacy and execution time for hyperparameter tunning of three kernel functions in SVM classification models including linear kernel, polynomial kernel and RBF kernel to recognize digit handwriting.

2. Research Methodology

Ethical considerations

Ethical exemption approval was obtained from the ethics committee of the Faculty of Public Health, Mahidol University.

Dataset

This study is a comparative study with error analysis based on investigations of two benchmark datasets in computer vision, including

1. MNIST (Modified National Institute of Standards and Technology) dataset consists of 60,000 training and 10,000 testing handwritten digit images (10 classes: numbers 0-9), each class represented in 28x28 greyscale pixels of size 7. This dataset was obtained from https://yann.lecun.org/exdb/mnist/index.html with idx3-ubyte and idx1-ubyte formats for images and labels, respectively.
2. EMNIST Digits dataset containing 280,000 samples of digits, the extension of the MNIST dataset developed by Cohen et al., 2017 (10 classes: numbers 0-9), in order to evaluate by machine learning models on a more extensive and class-balanced collection of handwritten digits, allowing models to be tested more fairly across different digit categories and avoiding bias from uneven class distributions, consists of 240,000 training images and 40,000 testing images, each 28x28 greyscale pixels in size. This dataset was obtained from https://biometrics.nist.gov/cs_links/EMNIST/gzip.zip.

Software and algorithms for model fitting and evaluation

The model training and evaluation in this study were performed by using Python 3 and the scikit-learn library, a widely used in machine learning framework. Computations were executed on a high-RAM NVIDIA T4 GPU with CPU support, accelerated through CUDA, a GPU-based parallel computing framework, to enhance computational efficiency. All experiments were conducted within the Google Colab Pro environment. The Support Vector Machine (SVM) algorithm was applied with three kernel functions: linear, polynomial, and radial basis function (RBF). The linear kernel serves as a baseline due to its computational efficiency and suitability for linearly separable data. The polynomial kernel expands the feature space, allowing the model to capture non-linear relationships that involve interactions between features. The RBF kernel, the most widely used, provides high flexibility in modeling complex non-linear boundaries by adjusting the gamma parameter to control the influence of individual data points. These kernel functions were compared to highlight their different capabilities in handling digit recognition tasks, emphasizing the importance of kernel selection in the implementation of SVM.

Once the best hyperparameters were obtained via GPU-accelerated grid search, SVM models with linear, polynomial, and RBF kernels were trained on each dataset. For performance evaluation, stratified k-fold cross-validation was applied with k ranging from 3 to 10, ensuring balanced class representation in all folds.

Model performance was assessed by using standard classification metrices: accuracy, precision, recall, and F1-score. All metrices were calculated by using macro-averaging to obtain account for class imbalance in multi-class classification. These kernel functions for SVM model can be shown in Table 1.

Table 1 Kernel functions formula *and parameter* in this study

Kernel function type	Formula	Parameter(s)
Linear Kernel	$K(x, x_i) = \sum_{j=1}^{p} x_i\, x_j$	- C- Regularization
Polynomial kernel of degreed	$K(x, x_i) = \left(1 + \sum_{j=1}^{p} x_{ij}\, x_{i'j}\right)^{d}$	- C-Regularization - Degree (d)
Radial Basis function (RBF)	$K(x, x_i) = \exp\left(-\gamma \frac{\left\|\|x - x_i\|\right\|^2}{\sigma^2}\right)$	- C-Regularization - Degree (d) - Gamma (γ)

Data preprocessing

Two benchmark datasets, MNIST and EMNIST Digits, have been split into training and testing datasets. Image pixel data was normalized in the range of -1 to 1, followed by PCA (Principal Component Analysis), which was conducted for dimension reduction while retaining the components that preserved 90% of variance to save computational cost. The data were subsequently converted from NumPy arrays to CuDF arrays, a GPU-compatible data structure designed for integration with CuDF and CuML libraries optimized for GPU acceleration.

This conversion aimed to make the data type compatible with the GPU-based library framework, which enhances computational efficiency during grid search hyperparameter tuning by leveraging the parallel processing capabilities of GPUs as hardware accelerators.

Hyperparameter tuning

Hyperparameter optimization was conducted by using GPU-accelerated grid search, enabling faster tuning across different kernel-specific parameters:

1. For the linear kernel: the tuning is regularization parameter C
2. For the polynomial kernel: the tunings are parameters C, and degree of polynomial (d)
3. For the Radial Basis Function (RBF) kernel: the tuning parameters are C and radius γ (gamma)

3. Results

Distribution of Labels in dataset

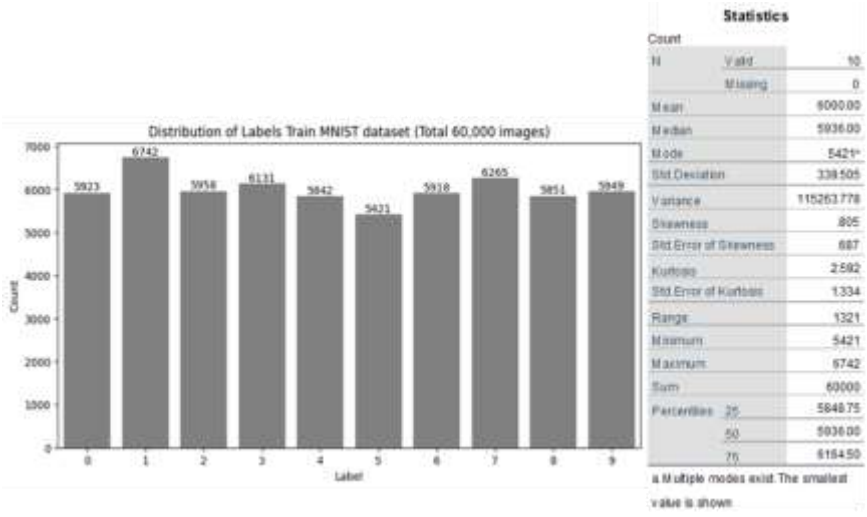


Figure 1 Distribution of Labels in MNIST Training dataset

Figure 1 presents the histogram and descriptive statistics table for the image of each label in the MNIST training dataset, comprising a total of 60,000 images, have counts ranging from 5,421 for digit '5' to 6,742 for digit '1' (range = 1,321), whereas the mean equals 6,000 and the SD is 339.505. This can assume the uniformly distribution of labels across the ten classes (digits 0-9), even if there is some variation of numbers per class.

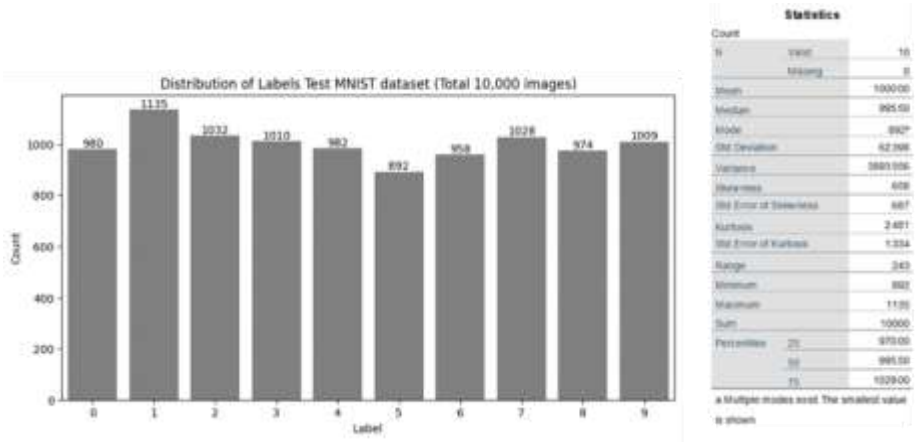


Figure 2 Distribution of Labels in MNIST test dataset

Figure 2 show the result in the test dataset is also similar to the training dataset, containing a total of 10,000 images with counts ranging from 892 for digit '5' to 1,135 for digit '1' (range = 243), whereas the mean equals 1,000 and the SD is 62.398. This result also assumes the uniformly distribution of labels across the ten classes (digits 0-9), with even some variation of numbers per class.

The histogram of Figure 3, comprises a total of 240,000 images, contains 24,000 images for each class of the EMNIST Digits training dataset. This is clearly indicated by the uniformly distribution of labels across the ten classes (digits 0-9), without any variation of numbers per class. Similarly, the EMNIST Digits test dataset containing 40,000 images with 4,000 images for each of the ten classes.

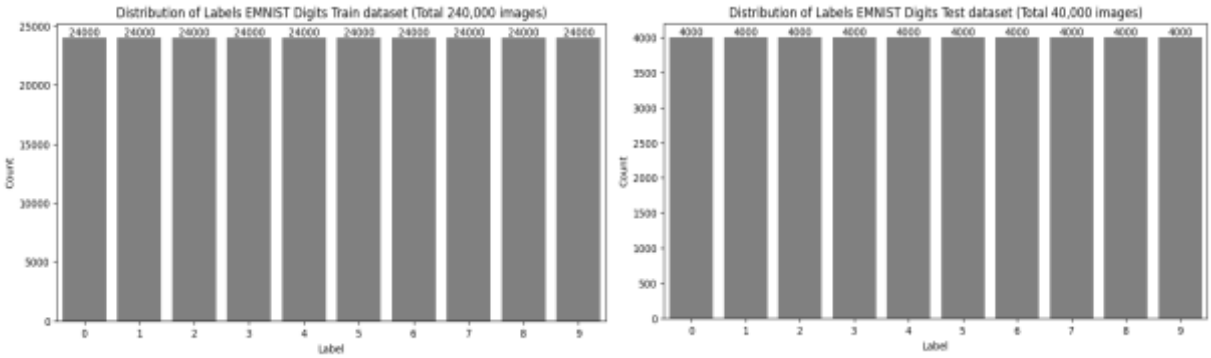


Figure 3 Distribution of Labels in EMNIST Digits train & test dataset

Table 2 Result of hyperparameter tunning of three kernel functions

Kernel function	Dataset	C (regularization)	Degree	Gamma	Execution Time (S)
Linear	MNIST	0.0727895384398315	-	-	50.61
Polynomial	MNIST	2.1544	3	-	2084.56

Table 2 Result of hyperparameter tunning of three kernel functions (etc.)

Kernel function	Dataset	C (regularization)	Degree	Gamma	Execution Time (S)
RBF	MNIST	10	-	0.01	556.99
Linear	EMNIST Digits	3.2903445623126677	-	-	62.93
Polynomial	EMNIST Digits	10	3	-	654.94
RBF	EMNIST Digits	10	-	0.01	3936.45

Table 2 represents the hyperparameter tuning of the three kernel functions linear, polynomial, and RBF across two standard benchmark datasets for handwritten digit classification: MNIST and EMNIST Digits. The table includes each model's selected regularization parameter C, kernel-specific parameters (including degree for polynomial and gamma for RBF), and total execution time in seconds required for grid search hyperparameter tuning on the GPU-enabled environment.

In the MNIST dataset, the RBF kernel achieved optimal performance at C=10, with gamma = 0.01, requiring 556.99 seconds for complete tuning. In contrast, the polynomial kernel (degree = 3) on the same dataset with C = 2.154 took significantly longer, at 2,084.56 seconds, suggesting that polynomial mapping leads to higher computational cost.

For the EMNIST Digits dataset, although it contains four times more samples than MNIST dataset, the total execution time for polynomial kernel tuning (C=10, degree=3) was only 654.94 seconds, which is lower than MNIST's polynomial kernel time. This unexpected trend may reflect better convergence behavior, more effective class separability in PCA-reduced space, or GPU memory caching effects benefiting the second experiment in sequence.

The linear kernel was the fastest across both datasets, as this required only 50.61 seconds and 62.93 seconds for MNIST and EMNIST Digits, reflecting its computational efficiency and the absence of non-linear transformation overhead. While linear kernels provide a useful baseline, their performance was consistently lower than the polynomial RBF kernel in terms of classification accuracy, precision, recall, and F1-score.

Principle Component Analysis (PCA)

PCA was conducted to reduce data dimension while still preserving the essential structure of the data. This study set 90% of total variance explanation. The number of dimensions was reduced from 784 to equal the number of PCA vectors. This result obtained the number of PCA vectors for the MNIST and EMNIST Digits datasets, which are 60 and 87 components respectively. Figures 4 and 5 represent 2-dimensional PCA projections of MNIST and EMNIST digits training dataset across all digit classes (0–9).

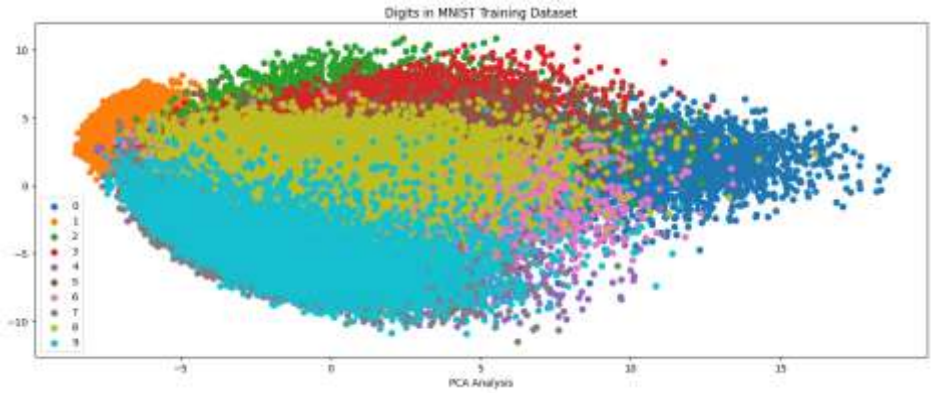


Figure 4: 2-Dimensional PCA projection of MNIST training dataset

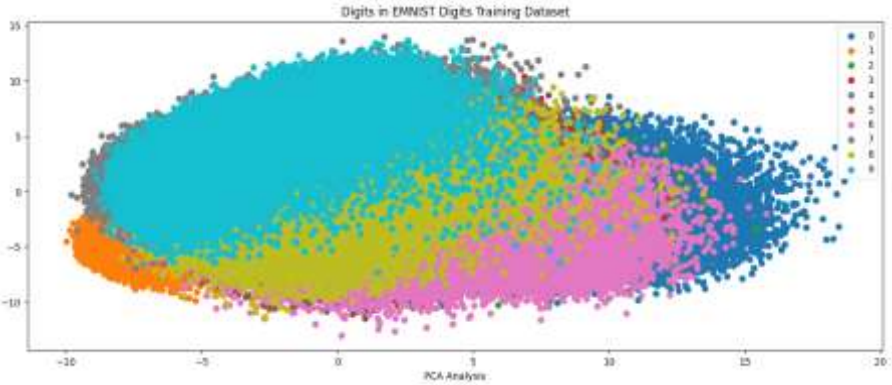


Figure 5:2-Dimensional PCA projection of EMNIST Digits training dataset

The clearly separated clusters with minimal overlapping among digit classes indicate the success of dimensionality reduction. In the EMNIST Digits projection, distinct groupings such as those representing digits 0, 1, and 8 suggest that the PCA retained sufficient discriminative information to preserve inter-class variance.

In contrast, the PCA projection of MNIST dataset reveals significant class overlap, particularly in the central region of the plot. This suggests that the lower number of retained components may have led to loss of discriminative features, making the dataset less separable in reduced dimensions and potentially complicating the training process for kernel-based classifiers.

Model performance metrices

The results from three repeated measurements revealed that the average validation and test model performance of SVM with the kernel: linear, polynomial and RBF based on the MNIST and EMNIST Digits datasets can be summarized below by the Table 3.

Table 3 Model performance metrices results

Kernel function	Dataset	Validation Accuracy (± SD)	Test Accuracy (± SD)	Validation Precision (± SD)	Test Precision (± SD)	Validation Recall (± SD)	Test Recall (± SD)	Validation F1-Score (± SD)	Test F1-Score (± SD)
SVM-Linear (Best K = 9)	MNIST	93.88% (± 0)	94.27% (±5.77E-05)	93.8% (± 0)	94.18% (±5.77E-05)	93.80% (± 0)	94.18% (±5.77E-05)	93.79% (± 0)	94.17% (±0.00011547)
SVM-Linear (Best K = 8)	EMNIST Digits	92.09% (±5.77E-05)	91.90% (± 0)	92.07% (±5.77E-05)	91.89% (±0)	92.08% (±5.77E-05)	91.90% (±0)	92.06 (±0.000173205)	91.89% (±0)
SVM-Polynomial (Best K = 10)	MNIST	98.492% (±5.77E-05)	98.42% (±0)	98.479% (±0.00011547)	98.40% (±0)	98.48% (±0.00011547)	98.40% (±0)	98.48% (±1.00E-04)	98.40% (±0)
SVM-Polynomial (Best K = 9)	EMNIST Digits	99.15% (±0)	99.14% (±0.00023094)	99.16% (±0)	99.14% (±0.00023094)	99.15% (±0)	99.14% (±0.00023094)	99.15% (±0)	99.14% (±0.00023094)
SVM-RBF (Best K = 10)	MNIST	98.697% (±0)	98.67% (±0)	98.691% (±5.77E-05)	98.67% (±0)	98.69% (±5.77E-05)	98.67% (±0)	98.69% (±5.77E-05)	98.67% (±0)
SVM-RBF (Best K = 9)	EMNIST Digits	99.27% (±0)	99.25% (±0)	99.27% (±5.77E-05)	99.25% (±0)	99.27% (±5.77E-05)	99.25% (±0)	99.27% (±5.77E-05)	99.25% (±0)

To compare the validation and test performance metrices, the observed differences were minimal, suggesting a consistent generalization capability of the model. This close alignment between validation and test results is indicative of model robustness, implying that the selected kernel functions and hyperparameters are not overfitting to the validation folds.

The SVM-RBF (Support Vector Machine with Radial Basis Function kernel) achieved superior performance across both datasets, consistently outperforming other kernel configurations in terms of accuracy, precision, recall, and F1-score. The RBF kernel demonstrated robust generalization on the more complex and diverse EMNIST Digits dataset, indicating its ability to effectively capture non-linear patterns inherent in high-dimensional handwritten digit data.

On the other hand, the SVM-linear kernel gave the lowest test accuracy among all classifiers, showing clear limitations when dealing with non-linearly separable data. The polynomial kernel gave moderate results, better than the linear kernel but still lower than the RBF. This suggests that while polynomial transformation enhances the model learn more complex decision boundaries, it may not be as flexible or efficient as the RBF kernel, especially in datasets with overlapping class features.

4. Discussion

Dataset Characteristics

The input dataset obtained from Lecun and Cortes (2 0 1 0) for MNIST and Cohen (2017)[11-12], for the EMNIST digits dataset, which consists of greyscale images transformed and normalized to 28×28 pixels, includes additional blank padding to center the mass of each digit within the image frame. The distribution of pixel intensities is clearly bimodal, indicating a significant contrast between foreground (digits) and background (blank space), which facilitates the segmentation of digit shapes. This characteristic not only enhances the clarity of the digits but also contributes to improved feature extraction during the early stages of SVM. Such preprocessing ensures that the spatial alignment of digits is standardized across samples, reducing variability due to positional shifts and thus improving model convergence and generalization.

Balanced class

The EMNIST Digits dataset, which contains a larger number of samples and greater feature diversity, along with perfectly balanced class distributions, provides a more complex and informative representation than the original MNIST. These properties had a noticeable impact on the performance metrices of the SVM classifier. This improvement is particularly evident due to the use of the RBF kernel, which effectively captures non-linear decision boundaries in high-dimensional space. The RBF kernel enables the SVM to map spatial hierarchies and fine-grained variations in handwriting patterns, allowing for improved generalization over diverse character structures.

Support Vector Machine

Compared with prior studies, the accuracy result depends on the type of kernel function, such as linear kernels, polynomial kernels, hyperbolic tangents, Gaussian radial basis functions, and Laplacian radial basis functions.

From literature review, the SVM with Gaussian RBF kernel function consistently outperforms other kernel types in terms of classification accuracy, particularly in image-based handwriting recognition tasks such as MNIST and EMNIST, due to its ability to model non-linear decision boundaries effective[9-10]. Panjaitan et al,2025[9] use SVM with RBF and reported that the test accuracy for the MNIST dataset was approximately 98.5%. In contrast, this testing accuracy result using SVM-RBF is 98.67%. Based on this results, SVM-RBF test accuracy in MNIST and EMNIST digits was achieved at least 98.67% and 99.25% in the MNIST and EMNIST digit datasets respectively, indicating performance increasing when the model was applied to the EMNIST digit dataset, which is the larger size and has more diversity of features. Whereas, the linear kernel obtained the lowest test accuracy: 94.28% and 91.90% for the MNIST and EMNIST digits datasets, respectively.

In the context of hyperparameter tuning, this study found that the linear kernel SVM exhibited the shortest execution time, while the RBF kernel SVM required the longest time execution. Balancing between computational cost and model performance, the polynomial kernel emerged as a promising compromise, offering moderate execution time while maintaining competitive classification accuracy.

Moreover, Ghadekar et.al,2018[11], using Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) as method of feature extraction, reported the accuracy achieved at 97.74% for MNIST dataset. Indeed, PCA, a dimensional reduction in pre-processing step before fitting SVM could reducing training times and improved model performance [12].

5. Conclusion

This study investigated the performance and computational cost of Support Vector Machines SVMs) with different kernel functions linear, polynomial, and RBF. In the context of handwriting character recognition. The results revealed that the linear kernel SVM achieved the shortest execution time during hyperparameter tuning, making it suitable for applications with limited computational resources. In contrast, the RBF kernel SVM demonstrated the highest

classification accuracy but required significantly longer processing time. Notably, the polynomial kernel offered a balanced trade-off, yielding competitive performance with moderate computational demands. These findings highlight the importance of selecting an appropriate kernel function based on the specific requirements of the task, such as accuracy, interpretability, and computational efficiency.

6. Recommendations

Future work may explore hybrid or adaptive kernel strategies to further enhance both performance and computational efficiency. Additionally, future research could extend this study by conducting real-time deployment testing and utilizing broader, more diverse datasets to improve scalability and applicability—particularly in healthcare settings, where such models could be integrated into digital tools for the preliminary assessment of Parkinson's disease symptoms.

Moreover, additional studies to calculate model algorithm execution time especially hyperparameter tunning step, could be conducted to assess computational efficiency and scalability, especially when deploying the models in real-world or resource-constrained environments. Additionally, exploring ensemble methods that combine multiple classifiers may improve overall performance and robustness, suggesting a need for further investigation in future studies.

Acknowledgement

I would like to sincerely thank Dr. Prasong Kitidamrongsuk and all teachers for their helping and kindly guidance, together with their thoughtful feedback throughout every stage of this study. Their mentorship not only helped shape the direction of my research but also motivated me to stay focused and keep improving.

References

Alexander Filatov IK, Arseni Seregin, inventor HOLISTIC-ANALYTICAL RECOGNITION OF HANDWRITTENTEXT 2009.

Al-Hamadani M. Classification and analysis of the MNIST dataset using PCA and SVM algorithms. 2023; 71

Arica N, Yarman-Vural FT. An overview of character recognition focused on off-line handwriting. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2001;31(2):216-33.

Cohen G, Afshar S, Tapson JC, Schaik Av. EMNIST: an extension of MNIST to handwritten letters. ArXiv. 2017; abs/1702.05373.

Ghadekar P, Ingole S, Sonone D. Handwritten Digit and Letter Recognition Using Hybrid DWT-DCT with KNN and SVM Classifier2018. 1-6 p.

Ge, inventor Method and apparatus for recognizing cursive writing from sequential input information 1994 17 May 1994.

Guberman S, Rozentsveig VV. Algorithm for the recognition of handwritten text. Avtomatika Telemekanika. 1976;5:122-9.

Lavrenko V, Rath TM, Manmatha R, editors. Holistic word recognition for handwritten historical documents. First International Workshop on Document Image Analysis for Libraries, 2004 Proceedings; 2004 23-24 Jan. 2004

Liu C-L, Fujisawa H. Classification and Learning Methods for Character Recognition: Advances and Remaining Problems. 902007. p. 139-61.

Panjaitan FS, Simbolon RS, Batubara J. Handwritten Digit Recognition Using Support Vector Machine (SVM) with Radial Basis Function (RBF) Kernel. Journal Basic Science and Technology. 2025;14(1):10-8.

Sayre KM. Machine recognition of handwritten words: A project report. Pattern Recognition.1973;5(3):213-28.

Yann LeCun CC, Chris Burges. MNIST handwritten digit database. http://yann lecun com/exdb/mnist. 2010;7(23):6.