Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2025) 000–000

Procedia
Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

# Document Deterioration Classification Using Gabor Filters and Ensemble Classifiers Integrated With LIME

Anantha Hothri Inuguri, Advik Narendran, Hemanth Saga, Remya Sivan, Peeta Basa Pati*0000-0003-2376-4591

*Department of Computer Science and Engineering, Amrita School of Computing Bengaluru, Amrita Vishwa Vidyapeetham, India.*

**Abstract**

In this era, when document preservation is of grave importance, the introduction of innovations to improve the quality of palm-leaf manuscripts becomes significant for the preservation of cultural heritage. These manuscripts deteriorate in various ways, which makes them difficult to preserve and read. The unique patterns of deterioration that can take place across each section of the same document are not easily captured by conventional methods. The current approach adopted uses chunks of palm-leaf manuscripts from documents and classifies them into deterioration types using multiple advanced machine-learning models. Multiple ensemble methods are used further to tune and increase the accuracy. After that, chunks are reassembled and tailored image enhancement is done based on the results. Wrapper forward feature selection to enhance the classification quality of the chunks and additionally Local Interpretable Model-agnostic Explanations (LIME) explainability is used to analyze features contributiing to the different classes. The proposed model outperforms traditional classification techniques by integrating chunk-wise classification with ensemble methods. It achieves a significant accuracy improvement, with Bagging Random Forest (RF) yielding the highest accuracy of 92%, compared to Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) This demonstrates the model's superior ability to handle the complexities of palm-leaf manuscript deterioration, providing a more accurate and reliable classification for document preservation.



*Keywords:* Document Deterioration; Gabor filters; Image Processing; Feature Extraction; Classification; Palm leaf documents;

## 1. Introduction

The existing historical records related to the past provide a view of historical, cultural, and social norms. Among these treasures, the palm-leaf manuscripts contain much information on various aspects of ancient wisdom and tradi-

* Corresponding author. Tel.: +91-990-014-4270.
*E-mail address:* bp_peeta@blr.amrita.edu

1877-0509 

tional practices. Unfortunately, they are also manuscripts highly prone to the natural vagaries of environmental factors, material fragility, and aging [1]. Consequently, the question turns twofold: that of saving and, more particularly, that of interpreting and making them safely accessible for the future.

Traditional methods for document recovery are labor intensive and may not apply to various degrees of document deterioration or scale for solution [2]. In addition, such approaches are very often brittle, since most of them do not take into account the subtleties and characteristics of different parts of a single manuscript. In response to these challenges, a new automated solution for a detailed classification of palm-leaf manuscripts was specially adopted for this purpose. The core of this solution is a segmentation technique that breaks a manuscript digitally into smaller, manageable chunks and then performs the classification. The classification approach becomes important because it facilitates the systematic evaluation and conservation of manuscripts. By categorizing the state of each segment as good, medium, or bad it gives clarity on the level of deterioration which is then integrated into one image and labeled after that, and the focus is made in terms of preservation of the image. This fine line of categorization is most useful in the overall preservation since it enables a more accurate assessment of what precisely has been damaged. To achieve highly accurate and reliable classes, it is possible to utilize several machine learning models, as the features of each model will be taken advantage of to enhance the model's prediction ability.

By employing Gabor filters on the image chunks, features have been extracted, which provides a significant contribution to the classification of deteriorated levels. A systematic approach is followed to select the optimal Gabor parameters for effective feature extraction. Additionally, LIME has been used to enhance the interpretability of the model by identifying which features most influence the classification. Ensemble classifiers are also utilized, which, combined with the aforementioned techniques, contribute to the novelty of our approach.

The study contributes to the conservation of palm leaf manuscripts and thus fits the following global Sustainable Development Goals (SDGs). Support SDG 4: Quality Education by providing and ensuring that people embrace and learn cultural and historical aspects throughout their lives. Through the use of advanced machine learning techniques, it advances SDG 9: Industry, Innovation, and Infrastructure by supporting technology innovation. Finally, safeguarding cultural heritage contributes to SDG 11: This would fall under the United Nations Sustainable Development Goal number 11: Sustainable Cities and Communities.

The structure of this paper is as follows: the second section provides a literature review; the third describes the dataset; the fourth outlines the methodology; the fifth presents the findings and discussion; and the final section concludes the research.

## 2. Literature Survey

To construct a document degradation classification system that has high efficiency and accuracy, a careful literature survey is carried out referring to the relevant areas and the current achievements.

Sivan et al.[2] in their study present a method to classify palm-leaf heritage documents based on image quality using Discrete Cosine Transform (DCT) features and Vision Transformers (ViT). The study compares deep learning and statistical techniques, demonstrating that integrating DCT with ViT achieves superior classification performance with a 90% F1 score for test data.

Wang et al. [3] deal with the study of identifying damage in Sanskrit palm leaf manuscripts using the proposed SegFormer model. The work also shows how the proposed SegFormer methodology is useful in identifying characters to enhance the readability of manuscripts, and also for searching with context. The gamma variation and histogram balancing aid a lot in improving contrast and sharpness which is very relevant for OCR extraction. Similarly, Bipin Nair et al. [4] using OCR and deep learning, created a model for the classification of Malayalam palm leaf manuscripts when deteriorated. Here, an accuracy of 84% is maintained in the study. Asim et al. [5] Present a passive & two-stream deep network for document image classification with textual & visual inputs. The approach also minimizes the usage of an OCR whose performance is enhanced by 4. 5% in comparison with state of the art systems on the Tobacco-3482 dataset, which proves better performance in the document categorization task. Jadli et al. [6] attempt document image classification in which deep transfer learning and feature reduction are applied in the research approach. This paper improves the accuracy and efficiency of the classification system through transfer learning and optimization of features, making the classification system stronger and usable. The versatility of texture analysis in document processing is underscored, and Nair et al.[7] provides an innovative approach to ancient epic manuscript binarization and classification using false color spectralization and the VGG-16 model, showcasing the power of deep learning in handling unique document types. Sivan et al.[8, 9] explored deep learning innovations in character recognition from palm leaves, demonstrating neural networks' potential in deciphering ancient scripts and preserving cultural heritage. They further conducted a comparative study of deep learning models for recognizing palm leaf Malayalam characters, providing insights into the performance of different architectures for this task.

Saddami et al. [10] focus generally on the identification of degradation in the images of ancient documents, a few pretrained CNN models are employed, such as Resnet 101 and Shufflenet. These subcategories are bleed-through, faint-text, and others. For unblind stages, Shufflenet obtained the accuracy of the range of 100%, which proves its efficiency and the reasonable volume of computations. Shahira and Lijiya [11] in their paper address the challenge of helping visually impaired people by categorizing document images such as graph and equations. Employing CNNs and bounding box detection, it extracts data from charts to make visual information available to the screen reader for visually impaired students improving the educational resources. Sevim et al. [12] proposed Vision transformers (ViT) for classifying the document images that perform better than CNNs. Hence, ViT models do not deteriorate in accuracy, and their use presents a good option for general and efficient document classification, as well as improving the classification performance. Noce et al. [13] have presented a novel document image classification method that combines visual and textual information. The preprocessing step involves embedding additional textual information extracted via OCR and NLP into document images. This combined approach significantly improves the classification results of a neural network, achieving an overall increase in the classification accuracy of documents of up to 30% for specific subsets. The method demonstrates the potential for enhanced document processing and analysis, particularly in datasets with visually similar document classes. Shahkolaei et al. [14] have explained the comparison analysis of various image quality assessment metrics for degraded document images. They introduced the Multi-distortion Document Quality Measure (MDQM) for assessing the quality of physically degraded documents. The MDQM metric, based on spatial and frequency image features, achieved the best performance with high correlation to human judgments. The proposed method includes a degradation classification model using SVM to estimate the probability of four common degradation types: paper translucency, stain, reader annotations, and worn holes. The study demonstrated MDQM's efficacy for classification and its moderate complexity, showing promise for automatic degradation modeling in future work .

Chen et al. [15] deal with the problem of image classification specifically to Chinese local gazetteers, which enhance the efficiency of the historical documents' database. Therefore, this research demonstrates how complex image classification methods can be applied in the preservation of cultural heritage. Xu et al. [16] also propose LayoutLM that enables the co-learning of text and layout information in document images. Outperforming the state-of-the-art in tasks such as form and receipt recognition, LayoutLM greatly improves the document image understanding, with the released code and pre-trained model for community use.

Nair et al.[17] introduced a deep binarization model using ResNet for ancient horoscopic palm leaf binarization, showcasing the use of advanced CNN architectures to digitize and analyze ancient texts. These studies highlight ongoing efforts to improve digital restoration and analysis of ancient and degraded documents, using both traditional image processing and advanced machine learning techniques. Finally, Sivan et al. [**?** ] utilized a discrete cosine transform integrated vision transformer to classify palm leaf documents into different classes based on the level of deterioration.

Classifying palm leaf documents is essential for digitization tasks, as pre-processing and segmentation should be tailored to the level of deterioration [**?** ]. However, the literature survey conducted reveals that very little work has been reported on classifying these documents according to their deterioration levels. Hence, it is proposed to use optimal Gabor-based feature extraction from image fragments for palm leaf classification based on deterioration level.

## 3. Dataset

The dataset used in this study is specifically designed for the assessment of deterioration and enhancement of document images. It includes quantitative expressions about both document quality and degradation. The unlabeled images are collected from [18]. The dataset comprises a total of 433 images, each of which has been divided into smaller chunks measuring 256 pixels by 256 pixels for detailed analysis. Depending on the size of the original image, each palm leaf image yielded between 60 to 80 chunks, resulting in an entire dataset of 16,406 chunks. The class distribution within the dataset includes 4,624 chunks categorized as "Good," 5,234 chunks labeled as "Medium," and 6,548 chunks classified as "Bad" based on the deterioration level.

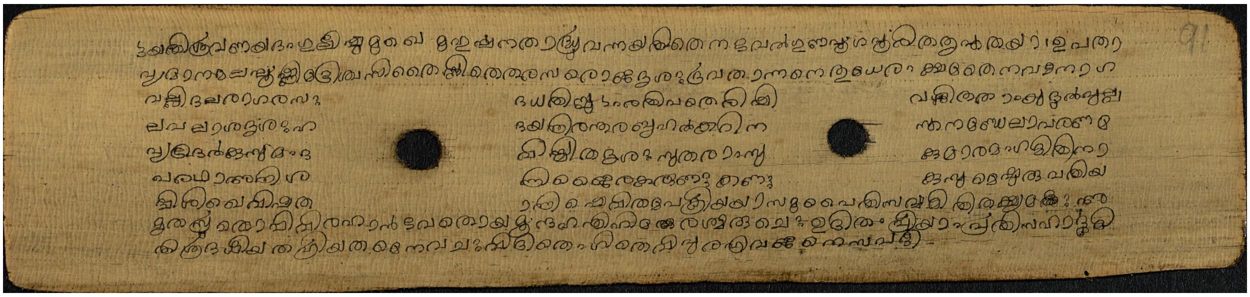


Fig. 1: Full Image of the palm leaf

Fig. 1. shows the palm leaf images and Fig. 2. shows the image chunks. The chunk-wise breakdown of the image will make detailed analysis easier. The document can be at different levels of deterioration, and this method also makes it possible to classify its different parts, thus ensuring appropriate enhancement treatment for every part of the document.

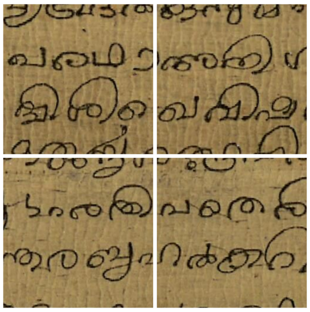


Fig. 2: The image broken down to multiple chunks

After creating the image chunks, the extraction of Gabor features from the chunks of the palm leaf images is performed. These are textural features in nature that help find the most important classification of the image and enhance the image. The Gabor filters were applied on the image chunks by varying the parameters such as gamma, sigma, and phi. These parameters play a very great role in the features that will be extracted: gamma determines the spatial aspect ratio of the filter, sigma determines the standard deviation of the Gaussian function that will influence the scale of the features, and phi determines the phase offset that will affect the orientation and edge detection capabilities of the filter.

Table 1: Optimal Gabor Parameters

| **Parameter** | **Gamma ($\gamma$)** | **Sigma ($\sigma$)** | **Phi ($\phi$)** | **Theta ($\theta$)** | **Lambda ($\lambda$)** |
|---|---|---|---|---|---|
| **Value** | 0.1 | 0.5 | 0 | $0, \pi, \pi/2, \pi/4, 3\pi/4$ | $2\pi/1, 2\pi/2, 2\pi/3, 2\pi/4, 2\pi/5$ |

By employing Gabor filters two main features were extracted from the process of filtering: Local Energy and Mean Amplitude, which measure the textural characteristics of the images. The features were calculated over a range of orientations (theta) and wavelengths (lambda) of the Gabor filter, capturing the overall textural information. The choice is made from a set of fine-tuned features of 900 by trying the combination of different values of gamma, sigma, and phi. Table 1 demonstrates the optimal Gabor parameters obtained after feature engineering.

## 4. Methodology

The proposed methodology for document classification based on degradation levels includes seven steps: 1. Chunking 2. Extracting Features 3. Feature Engineering 4. Classification 5. Comparing different values 6. Aggregation and 7. Explainability with LIME.

The experiments for this research were conducted on a system equipped by an Intel Core i5 processor, 8GB RAM, and storage of 256GB SSD. The system also features Nvidia GeForce graphics, which facilitated the computation-heavy tasks associated with image processing and machine learning models. The operating system used was Windows 10 Home, and Jupyter Notebook was employed as the main development environment. Libraries used include cv2 for image processing, numpy for numerical operations, matplotlib and seaborn for visualization, pandas for data analysis, scikit-learn for machine learning, lime for interpretability, and mlxtend for extended tools.

The model starts by taking an image of a document as input. The chunker module processes this image, breaking it down into smaller chunks. Next, the feature extractor module uses Gabor filters to extract features from these chunks. Feature engineering is utilized after feature extraction to identify and select the most relevant features from the extracted ones. After feature engineering, classifiers are applied to these selected features to classify the documents and compare their accuracy with the previous highest accuracy. If the new accuracy surpasses the previous highest accuracy, save the current accuracy and parameters as the new highest and optimal values. Repeat this process for all parameter sets. After identifying the optimal parameters, chunk labels are predicted. An aggregator then uses these labels to predict a label for the entire image. Additionally, LIME explainability is used to analyze how each feature contributes to the prediction for the current sample. The proposed method is shown in Fig. 3.

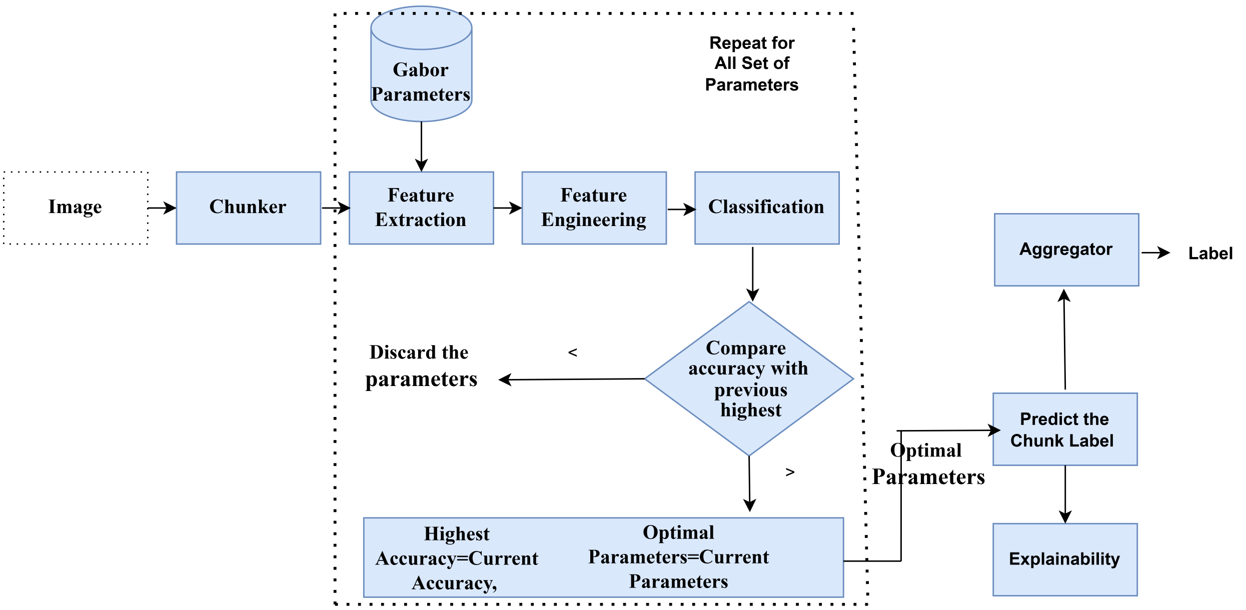



Fig. 3: Flow diagram of the proposed model

The chunker divides images into 256x256 chunks. All the generated chunks are visually inspected and manually labeled based on the degradation present in them. The possible classes are good, bad, and medium. Gabor filters are employed on the generated chunks to extract the features. The extracted features are the local energy and mean amplitude features. The choice of Gabor parameters is made from a set of 900 parameters by experimenting with different combinations of gamma, sigma, and phi, while theta and lambda remain constant. These are the parameters used for the experiment: Gamma: [0.1, 0.5, 1.0, 1.5, 2.0], Sigma: [0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0], Phi: [0, $\pi/4$, $\pi/2$, $3\pi/4$, $\pi$, $5\pi/4$, $3\pi/2$, $7\pi/4$, $2\pi$]. The constant parameters are: Theta: [0, $\pi$, $\pi/2$, $\pi/4$, $3\pi/4$], Lambda: [$2\pi/1$, $2\pi/2$, $2\pi/3$, $2\pi/4$, $2\pi/5$]. Gabor kernels are created for each combination of gamma, sigma, and phi, with constant theta and lambda, and are applied to the chunks. As part of feature

**Algorithm 1** Gabor Filter Feature Extraction and Best Hyperparameters

Initialize main directory path and list of label folders
Define Gabor filter parameters:
  `Gamma` = [0.1, 0.5, 1.0, 1.5, 2.0]
  `Sigma` = [0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0]
  `Phi` = [0, $\pi/4$, $\pi/2$, $3\pi/4$, $\pi$, $5\pi/4$, $3\pi/2$, $7\pi/4$, $2\pi$]
  `Theta` = [0, $\pi$, $\pi/2$, $\pi/4$, $3\pi/4$]
  `Lambda` = [$2\pi/1$, $2\pi/2$, $2\pi/3$, $2\pi/4$, $2\pi/5$]
Iterates over different Gabor filter parameter combinations and processes each image where the theta and lambda remain constant for all.
For each image, it generates a Gabor kernel, applies the filter, and calculates local energy and mean amplitude. The results are stored in a DataFrame and saved to a CSV file for each parameter combination.
Initialize a dictionary `best_performance` to track the best accuracy for each classifier.
Define a dictionary of classifiers:
  `SVM, Random Forest, Gradient Boosting, K-Nearest Neighbors, Naive Bayes, Logistic Regression, MLP Classifier, Decision Tree, AdaBoost, Bagging (Random Forest), Boosting (AdaBoost).`
**for** each file in the directory **do**
  Load the CSV file into DataFrame `data`.
  Extract target labels `y` from `data`.
  Extract features `X` from `data`, excluding `ImageName` and `Label`.
  Standardize the features using `StandardScaler`.
  Encode target labels using `LabelEncoder`.
  **for** each classifier in the dictionary **do**
    Fit the classifier on `X` and `y`.
    Predict the target labels using the trained classifier.
    Calculate the accuracy of the predictions.
    **if** current accuracy is greater than best recorded accuracy **then**
      Update `best_performance` dictionary with the current accuracy and file name.
    **end if**
  **end for**
**end for**
**for** each model in `best_performance` **do**
  Print the best file and accuracy for the model.
**end for**
**Final Output Values Selected:**
Gamma ($\gamma$) = 0.1
Sigma ($\sigma$) = 0.5
Phi ($\phi$) = 0

engineering, forward feature selection is employed for each combination. After performing forward feature selection, 22 features were chosen from the original set of 25 features. These selected features were then used by classifiers to perform classification tasks. The classification accuracies obtained were compared with the previously recorded highest accuracies for each classifier. If the current accuracy is higher than the previous one, the current parameters are considered the best parameters. This process continues for all the recorded parameters. Afterward, the optimal parameters are obtained to create the Gabor kernel. The optimal parameter selection is described in the given algorithm. The classification process for the image chunks using the best chosen parameters from the Gabor filter data of the palm leaf involves using a series of machine learning models coupled with an ensemble method to increase the overall accuracy. A variety of classifiers are employed, each chosen for its own unique strengths. Support Vector Machines (SVM) and RF for managing non-linear relationships, Naïve Bayes (NB) for probabilistic predictions, and

Multi-Layer Perceptron (MLP) for dealing with high-dimensional data. Other models used included KNN, Logistic Regression (LR), Decision Tree (DT), and combinations of bagging and boosting with RF and AdaBoost, respectively. To validate the classification model's performance, GridSearchCV cross-validation was employed. The dataset was randomly split into training and testing subsets for each fold, optimizing hyperparameters for the best performance Each classifier then made predictions on a new set of image chunks, which were then compiled into a data frame. To further refine the accuracy, a voting classifier is employed, combining all individual predictions into a final decision based on a majority vote. This ensemble strategy is chosen to compensate for the limitations of individual models and harness their combined strengths, ensuring a more robust and accurate classification. This comprehensive method aims to produce results that are not only more accurate but also more reliable.

After dividing each image into 256x256 pixel chunks and classifying them separately, the predictions for each chunk are loaded from an Excel file. The image names are extracted without the chunk numbers, and chunks are grouped by their parent image. An aggregation function, specifically simple voting, combines the classifications of all chunks to determine the final classification for the entire image. This final classification, along with the image name, is stored in a DataFrame and written to an Excel file. This method ensures accurate image classification by integrating chunk predictions and calculating the overall accuracy for each model.

The model is evaluated using precision, recall, F1-score, support for all the good, bad, and medium chunks, and integrated images. Detailed classification reports of the classifiers are further elaborated on based on the training data. The feature contributions from the model predictions using LIME are also implemented to ensure classifier interpretability.

To validate the classification task, various image enhancement techniques are applied to the predicted chunks. Six different enhancement techniques are applied to the predicted chunks: negative transformation, log transformation, gamma correction, piecewise linear transformation, histogram equalization, and intensity transformation. Subsequently, the processed images are binarized using the Otsu method to obtain binary images. The quality of each enhancement is assessed in terms of the Peak Signal-to-Noise Ratio (PSNR), calculated between the original grayscale images and their respective enhanced binarized versions. From the experiment, it is evident that the same technique is not suitable for all classes of documents. Different classes of documents require different sets of enhancement techniques. This approach can significantly boost digitization accuracy by selecting appropriate enhancement methods for each type of document.

## 5. Results and Analysis

The section reports the results and analysis of the work that litigates, in adapting the strategic approach detailed previously, which comprised the areas of segmentation, feature selection, classification and enhancement, to uncover the critical insights indispensable for document preservation and analysis.

One of the essential parts is image segmentation, which divides the image into smaller, manageable, accurate components, which makes the process faster. Taking this approach, the analysis compares the performance of the classifier models with and without feature selection and then looks at several different image enhancement features.

The best hyperparameters are selected by employing the method explained in the methodology section. The parameter values chosen are as follows: Gamma ($\gamma$): 0.1, Sigma ($\sigma$): 0.5, Phi ($\phi$): 0, Theta ($\theta$): 0, $\pi$, $\pi/2$, $\pi/4$, $3\pi/4$, Lambda ($\lambda$): $2\pi/1$, $2\pi/2$, $2\pi/3$, $2\pi/4$, $2\pi/5$. These parameters are used to process the entire dataset, and the results are presented based on these values.

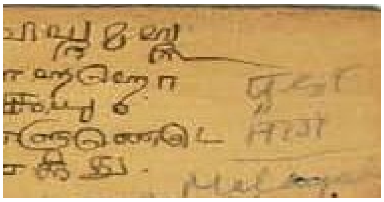
In the model metrics shown in Table 2, it is observed that, generally, feature processing has favorable theoretical properties of stability and robustness. Using feature selection, it can be seen that the models achieve stable metrics levels (precision, recall, F1-score, and accuracy) where better generalization is seen. A total of 22 features were selected using the forward feature selection method. Despite this, there is a slight improvement in the models without feature selection, even though in terms of raw accuracy the performance is better than the former. This higher precision is mainly because the models are being trained with more features, which might include some that are relevant, but which few experts would have guessed to be meaningful.

Table 2: Model Performance Metrics Across Different Classifiers with Feature Selection

| Metric | SVM | RF | GB | KNN | NB | LR | MLPC | DT | Bagging (RF) | Boosting (AdaBoost) | Voting |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Precision** | 0.73 | 1 | 0.84 | 0.82 | 0.63 | 0.72 | 0.8 | 1 | 1 | 0.73 | 0.85 |
| **Recall** | 0.69 | 1 | 0.82 | 0.82 | 0.65 | 0.71 | 0.8 | 1 | 1 | 0.73 | 0.82 |
| **F1** | 0.64 | 1 | 0.81 | 0.82 | 0.58 | 0.66 | 0.79 | 1 | 1 | 0.7 | 0.81 |
| **F1(B)** | 0.68 | 1 | 0.83 | 0.83 | 0.6 | 0.7 | 0.81 | 1 | 1 | 0.73 | 0.84 |
| **F1(G)** | 0.79 | 1 | 0.86 | 0.87 | 0.76 | 0.8 | 0.85 | 1 | 1 | 0.81 | 0.86 |
| **F1(M)** | 0.18 | 1 | 0.64 | 0.66 | 0.05 | 0.23 | 0.58 | 1 | 1 | 0.3 | 0.63 |
| **Accuracy** | 0.69 | 1 | 0.82 | 0.82 | 65 | 0.71 | 0.8 | 1 | 1 | 0.73 | 0.82 |

However, the subset of attributes constitutes a more practical and effective representation. This approach is used because when a few features are discarded, the models are trained faster and they need less computational power, which makes them more scalable and efficient, while processing huge data sets. Evaluating the models using multiple measures, such as performance metrics, enables a comprehensive assessment of each algorithm. Fig. 4 shows a few examples of the classification of chunks made by the voting classifier.

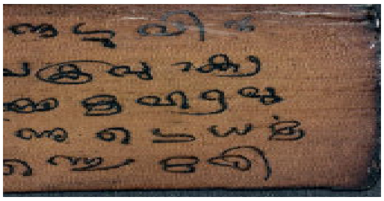
(a)



(b)

(c)

(d)



(e)

(f)

Fig. 4: Correctly and incorrectly classified image chunk by the voting classifier. (a) shows more deteriorated image chunk is correctly classified, (b) shows the less deteriorated image chunk is correctly classified, (c) shows the medium deteriorated image chunk is correctly classified, (d) shows the more deteriorated image chunk is classified as medium deteriorated, (e) shows the less deteriorated chunk is classified as medium deteriorated, and (f) shows the medium deteriorated image chunk is classified as more deteriorated.

Fig. 5. is a visualization of the LIME explanation for a bad sample. LIME is employed for a low-quality image, resulting in a prediction of bad quality with a confidence of 0.74%. The LIME results indicate that the features, highlighted in green boxes, significantly contributes to accurately predicting the image as of bad quality.

Table 3 shows the accuracy of different machine learning models for image classification after integrating chunk classifications. Voting mechanisms combine chunk predictions to determine the overall image classification, from which model accuracies are calculated. The RF and bagging RF models give the highest accuracies of 0.92, which appeared to be the most appropriate for this dataset. The decisiveness of the DT and voting ensemble gradually gives good accuracies of around 0.87 and 0.71, respectively. Less accurate accuracies are maintained by the SVM, NB, LR, MLP Classifier, and Bagging of AdaBoost, linking them to lesser suitability. From this, more importance will have to be put on the applicability of the KNN and RF in future works and probably use the strengths of both models to the creation of a classifier for higher performance.

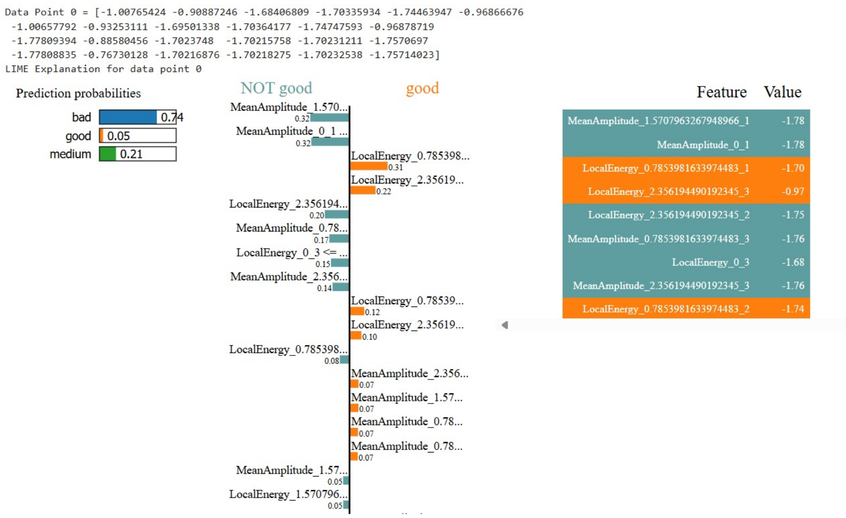



Fig. 5: Lime Explanation of Data Point

Table 3: Model Accuracies After Integrating Chunks to Full Images with Feature Selection

| **Model** | **SVM** | **RF** | **KNN** | **NB** | **LR** | **MLP Classifier** | **DT** | **Bagging (AdaBoost)** | **Bagging (Random Forest)** | **Voting Ensemble** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Accuracy** | 0.54 | 0.92 | 0.64 | 0.52 | 0.56 | 0.68 | 0.87 | 0.6 | 0.92 | 0.71 |

Table 4: Comparison of proposed model results with the existing literature

| | **Current Paper (Gabor + Ensemble Models)** | **Compared Paper 1 (Deep Learning Techniques)[2]** | **Compared Paper 2 (OCR + Voting Process)[4]** |
|---|---|---|---|
| **Classifiers Used** | RF, SVM, KNN, Naive Bayes, Logistic Regression, Decision Tree, Bagging, Boosting, MLP | ViT, Custom CNN, VGG16, VGG19, ResNet152v2, Decision Tree, Logistic Regression | Tesseract 4 OCR with deep learning |
| **Best Model/Classifier** | Random Forest (RF) - Accuracy 92% | Vision Transformer (ViT) with DCT features - F1 score of 90% | Tesseract OCR - Accuracy 84.5% (with voting) |
| **Accuracy** | 92% | 90% | 83% (without voting), 84.5% (with voting) |
| **Preprocessing** | Chunk-wise segmentation and Gabor parameter tuning | SMOTE for balancing dataset, resizing to 100x300 pixels | Dataset analysis and bag of words repository creation |
| **Evaluation Metrics** | Precision, Recall, F1-Score, Accuracy | Accuracy, Precision, Recall, F1-Score | F1-Score of 0.90, Accuracy |

This report investigates the use of Gabor filters for the extraction of features from palm-leaf documents, employing machine learning classifiers such as Random Forest, SVM, and KNN. It enhances performance through ensemble methods and utilizes LIME for feature interpretability. In contrast, Sivan et al.[2] employ a combination of Discrete Cosine Transform and Vision Transformer, focusing on deep learning to boost classification accuracy but lacking interpretability. Additionally, Bipin Nair et al. [4] provides another method which targets six categories of Malayalam palm leaf manuscripts, following a three-phase approach: dataset analysis, building a bag of words repository, and recognition and classification using a voting mechanism. This method utilizes Tesseract 4 OCR for text extraction. All the three methods aim to classify manuscript deterioration. The current technique excels in explainability and ensemble methods, while the alternative approach highlights the effectiveness of deep learning. A detailed comparison is provided in Table 4.

To justify the classification of historical documents into different classes based on the deterioration level, some traditional enhancement techniques along with Otsu binarization have been employed on the image documents. From the result obtained in Fig. 6, it is clear that traditional methods and Otsu binarization work well for good documents.

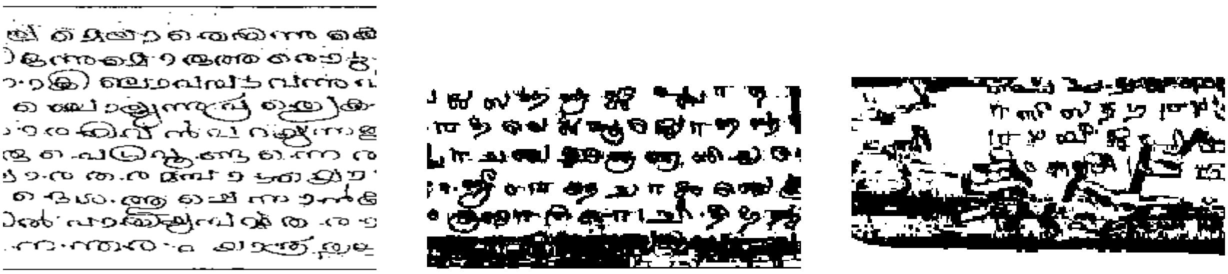


Fig. 6: Images processed Using Otsu Binarization (a) shows binarization of good image, (b) binarization of bad image, (c) shows binarization of medium images.

From the observations in Table 5, the combination of Intensity Transformation and Otsu's thresholding emerged as the best-performing technique for good-classified palm leaf manuscripts, achieving the highest PSNR of 28.57 dB and demonstrating superior noise suppression and detail enhancement.

Table 5: PSNR (dB) Values for Binarized Images of Good Images

| Enhancement Technique | Negative Log | Gamma | Piecewise | Linear | Histogram | Intensity |
|---|---|---|---|---|---|---|
| **Good Images** | 28.5 | 28.49 | 28.49 | 28.5 | 28.48 | 28.57 |

In real-life conditions, especially when preserving old writing, it is crucial to identify which enhancement methods effectively restore the appearance of the document while maintaining its integrity. Different types of deterioration, such as fading, stains, or physical damage, may require tailored solutions. The ability to select the optimal technique for each specific case ensures that the historical value of the document is preserved without introducing artifacts or losing crucial details, thus reinforcing the need for adaptive and document-specific restoration methods.

## 6. Conclusion

This study highlights the essential role of automated document classification in preserving cultural heritage. Using machine learning techniques such as Gabor filters and ensemble methods, significant improvements in classification accuracy have been achieved. This research not only contributes to the field of document preservation, but also lays the groundwork for future innovations in image classification, particularly for historical artifacts.

A key focus of this work was the classification of palm-leaf manuscripts based on their degradation levels. By categorizing the manuscripts as good, medium or bad chunk-wise, a series of improvements was systematically explored and implemented. The novel technique combines Gabor feature extraction with an ensemble classifier, providing a more accurate solution for identifying and enhancing deteriorated manuscripts. This analysis demonstrates that the use of ensemble methods significantly improves the accuracy of classification, offering a valuable tool for the preservation and restoration of cultural heritage.

## 7. Future Enhancement

For future work, integrating deep learning algorithms, particularly convolutional neural networks (CNNs) or generative adversarial networks (GANs), could enhance the ability to autonomously select the optimal restoration methods based on the unique features of each manuscript. These neural networks could adaptively learn from the characteristics of each document, enabling more precise enhancement techniques for different levels of degradation.

Additionally, incorporating advanced deep learning methods for detecting and correcting areas with significant damage, such as missing sections or faded text, could further improve the restoration process. Techniques like auto-encoders or image inpainting could be used to fill in missing details, ensuring a more refined restoration. It is suitable for a wider range of historical documents beyond palm-leaf manuscripts.

## Acknowledgements

We would like to express our sincere gratitude to Amrita Vishwa Vidyapeetham for the infrastructure and support provided towards the completion of this project. We acknowledge that we used Grammarly to correct some sentences in this article.

## References

[1] Sulaiman, A., Omar, K. and Nasrudin, M.F., 2019. Degraded historical document binarization: A review on issues, challenges, techniques, and future directions. Journal of Imaging, 5(4), p.48.

[2] Sivan, R., Pati, P.B. & Kesiman, M.W.A. Image quality determination of palm leaf heritage documents using integrated discrete cosine transform features with vision transformer. IJDAR (2024). https://doi.org/10.1007/s10032-024-00490-x

[3] Wang, Y., Wen, M., Zhou, X., Gao, F., Tian, S., Jue, D., Lu, H. and Zhang, Z., 2024. Automatic damage identification of Sanskrit palm leaf manuscripts with SegFormer. Heritage Science,12(1), p.8.

[4] Bipin Nair, B.J., Shobha Rani, N. and Khan, M., 2023. Deteriorated image classification model for malayalam palm leaf manuscripts. Journal of Intelligent & Fuzzy Systems, 45(3), pp.4031-4049.

[5] M. N. Asim, M. U. G. Khan, M. I. Malik, K. Razzaque, A. Dengel and S. Ahmed, "Two Stream Deep Network for Document Image Classification," 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 2019, pp. 1410-1416, doi: 10.1109/ICDAR.2019.00227.

[6] Jadli, A., Hain, M. and Hasbaoui, A., 2021. An improved document image classification using deep transfer learning and feature reduction. International Journal, 10(2), pp.549-557.

[7] Nair, BJ Bipin, KV Aadith Raj, M. Kedar, S. Pai Vaishak, and E. V. Sreejil. "Ancient Epic Manuscript Binarization and Classification Using False Color Spectralization and VGG-16 Model." Procedia Computer Science 218 (2023): 631-643.

[8] R. Sivan, T. Singh and P. B. Pati, "Malayalam Character Recognition from Palm Leaves Using Deep-Learning," 2022 OITS International Conference on Information Technology (OCIT), Bhubaneswar, India, 2022, pp. 134-139, doi: 10.1109/OCIT56763.2022.00035.

[9] R. Sivan, S. Palaniswamy and P. B. Pati, "Comparative Study of Deep Learning models to Recognize Palm Leaf Malayalam Characters," 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/CSITSS57437.2022.10026392.

[10] K. Saddami, K. Munadi and F. Arnia, "Degradation Classification on Ancient Document Image Based on Deep Neural Networks," 2020 3rd International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2020, pp. 405-410, doi: 10.1109/ICOIACT50329.2020.9332042.

[11] K. C. Shahira and A. Lijiya, "Document Image Classification: Towards Assisting Visually Impaired," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, 2019, pp. 852-857, doi: 10.1109/TENCON.2019.8929594.

[12] Sevim, S., Omurca, S.İ. and Ekinci, E., 2022, February. Document image classification with vision transformers. In International Congress of Electrical and Computer Engineering (pp. 68-81). Cham: Springer International Publishing.

[13] Noce L, Gallo I, Zamberletti A, Calefati A (2016) 'Embedded textual content for document image classification with convolutional neural networks'. In Proceedings of the 2016 ACM symposium on document engineering (pp. 165-173).

[14] Shahkolaei A, Beghdadi A, Cheriet M (2019) Blind quality assessment metric and degradation classification for degraded document images. Signal Processing: Image Communication, 76, pp.11-21.

[15] Chen, J.A., Hou, J.C., Tsai, R.T.H., Liao, H.M., Chen, S.P. and Chang, M.C., 2024. Image classification for historical documents: a study on Chinese local gazetteers. Digital Scholarship in the Humanities, 39(1), pp.61-73.

[16] Xu, Y., Li, M., Cui, L., Huang, S., Wei, F. and Zhou, M., 2020, August. Layoutlm: Pre-training of text and layout for document image understanding. In Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 1192-1200).

[17] BB. Nair B J and A. S. Nair, "Ancient Horoscopic Palm Leaf Binarization Using A Deep Binarization Model - RESNET," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 1524-1529, doi: 10.1109/ICCMC51019.2021.9418461.

[18] Shiju Alex. (n.d.). Shiju Alex personal website. Retrieved October 8, 2024, from http://shijualex.in/