Classification of heterogeneous Malayalam documents based on structural features using deep learning models

Bipin Nair Balakrishnan Jayakumari, Amel Thomas Kavana
Department of Computer Science, Amrita School of Arts and Sciences, Mysuru Campus, Amrita Vishwa Vidyapeetham,
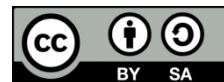Tamil Nadu, India

Article Info	ABSTRACT
	The proposed work gives a comparative study on performance of various pretrained deep learning models for classifying Malayalam documents such as agreement documents, notebook images, and palm leaves. The documents are classified based on their visual and structural features. The dataset was manually collected from different sources. The method of research proceeds with preprocessing, feature extraction, and classification. The proposed work deals with three fine-tuned deep learning models such as visual geometry group-16 (VGG-16), convolutional neural network (CNN) and AlexNet. The models attained high accuracies of 99.7%, 96%, and 95%, respectively. Among the three models, the fine-tuned VGG-16 model was found to perform better attaining a very high accuracy on the dataset. As a future work, methods to classify the documents based on content as well as spectral features can be developed. *This is an open access article under the [CC BY-SA](#) license.*

Corresponding Author:

Bipin Nair Balakrishnan Jayakumari
Department of Computer Science, Amrita School of Arts and Sciences
Amrita Vishwa Vidyapeetham, Tamil Nadu 641112, India
Email: bipin.bj.nair@gmail.com

1. INTRODUCTION

Ancient documents reveal the history of the people, nation as well as tradition. The preservation and segregation of these documents is a tedious process. The documents get degraded over time due to various natural factors like aging, environmental factors, and accidental errors [1]. Digitization of documents is an effective method for the preservation as well as classifying the documents.

Document classification will throw a new light on the new era of digitization and categorization. As the need for digitization increases, the requirement for the classification of these digitized documents comes into play. Document classification is the process of classifying or grouping documents into various categories based on the structural features. Document segregation and storage is an important step in information management and retrieval [2]. As each document belongs to different categories, classifying the documents manually would consume more time and effort. A deep learning-based method is proposed to check which classifier obtains high accuracy for the classification of ancient Malayalam documents based on structural features. The documents are classified into three categories namely palm leaves, agreement copies, and notebook images.

Mushtaq *et al.* [1] investigated a deep learning-based convolutional neural network (CNN) model for spectral image classification on datasets which consist of 10, 10 and 50 classes respectively have got a good accuracy of 99.04%, 99.49%, and 97.57% were reported respectively for each dataset. In [2], [3] a deep model based on CNN was adopted for classifying documents that was experimented on various datasets containing over 5,200 documents got remarkable accuracies compared to the traditional models. Quan *et al.* [4] discussed a multispectral fusion-based classification on radar image and sentinel-2A datasets with

adequate accuracy. Prieto *et al.* [5] proposed a deep model probabilistic indexing to perform classification on un-transcribed manuscript images from Spanish *Archivo General de Indias* that could achieve an accuracy of 70% on 3 different classes. Nasir *et al.* [6] investigated a deep model AlexNet and visual geometry group-19 (VGG-19) on Tobacco-3842 dataset consisting of 10 different image classes and a dataset containing 15 different classes achieved an exceptional accuracy rate of 93.1%. Bakkali *et al.* [7] discussed a deep classification model Nasnet on Tobacco-3842 to give an overall accuracy of 99.7% over 10 different classes. A classification graph based neural network model was investigated by Mandivarapu *et al.* [8] on insurance and Tobacco 3842 dataset which give an adequate accuracy of 91% for 11 classes and 77.5% for 10 different classes. Chen *et al.* [9] experimented a decomposition-based hyperspectral classification approach on the Sar images dataset which achieved a remarkable accuracy of 99.7%. CNN was found to be effective for classifying hyperspectral images [10], [11] resulting in faster processing and higher classification accuracies over various hyperspectral image datasets.

A novel model was investigated by Gayathri and Kannan [12] on a dataset containing over 3,000 images to classify Ayurvedic documents which gave promising results. Jimenez *et al.* [13] proposed a deep neural network model to classify Covid-19 literature on the LitCovid and Covid-19 open research dataset (CORD-19) datasets containing 8 classes each showed better accuracies than the existing traditional methods. A deep learning-based Squeeze net model on Tobacco-3842 dataset consisting of 10 different classes was discussed by Hassanpour *et al.* [14] to classify documents based on visual features with an accuracy of 77% was obtained. Kanchi *et al.* [15] discussed a deep multi model-based approach to classify documents on datasets containing 16 and 10 classes respectively. The proposed approach obtained an accuracy of 90.3%. A deep active learning-based approach was investigated by Hemmer *et al.* [16] for the classification of images that attained an accuracy of 90% on the Modified National Institute of Standards and Technology (MNIST) and Cifar-10 datasets. Indraswari *et al.* [17] proposed a mobile net-based classification of melanoma images which achieved an accuracy of 85% over four different datasets containing images belonging to two classes. Ahmed *et al.* [18] investigated a deep neural net model with attention mechanism for Bangla document classification on a manually collected dataset. The proposed model obtained an accuracy rate of 86.56% was obtained over 13 document classes. Pan *et al.* [19] experimented with an ontology-driven approach to classify scientific literature that achieved a score of 95% on DBLP dataset. Jiang *et al.* [20] proposed three various deep models for technical documents for the classification that could yield a decent accuracy rate of 77.9% over 50 distinct classes. A deep learning-based adaptive multiscale segmentation method was proposed by Zhao *et al.* [21] on Indian Pines, Salinas Scene and University of Pavia datasets containing 16.9 and 15 classes respectively on which accuracies of 94.312%, 99.217%, and 92.693% were obtained. A deep learning-based hybrid machine learning based model was developed by Swetanisha *et al.* [22] for classifying multispectral images on Landsat-8 dataset containing 7 different classes of satellite images could attain decent accuracy scores. In [23], [24] a combined approach of deep learning and machine learning models were used to perform multiclass classification of documents which could yield exceptionally high accuracy values. Deep neural network-based models were proposed in [25]–[27] for classification of hyperspectral images on the Indian Pines and University of Houston and Salinas Seas datasets containing over 16 image classes. The proposed models could attain a high accuracy rate varying between 90-99%. Jayakumari and Nair [28] proposed a deep learning based ResNet model to perform binarization of ancient horoscopic palm leaf images which attained a very high accuracy of 95.38% on a manually collected dataset consisting of ancient horoscopic palm leaves. Deep learning is definitely the pick of the bunch when the problem requires processing of huge and unstructured deep model data processing [29]. The further sections of paper are described below as methods, results and discussion, and conclusion.

2. METHOD

The proposed method classifies ancient documents such as agreement copies, palm leaf manuscripts, and notebook images based on their structural features. The methodology has three approaches using three different deep learning models such as CNN, fine-tuned VGG-16 and modified AlexNet along with various enhancement methods for classifying the documents. Each model is evaluated on the basis of the accuracy obtained over the dataset. The model that performs better on the dataset is identified from the evaluation results.

2.1. Data collection

The proposed work is to classify Malayalam documents belonging to three different categories of documents such as agreement copies, palm leaves, and notebook images. The datasets used for our proposed work are manually collected from people as well as from online repositories. The datasets obtained contained degraded documents which made them difficult to be classified because of the same pattern of Malayalam characters present in all the documents.

Table 1 displays the details about the datasets, their sources, and the number of samples collected. The agreement copies were manually collected from various Taluk offices as well as from the internet. The palm leaf data set is collected from various people in Kerala as well as from Varikkasheri Mana, Palakkad, Kerala; and notebook images [30] are collected from various schools and colleges of Kottayam Kerala. For the research, 1,500 samples of each category of documents were collected.

Table 1. Dataset collection details

S.no	Source	Dataset	Count
1	Internet, Taluk offices	Agreement copies	1,500
2	Kerala Manas	Palm leaves	1,500
3	Universities, Schools, Online repositories	Notebook images	1,500

2.2. Preprocessing

In the initial approach using CNN, the input image is resized to 224×224 and is converted into grayscale format. The grayscale image then undergoes Otsu thresholding for image enhancement. Thresholding is done to binarize the input image based on its pixel intensity values. The Otsu enhancement process uses (1).

$$\sigma^2(x) = \omega_{bg}(x)\sigma^2_{bg}(x) + \omega_{fg}(x)\sigma^2_{fg}(x) \tag{1}$$

In (1), $\omega_{bg}(x)$ and $\omega_{fg}(x)$ are the probability of the number of pixels of each class at a threshold value of X. σ^2 represents the color value variance. In the second approach using VGG-16, the enhancement process is done by normalizing the RGB values for each pixel of the input document image. Here, the mean pixel value is reduced from each pixel in this process. The image is normalized and resized to 224×224. The normalization enhancement process uses (2).

$$yi = \frac{si - min(s)}{max(s) - min(s)} \tag{2}$$

In (2), s is the input data that ranges from s_1 to s_n, and y_i becomes the ith normalized data. In the approach using AlexNet, the input image is initially normalized by rescaling. The image is then resized to 227×227 as it is the standard input size for AlexNet architecture. In third approach also uses the same pre-processing method as normalization.

2.3. Classification using CNN

In Figure 1, after pre-processing the image is taken as the input for the CNN model. The model consists of three stacks of convolutional and MaxPooling layers, a flattening layer, and three dense layers. The input image of size 224×224 is passed through a convolutional layer having 32 filters of size 3×3 with ReLU activation function from which it is passed to a max pooling layer of filter size 2×2. The image is then inputted into the next convolution layer with filter having 64 filters of size 3×3. It is then passed to the MaxPooling layer of filter size 2×2 that follows. The image is forwarded to the next convolutional layer having number of filters 64 of size 3×3. It is then inputted to the next MaxPooling layer of filter size 3×3 and a dropout layer that follows. The output image is flattened and fed into the hidden dense layer with 128 filters and ReLU activation function. The layer that follows is the output layer trained for three classes.
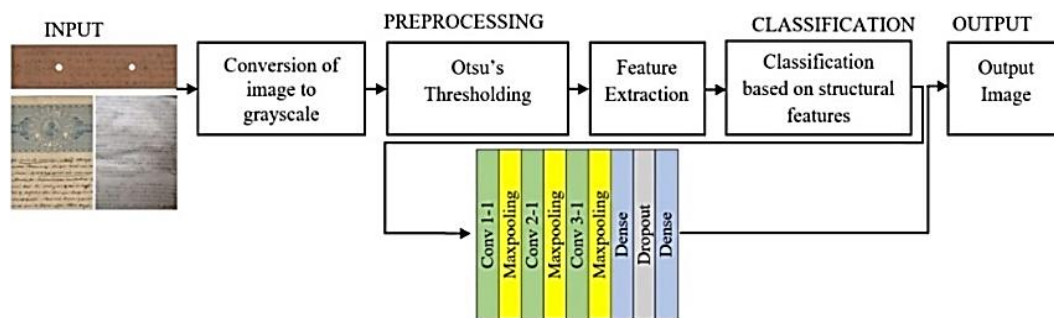

Figure 1. Classification using CNN

2.4. Classification using fine-tuned VGG-16

In Figure 2, the enhanced image is passed through the VGG-16 network which is a stack of two convolutional layers having 64 filters of size 3×3, from which it is passed to max pooling layer of filter size 2×2. The image size is reduced to 112×112×64 after max pooling. The image is then passed to the next stack of two convolution layers and a max pooling layer where the same process is repeated and as a result, an output image of size 56×56×128 is obtained. This is followed by the next stack containing three convolution layers of kernel size 256 which makes the output size 28×28×256. The next two consecutive stacks again contain three convolution layers and 512 filters each. After the image passes through the two stacks, the output will be of size 7×7×512. The obtained output is flattened and passed to a stack of three fully connected layers. The model is fine-14 by replacing the final fully connected dense layer which serves as the output layer.
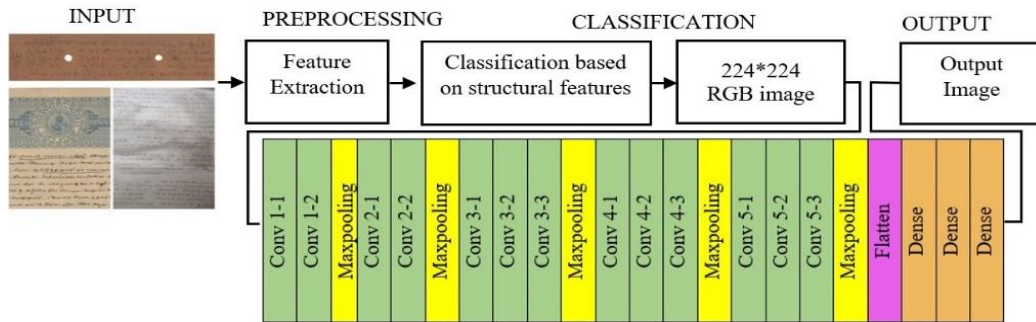

Figure 2. Classification using fine-tuned VGG-16

Figure 3 depicts pseudo code of the VGG-16 model. The processing method displays how the image is processed by the VGG-16 model. In the figure, I_{img} stands for the input image which is of size 224×224 which is fed into the VGG-16 model, Stack stands for the stack of convolutional and MaxPooling layers, conv stands for convolutional layer, L stands for layer, Maxpool is the abbreviation for MaxPooling layer, Flat is used to denote flattening layer and O_{img} is the output image.

$$I_{img}(224 * 224) \rightarrow VGG - 16$$

```
processing VGG-16()
{
  Stack L (1 to 5)
  {
    Conv L(1 to 13)
      Maxpool L(1 to 5)
  }
  Flat L (1 o 1)
  Dense (1 to 3)
}
O_img
```

Figure 3. VGG-16 pseudo code

2.5. Classification using modified AlexNet

The AlexNet model in Figure 4 consists of five stacks of convolutional layers activated by ReLU activation. Each convolution layer is followed by a batch normalization layer. The input image if size 227×227 is passed to the input convolution layer with 32 filters of size 11×11 with an activation function ReLU which is then forwarded to the batch normalization layer. The output image from the first stack of layers is then sent to a MaxPooling layer and further forwarded to the next stack of layers. Here the number of filters of the convolutional layer is increased to 64 and the filter size becomes 5×5. After MaxPooling, the

modified image is sent to the following stack of three convolutional layers with 128, 128, and 256 filters respectively of 3×3 size each. The image is then flattened and forwarded to a couple of dense layers each with ReLU activation. It is then forwarded to the output layer that has been trained for classifying the three classes. The output layer is activated by SoftMax activation.
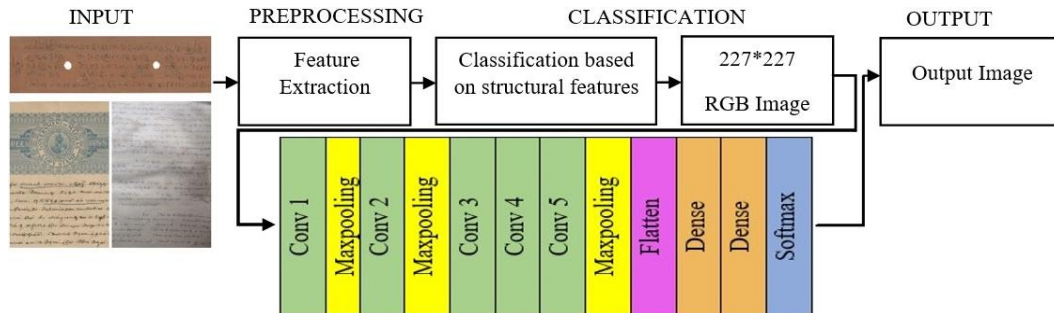

Figure 4. Classification using modified AlexNet

3. RESULTS AND DISCUSSION

The experiment was carried out on a dataset consisting of a total of 4,500 images belonging to three different classes. Out of which 3,600 images were used for training, 705 images for testing and 195 images for predictions. The models were tested on the dataset and the performance of each model has been assessed based on the performance evaluation metrics. Tables 2 to 4 show the performance evaluation metrics of each model.

<table>
<tr><td colspan="5">Table 2. Performance evaluation result of
classification using CNN</td></tr>
<tr><td>Epochs</td><td>Accuracy</td><td>Precision</td><td>Recall</td><td>F1 score</td></tr>
<tr><td>1</td><td>91.34</td><td>0.590</td><td>0.670</td><td>0.590</td></tr>
<tr><td>2</td><td>94.58</td><td>0.590</td><td>0.860</td><td>0.660</td></tr>
<tr><td>3</td><td>93.22</td><td>0.630</td><td>0.710</td><td>0.560</td></tr>
<tr><td>4</td><td>94.55</td><td>0.570</td><td>0.570</td><td>0.340</td></tr>
<tr><td>5</td><td>94.85</td><td>0.500</td><td>0.620</td><td>0.500</td></tr>
<tr><td>6</td><td>96.23</td><td>0.420</td><td>0.560</td><td>0.400</td></tr>
<tr><td>7</td><td>91.20</td><td>0.590</td><td>0.600</td><td>0.490</td></tr>
<tr><td>8</td><td>92.05</td><td>0.590</td><td>0.600</td><td>0.490</td></tr>
<tr><td>9</td><td>90.56</td><td>0.690</td><td>0.650</td><td>0.550</td></tr>
<tr><td>10</td><td>88.11</td><td>0.750</td><td>0.600</td><td>0.590</td></tr>
</table>

<table>
<tr><td colspan="5">Table 3. Performance evaluation result of
classification using fine-tuned VGG-16</td></tr>
<tr><td>Epochs</td><td>Accuracy</td><td>Precision</td><td>Recall</td><td>F1 score</td></tr>
<tr><td>1</td><td>88.6</td><td>0.850</td><td>0. 850</td><td>0.850</td></tr>
<tr><td>2</td><td>84.8</td><td>0.900</td><td>0.900</td><td>0.900</td></tr>
<tr><td>3</td><td>81.1</td><td>0.850</td><td>0.850</td><td>0.850</td></tr>
<tr><td>4</td><td>85.6</td><td>0.900</td><td>0.900</td><td>0.900</td></tr>
<tr><td>5</td><td>91.6</td><td>0.800</td><td>0.800</td><td>0.800</td></tr>
<tr><td>6</td><td>94.2</td><td>0.850</td><td>0.850</td><td>0.850</td></tr>
<tr><td>7</td><td>99.7</td><td>0.950</td><td>0.950</td><td>0.950</td></tr>
<tr><td>8</td><td>95.5</td><td>0.900</td><td>0.900</td><td>0.900</td></tr>
<tr><td>9</td><td>93.5</td><td>0.880</td><td>0.880</td><td>0.880</td></tr>
<tr><td>10</td><td>94.8</td><td>0.840</td><td>0.840</td><td>0.840</td></tr>
</table>

Table 4. Performance evaluation result of classification using AlexNet

Epochs	Accuracy	Precision	Recall	F1 score
1	43.8	0.390	0.490	0.490
2	68.3	0.590	0.860	0.660
3	85.3	0.630	0.710	0.500
4	91.2	0.820	0.670	0.540
5	95.0	0.410	0.620	0.440
6	94.4	0.720	0.880	0.600
7	95.1	0.590	0.600	0.490
8	92.4	0.590	0.600	0.490
9	82.0	0.690	0.650	0.550
10	86.8	0.630	0.700	0.690

Table 2 displays the accuracy, precision, recall, and F1-score values obtained by the proposed CNN method. The accuracy is found to be increasing gradually with the number of epochs. The maximum accuracy is obtained at the 6th epoch after which the accuracy tends to decrease gradually. The precision, recall, and F1-score values varied inconsistently with each epoch.

From Table 3, it is observed that the VGG-16 model accuracy tends to decrease in the initial epochs. However, from the 4th epoch, accuracy increases gradually. The highest accuracy is obtained at the 7th epoch after which a decline in the accuracy values can be witnessed. The model could successfully classify the input documents into the respective classes and achieved a very high accuracy of 99.7% on the dataset. The values such as precision, recall, and F1-score were found to be high and were balanced as the epochs varied.

Table 4 shows the performance of the modified AlexNet on the test dataset. The model accuracy improves as the epoch's increases and reaches a maximum of 95.5% at the 7th epoch. However, a decline was observed in accuracy after the 7th epoch. The precision, recall, and F1-score values were found to be decreasing after the sixth epoch.

Figure 5 depicts the accuracy, loss values of the CNN, fine-tuned VGG-16 and modified AlexNet models while training. From the accuracy loss graph of CNN, it can be inferred that the accuracy of the model increases with the epoch and reaches a very high value near 90%. Meanwhile, the loss reduces after the initial epoch, it finally reaches a very low value. For the VGG-16 model as the number of epoch's increases, the accuracy increases gradually and reaches a value in the range of 80 to 95% whereas the loss reduces to a value in the range 0.4% to 0.6%. Finally, the performance of the modified AlexNet model is depicted in the third graph. It is observed that the loss sharply declined in the model loss with the decrease in the epochs. Meanwhile, accuracy is found to be increasing with the number of epochs and reaches a value in the range 90 to 100%.
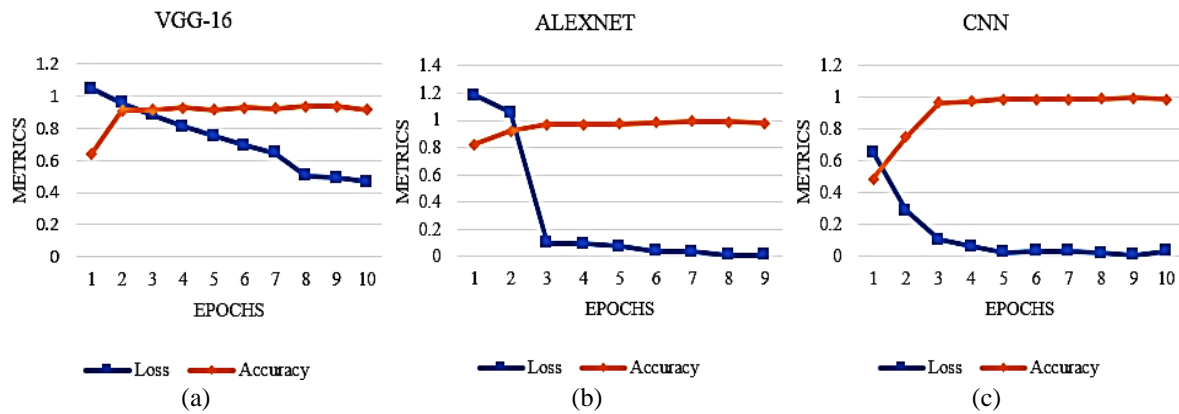

Figure 5. The training accuracy, loss graphs of (a) the CNN model, (b) fine-tuned VGG-16 model, and (c) modified AlexNet model

Table 5 shows the misclassifications by AlexNet. Compared to CNN and VGG-16 model, it was observed that in AlexNet notebook images were often confused with agreement images. A few palm leaf images were also misclassified as notebook images.

The graph in Figure 6 depicts the extent of homogeneity among the datasets of the three different classes.it clearly shows the structure wise similarity among notebook images, agreement copies as well as palm leaves. From the graph we can conclude that notebook and agreement images have more similarity due to which more documents are misclassified as notebook instead of agreement copies and agreements to notebook.

Table 5. Misclassifications by AlexNet model

| S.no	Document	Misclassification
1	Agreement	Classified as notebook
2	Agreement	Classified as notebook
3	Palmleaf	Classified as agreement
4	Notebook	Classified as agreement
5	Palmleaf	Classified as notebook
6	Agreement	Classified as notebook
7	Notebook	Classified as agreement
8	Notebook	Classified as agreement
9	Palmleaf	Classified as notebook
10	Agreement	Classified as notebook

Homogeneity of classes
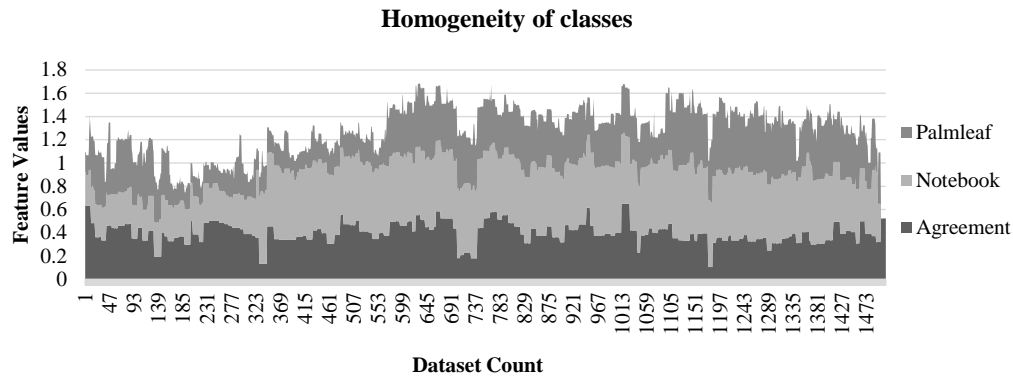
Figure 6. The homogeneity values between the three classes

4. CONCLUSION

The fine-tuned VGG-16 model was found to be achieving dominant results over the other two models with a remarkable accuracy rate of 99.7%. The CNN model achieved a good accuracy score of 96% whereas the modified AlexNet achieved an impressive accuracy of 95%. The AlexNet was found to give more misclassified results. The proposed approach can be used to perform document classification based on structural and visual features. This method of automatic classification of documents can be used to replace manual classification of documents for purposes such as document digitization, and cataloguing. The future work is to increase the number of classes by including several different categories of documents and to perform intra class classification based on textual as well as spectral contents.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Z. Mushtaq, S.-F. Su, and Q.-V. Tran, "Spectral images based environmental sound classification using CNN with meaningful data augmentation," *Applied Acoustics*, vol. 172, Jan. 2021, doi: 10.1016/j.apacoust.2020.107581.

[2]   R. Arief, A. Benny Mutiara, T. Maulana Kusuma, and H. Hustinawaty, "Automated hierarchical classification of scanned documents using convolutional neural network and regular expression," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, pp. 1018–1029, Feb. 2022, doi: 10.11591/ijece.v12i1.pp1018-1029.

[3]   M. M. Rahman, R. Sadik, and A. A. Biswas, "Bangla document classification using character level deep learning," in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Oct. 2020, pp. 1–6, doi: 10.1109/ISMSIT50672.2020.9254416.

[4]   Y. Quan, Y. Tong, W. Feng, G. Dauphin, W. Huang, and M. Xing, "A novel image fusion method of multi-spectral and SAR images for land cover Classification," *Remote Sensing*, vol. 12, no. 22, Nov. 2020, doi: 10.3390/rs12223801.

[5]   J. R. Prieto, V. Bosch, E. Vidal, C. Alonso, M. C. Orcero, and L. Marquez, "Textual-content-based classification of bundles of untranscribed manuscript images," in *2020 25th International Conference on Pattern Recognition (ICPR)*, Jan. 2021, pp. 3162–3169, doi: 10.1109/ICPR48806.2021.9412688.

[6]   I. M. Nasir *et al.*, "Pearson correlation-based feature selection for document classification using balanced training," *Sensors*, vol. 20, no. 23, Nov. 2020, doi: 10.3390/s20236793.

[7]   S. Bakkali, Z. Ming, M. Coustaty, and M. Rusinol, "Cross-modal deep networks for document image classification," in *2020 IEEE International Conference on Image Processing (ICIP)*, Oct. 2020, pp. 2556–2560, doi: 10.1109/ICIP40778.2020.9191268.

[8]   J. K. Mandivarapu, E. Bunch, Q. You, and G. Fung, "Efficient document image classification using region-based graph neural network," *preprint arXiv:2106.13802*, Jun. 2021, [Online]. Available: http://arxiv.org/abs/2106.13802.

[9]   G. Chen, L. Wang, and M. M. Kamruzzaman, "Spectral classification of ecological spatial polarization SAR image based on target decomposition algorithm and machine learning," *Neural Computing and Applications*, vol. 32, no. 10, pp. 5449–5460, May 2020, doi: 10.1007/s00521-019-04624-9.

[10] A. Paul, S. Bhoumik, and N. Chaki, "SSNET: an improved deep hybrid network for hyperspectral image classification," *Neural Computing and Applications*, vol. 33, no. 5, pp. 1575–1585, Mar. 2021, doi: 10.1007/s00521-020-05069-1.

[11] A. Sha, B. Wang, X. Wu, and L. Zhang, "Semisupervised classification for hyperspectral images using graph attention networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 157–161, Jan. 2021, doi: 10.1109/LGRS.2020.2966239.

[12] M. Gayathri and R. J. Kannan, "Ontology based concept extraction and classification of ayurvedic documents," *Procedia Computer Science*, vol. 172, pp. 511–516, 2020, doi: 10.1016/j.procs.2020.05.061.

[13] B. Jimenez Gutierrez, J. Zeng, D. Zhang, P. Zhang, and Y. Su, "Document classification for COVID-19 literature," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020, pp. 3715–3722, doi: 10.18653/v1/2020.findings-emnlp.332.

[14] M. Hassanpour and H. Malek, "Learning document image features with squeeze net convolutional neural network," *International Journal of Engineering*, vol. 33, no. 7, Jul. 2020, doi: 10.5829/ije.2020.33.07a.05.

[15] S. Kanchi, A. Pagani, H. Mokayed, M. Liwicki, D. Stricker, and M. Z. Afzal, "EmmDoc classifier: Efficient multimodal document image classifier for scarce data," *Applied Sciences*, vol. 12, no. 3, Jan. 2022, doi: 10.3390/app12031457.

[16] P. Hemmer, N. Kühl, and J. Schöffer, "DEAL: Deep evidential active learning for image classification," in *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2022, pp. 171–192, doi: 10.1007/978-981-16-3357-7_7.

[17] R. Indraswari, R. Rokhana, and W. Herulambang, "Melanoma image classification based on MobileNetV2 network," *Procedia Computer Science*, vol. 197, pp. 198–207, 2022, doi: 10.1016/j.procs.2021.12.132.

[18] M. Ahmed, P. Chakraborty, and T. Choudhury, "Bangla document categorization using deep RNN model with attention mechanism," in *Cyber Intelligence and Information Retrieval*, 2022, pp. 137–147.

[19] Z. Pan, P. Soong, and S. Rafatirad, "Ontology-driven scientific literature classification using clustering and self-supervised learning," *EasyChair Preprint*, pp. 1–19, 2022.

[20] S. Jiang, J. Hu, C. L. Magee, and J. Luo, "Deep learning for technical document classification," *IEEE Transactions on Engineering Management*, pp. 1–17, 2022, doi: 10.1109/TEM.2022.3152216.

[21] C. Zhao, B. Qin, S. Feng, and W. Zhu, "Multiple superpixel graphs learning based on adaptive multiscale segmentation for hyperspectral image classification," *Remote Sensing*, vol. 14, no. 3, Jan. 2022, doi: 10.3390/rs14030681.

[22] S. Swetanisha, A. R. Panda, and D. K. Behera, "Land use/land cover classification using machine learning models," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, pp. 2040–2046, Apr. 2022, doi: 10.11591/ijece.v12i2.pp2040-2046.

[23] C. Liu, J. Li, Q. Tang, J. Qi, and X. Zhou, "Classifying the Nunivak Island coastline using the random forest Integration of the sentinel-2 and ICESat-2 data," *Land*, vol. 11, no. 2, Feb. 2022, doi: 10.3390/land11020240.

[24] İ. Yelmen, "Multi-class document classification based on deep neural network and Word2Vec," *Journal of Aeronautics and Space Technologies*, vol. 15, no. 1, pp. 59–6, 2022.

[25] P. M. Rajegowda and P. Balamurugan, "A neural network approach to identify hyperspectral image content," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 4, pp. 2115–2125, Aug. 2018, doi: 10.11591/ijece.v8i4.pp2115-2125.

[26] C. Shi, D. Liao, T. Zhang, and L. Wang, "Hyperspectral image classification based on 3D coordination attention mechanism network," *Remote Sensing*, vol. 14, no. 3, Jan. 2022, doi: 10.3390/rs14030608.

[27] C. Ding *et al.*, "Hyperspectral image classification promotion using clustering inspired active learning," *Remote Sensing*, vol. 14, no. 3, Jan. 2022, doi: 10.3390/rs14030596.

[28] B. N. B. Jayakumari and A. S. Nair, "Ancient horoscopic palm leaf binarization using A deep binarization model - RESNET," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Apr. 2021, pp. 1524–1529, doi: 10.1109/ICCMC51019.2021.9418461.

[29] A. J. Moshayedi, A. S. Roy, A. Kolahdooz, and Y. Shuxin, "Deep learning application pros and cons over algorithm," *EAI Endorsed Transactions on AI and Robotics*, vol. 1, pp. 1–13, Feb. 2022, doi: 10.4108/airo.v1i.19.

[30] N. Sharma, "Vedic literature in Malayalam," Ayur Malayalam Notes, 2021. https://archive.org/details/AyurMalayalamNotes (accessed Aug. 07, 2021).

BIOGRAPHIES OF AUTHORS

Bipin Nair Balakrishnan Jayakumari 🆔 📊 SC 🔵 is currently an assistant professor at the Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwavidyapeetham, Mysuru Campus. His research interests include document image processing, medical image processing, and image analysis. He has published more than 60 research papers in reputable national and international journals. All those research papers are indexed in leading indexing databases such as Scopus, and SCI. He can be contacted at bipin.bj.nair@gmail.com.

Amel Thomas Kavana 🆔 📊 SC 🔵 is a final year student of Integrated MCA pursuing a dissertation in the area of document image classification and deep learning at Amrita School of Arts and Sciences, Mysuru Campus. His area of research includes document image processing and image classification. He can be contacted at amel.t.k05@gmail.com.