STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, November 2025, pp 2624–2642.
Published online in International Academic Press (www.IAPress.org)

Ensemble Learning and K-means Models for Lung and Colon Cancer Classification

Abdelwahid Oubaalla [1,*], Hicham El Moubtahij [2], Nabil EL AKKAD [1]

[1]*Applied Sciences and Emerging Technologies Laboratory, Sidi Mohamed Ben Abdellah University, Fez, Morocco*
[2]*Higher School of Technology, University of Ibn Zohr, Agadir, Morocco*

Abstract According to World Health Organization (WHO) statistics, cancer remains one of the leading causes of death worldwide. The highest number of cancer-related deaths is caused by lung cancer, with approximately 1.8 million fatalities (18.7%), followed by colorectal cancer, responsible for around 900,000 deaths (9.3%). These death rates are increasing in developing countries, where human, material, and technological resources for early detection are sometimes limited. In Morocco, for example, according to WHO statistics, the mortality rate for lung cancer stands at 21.6%. Diagnosis of histopathological images is one of the most effective ways of confirming or denying the existence of this type of cancer. Traditionally, this analysis is done manually by pathologists, which makes this process time-consuming and the outcome largely depends on the expertise of the pathologist. Automating this process can enable early detection and considerably increase the chances of cure. Thanks to the remarkable results achieved using machine learning techniques, many research projects have attempted to capitalize on these advances and apply them to automate and improve cancer detection accuracy. Despite advancements in deep learning-based classification, achieving consistently high accuracy remains a challenge. In this paper, we propose a new approach that uses the K-means model and gamma correction function to preprocess histopathological images from the LC25000 dataset, and transfer learning and ensemble learning to enhance the classification performance. We have combined two models based on VGG16 and DenseNet pre-trained models. This approach enabled us to achieve an accuracy of 99.96%, which illustrates the importance of combining unsupervised models, transfer learning and ensemble learning to improve the accuracy of histopathological images classification.

Keywords Histopathological Image, Lung and Colon Cancer, Deep Learning, Ensemble Learning, K-means Clustering, Gamma Correction.

DOI: 10.19139/soic-2310-5070-2833

1. Introduction

Cancer remains one of the world's leading causes of death, posing a major challenge for healthcare systems due to the lack or high cost of early diagnosis and analysis. According to recent statistics from the International Agency for Research on Cancer (IARC), there will be almost 20 million new cases of cancer by 2022, of which 12.4% will be lung cancer and 9.6% colon cancer, which caused around 9.7 million deaths [1]. Lung cancer alone accounts for 18.7% of cancer deaths, while colon cancer is responsible for 9.3%. Given these alarming figures, developing effective early diagnostic techniques has become increasingly urgent to improve survival rates and reduce the adverse effects associated with these diseases.

Cancer begins with abnormal transformations in cell function. These cells undergo genetic mutations that disrupt their normal cycle of division and growth. These life-cycle alterations lead to unlimited division and growth of these cells, culminating in the creation of a cancerous tumor. These cells undergo observable alterations concerning

*Correspondence to: Abdelwahid Oubaalla: abdelwahid.oubaalla@usmba.ac.ma). Applied Sciences and Emerging Technologies Laboratory, Sidi Mohamed Ben Abdellah University, Fez, Morocco.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

variations in color, shape, and size of cells and nuclei [2]. According to health research, such changes are linked to risk factors including smoking, chemical exposure, obesity, excessive alcohol consumption, physical inactivity, poor nutrition, genetic predisposition, and infections. [3].

Early diagnosis of cancer is essential to maximize the chances of therapeutic success. Currently, several investigative methods are used, including analysis of tumor markers in blood tests, medical imaging (MRI, CT, ultrasound), and histopathological examination by biopsy [4]. Among these techniques, histopathological analysis remains the reference standard, despite its time-consuming nature and heavy dependence on the expertise of pathologists. Associated limitations include intra- and inter-observer variability and diagnostic errors, underlining the importance of reliable and accurate automation of this process [5]. In addition, these traditional methods are costly, invasive and may present risks such as false negatives or positives, thus limiting their effectiveness in widespread screening programs [6].

The use of Artificial Intelligence (AI) technologies in various fields, including healthcare, has become a trend over the last fifteen years. These technologies have become more and more widely used, particularly following the success of the convolutional neural network (CNN) AlexNet [7] in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [8], achieving a top-5 error of 15.3%. Since then, Machine Learning (ML) and Deep Learning (DL) have continued to deliver impressive results. They are used in the automation of various tasks such as image classification, segmentation [9], object recognition, text processing, sentiment analysis [10], cybersecurity applications, and so on.

AI technologies, particularly deep DL, have become a trend in Clinical Decision Support Systems (CDSS). These technologies have been adopted thanks to their revolutionary success in various fields, particularly computer vision. Also in the medical field, they have shown promising results in the accurate and rapid identification of cancerous pathologies by extracting complex and hidden features that are not extracted or detected by the naked eye [11]. The application of CNNs to histopathological images has proved particularly effective in distinguishing malignant from normal tissue, helping to significantly improve diagnostic accuracy and analysis efficiency [12]. The integration of DL models into diagnostic processes has enabled us to significantly reduce time and associated costs, while improving reproducibility of results and avoiding human error.

Despite the impressive results achieved using DL in cancer detection, several challenges persist. The lack or insufficiency of such datasets represents the most notable challenge. This is due to the complexity of the data collection and labeling process, which requires expert pathologists. The scarcity of high-quality datasets, such as the LC25000 dataset [13], poses a significant challenge to the effective generalization of DL models in real clinical settings. In addition, improved model performance often depends on image preprocessing techniques such as gamma correction and contrast enhancement [14], which play a vital role in revealing subtle visual features important for diagnosis [15]. The difficulty of obtaining high-quality, representative histopathological images, not least because of the heterogeneity of biological samples, is also a major obstacle.

In this paper, we explore the impact of two popular DL techniques, namely transfer learning and ensemble learning, to classify histopathological images of lung and colon cancer. Our approach consists of three steps. The first step is to improve image quality using the gamma correction function and the unsupervised learning model K-means.The second step consists of training our classification models. The final step is to combine the models from the second step into an ensemble learning model. The first two steps have been illustrated in previous work, and our aim in this paper is to show how the ensemble learning technique can improve classification performance. The aim of this study is therefore to provide a comprehensive, robust, and accurate solution aimed at automating the diagnosis of lung and colon cancers, likely to transform clinical practice by reducing diagnostic delays, improving operational efficiency, decreasing the costs incurred by traditional methods, and contributing to better therapeutic management of patients, particularly in resource-limited regions.

The rest of this paper is organized as follows: Section 2 explores previous research studies, with a focus on hybrid and ensemble learning approaches applied to the LC25000 dataset. Section 3 examines the various methods and materials utilized in this research. Section 4 presents the results, followed by a comprehensive discussion. Section 5 provides the conclusion.

2. Related Work

Histopathological image processing is one of the key elements in early cancer detection. The use of ML and DL in this field represents a very active area of research. Several surveys and literature reviews have explored the various advances made in this field [16], [17].

In this section, we provide an overview of research concerning the classification of histopathological images, focusing specifically on techniques that combine ensemble learning models. The selection of papers is based on their indexing in reputable sources and the performance metrics achieved. Specifically, we have included only articles reporting an accuracy exceeding 90%, ensuring a high standard of reliability and effectiveness in the methodologies discussed.

To explore the use of DL applications in colon and lung cancer detection, Patharia et al. [18] have conducted a comprehensive literature review. This review is focused on research carried out on the LC25000 dataset. This review systematically analyzed various DL architectures, including CNNs and transfer learning techniques. These architectures were evaluated based on their performance and clinical applicability. This review revealed the various challenges involved, particularly the scarcity of datasets available for research and the need for cooperation between different stakeholders to develop robust and efficient architectures that can be implemented in healthcare environments.

Echle et al. [19] survey the use of DL in cancer pathology. They have demonstrated that histopathology images reveal valuable and frequently unexploited information. Two types of tasks are distinguished in their research: simple tasks such as tumor detection or subtyping, and advanced tasks such as mutation or survival prediction. These approaches can help to achieve more accurate clinical decisions. The authors also highlight technical challenges, including image size, lack of standardization, and difficulty in generating new datasets. Despite this, DL has great potential. It could lighten the workload of pathologists and improve diagnostic accuracy. Clinical validation remains necessary for regular use.

The two-dimensional discrete Fourier transform and the two-dimensional discrete wavelet transform were used by Masud et al. [20] to capture, respectively, low-frequency and high-frequency details. These features were combined to create a complete representation of each image before being used to enter CNN modes. This approach achieved an accuracy of 96.33%.

Mesut Toğaçar [21] has proposed an approach based on using the pre-trained model for feature extraction of LC25000 images. Then, these features are optimized using the Equilibrium and Manta Ray Foraging Optimization (MRFO) algorithms. Subsequently, these enhanced images are used to train the Support Vector Machine (SVM) model, which has achieved an accuracy of 99.69%.

The combination of the extracted features, using First-Order Statistical Features, Gray-Level Co-occurrence Matrix, and Hu Invariant Moments enabled Hage Chehade et al. [22] to achieve an accuracy of 99% and an F1-score of 98.8% using XGBoost model for LC25000 images classification. Other models were tested in their research, namely SVM , Random Forest , Linear Discriminant Analysis and Multilayer Perceptron, but XGBoost was the one that achieved the best performance.

Talukder et al. [23] have used MobileNet for feature extraction and ensemble soft voting classifier for classification. This approach has achieved an accuracy of 99.30%. Multiple transfer learning models namely VGG16, VGG19, MobileNet, DenseNet169, and DenseNet201 were trained and evaluated, but the best accuracy was achieved by MobileNet.

A fast and efficient approach to classifying histopathological images of the lung was proposed in [24]. It combines a lightweight CNN model with an optimized LightGBM classifier. CNN automatically extracts image features after preprocessing. Then, LightGBM, thanks to its multiple threads, has ensured rapid classification and achieved an accuracy and sensitivity of 99.60% on LC25000 dataset. Farhadipour [25] proposes a method for detecting lung and colon cancers. Images are first enhanced with adaptive histogram equalization. Then, eight CNN models are tested. Four are in serial architecture: Cifar10Net, AlexNet, VGG19, and DarkNet19. The other four are in DAG architecture: GoogLeNet, SqueezeNet, EfficientNet-b0, and InceptionResNetV2. The VGG19 model gives the best results: 99.84% accuracy, with very good scores for sensitivity, specificity, F1-score and AUC. SqueezeNet

is the fastest. It offers a good compromise with 99.58% accuracy and just 0.56 ms per image. This method therefore combines quality and speed in order to better detect cancerous areas.

Bishnoi and Goel [26] propose a model called CD-CNN to detect lung cancer. It uses color and dilated convolutions to better classify images. Five color spaces were tested. The best is Hue-Saturation-Value (HSV). It gives the best results on three datasets: TCGA, CPTAC and LC25000. The model achieves between 97% and 99% accuracy. F1 scores are also very high, between 0.97 and 0.98. The AUC is as high as 0.984. The kappa score reaches 0.986. These results show that the model is fast, accurate and useful for diagnosis.

Attallah [27] proposed an approach that combines three lightweight CNNs (MobileNet, ResNet-18, and EfficientNetB0) by exploiting multiscale features extracted from two distinct layers (pooling and fully connected). Canonical correlation analysis and variable selection (ANOVA, Chi-Squared) are used for dimension reduction. This results in a compact and discriminative feature vector. This approach was evaluated on the LC25000 dataset and achieved 99.8% accuracy with a cubic SVM using only 50 selected features, while significantly reducing complexity compared to individual deep models.

Our contribution in this research consists in demonstrating, firstly, the importance of our k-means and gamma correction preprocessing approach in performance enhancement. On the other hand, the combination of our models using ensemble learning techniques has led to a substantial improvement in results.

3. Materials and Methods

In this section, we will present the different resources and techniques used in this research, as well as the methodology adopted. We first describe the dataset used, then the different preprocessing techniques applied to the images. Finally, we will explain the methodology adopted for the whole project.

3.1. LC25000 Dataset Description

Histopathological images are microscopic images taken from a slide of biological tissue. These tissues are often stained with dyes such as hematoxylin-eosin to separate the different components of the tissue. These images can be used to explore cellular and tissue architecture to observe abnormalities such as cancer cells, inflammation, or necrotic tissue. As part of our use of ML in the healthcare sector, they are used to train cancer detection models. In this research, we used the Lung and Colon Cancer Histopathological Image dataset (LC25000). This dataset is, publicly available online on the Kaggle and GitHub websites [13]. Its primary purpose is to support artificial intelligence (AI) based research for detecting lung and colon cancer. It contains 25000 color images of dimension 768 x 768 pixels divided into 5 classes (Figure 1) which are colon adenocarcinoma, normal colon tissue, lung adenocarcinoma, normal lung tissue and lung squamous cell carcinoma Each class comprises 5000 images, making it a balanced dataset, ideal for training ML models. These images were derived from 750 lung tissue images and 500 colon tissue images through data augmentation techniques, including left and right rotations as well as horizontal and vertical flips.

3.2. Preprocessing methods

The preprocessing phase involves preparing data for use in training models. In the case of image processing, this phase consists of improving image quality by removing noise, enhancing contrast, and highlighting features. In our approach, we used two image enhancement techniques: the K-means model and the gamma correction function.

K-means is an unsupervised method widely used in ML to classify data into a predefined number of groups called clusters. Each cluster groups data with similarities around a centroid representing the mean of the points assigned to it. The algorithm works iteratively, alternating between assigning points to the nearest centroids and updating the centroids until convergence is reached. Its main objective is to minimize intra-cluster variance, i.e. to make each group as homogeneous as possible. Easy to implement, fast and adapted to large volumes of data, K-means is used in many fields, notably in medical imaging to segment structures, in marketing for customer segmentation, and in computer vision to simplify the colors of an image. In our research, this method was applied

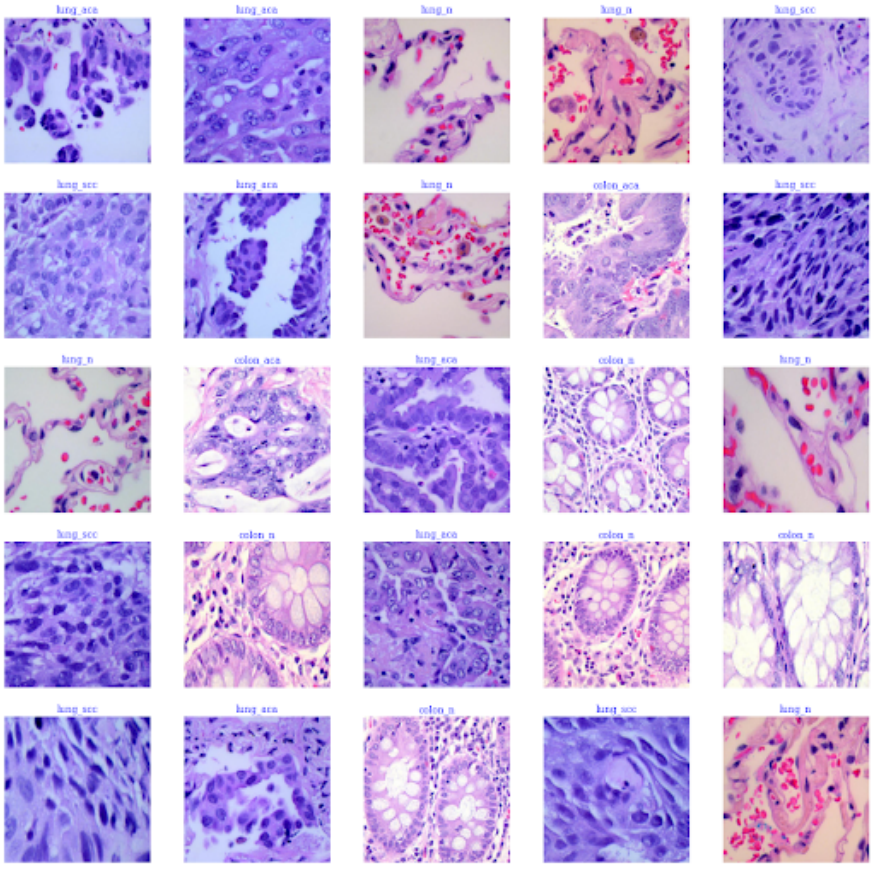

Figure 1. the Histopathological images of the LC25000 dataset

as a preprocessing technique to simplify histopathological images and enhance their visual clarity, which facilitates the task of classification by DL models.

In our approach, the choice of K-means clusters was not based on color similarity alone, so we have chosen to test K values between 5 and 30 to achieve optimal classification performance.

Gamma Correction is the second preprocessing method used in this approach to balance brightness. It is a non-linear mathematical transformation function that is applied to all pixels in an image, treating them as independent entities [28], to adjust the brightness.

$$g(x) = I_{\max} \left(\frac{I}{I_{\max}} \right)^{\gamma} \tag{1}$$

- I : the original pixel intensity.
- $I_{\mathbf{max}}$: the maximum intensity in the input image.
- γ : the gamma factor applied. If $\gamma < 1$, the image brightens; if $\gamma > 1$, it darkens; if $\gamma = 1$, nothing changes.

3.3. Classification Models

In our approach, we used two supervised techniques for image classification, namely transfer learning and ensemble learning.

Transfer learning is a powerful DL technique that uses the knowledge of a model pre-trained on a large dataset, such as the ImageNet. The model is then adapted to a new task with a smaller dataset. This approach reduces

training time and often improves performance. This method is very useful in healthcare tasks, where data is scarce and precious. The idea is to change and train the decision layers and keep the previous layers that have already learned to extract general image features such as horizontal and vertical lines, colors and shapes. In this work, we built our models based on two pre-trained models, VGG16 and DenseNet:

- **VGG (Visual Geometry Group)**: This is a model developed by Oxford University. It is used to detect and classify images [29]. Its architecture is simple: 13 convolution layers are divided into 5 blocks. After each block, there is a 2x2 max pooling layer. Then there are 3 fully connected layers. The input image must be 224 x 224 in size. The model classifies images into 1000 categories using a Softmax layer. Several versions are available, such as VGG16 and VGG19. They contain 16 and 19 layers respectively. This model is easy to use, but rather heavy: VGG16 has around 138 million parameters. It therefore requires a lot of memory and power. Despite this, it remains very popular for transfer learning.
- **DenseNet (Dense Convolutional Network)**: DenseNet is a more recent model. It consists of blocks called DenseBlock [30]. Each layer in a block uses the outputs of all preceding layers. This makes it easier to share information. The model learns more easily and avoids data loss. DenseNet is lighter than VGG. It has fewer parameters but gives very good results. It's also faster to train. Several versions are available, such as DenseNet-121 or DenseNet-169. It's an excellent choice for complex images, such as medical ones. Its dense architecture improves accuracy without increasing model size.

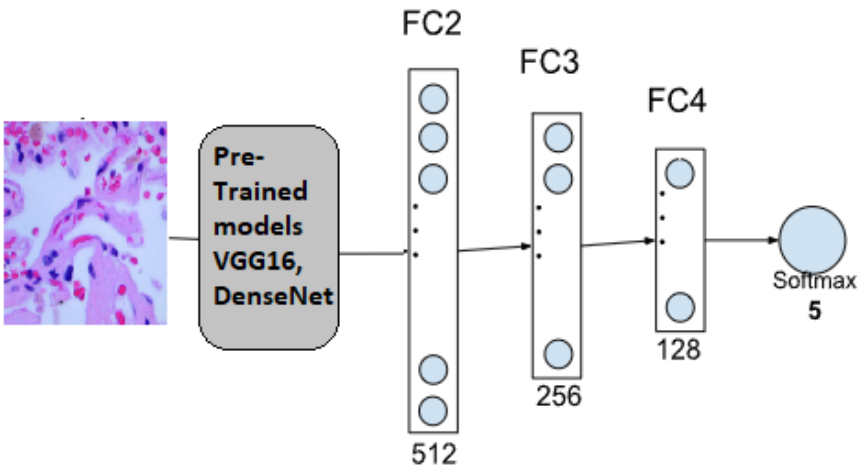

Figure 2. the Histopathological images of the LC25000 dataset

In our approach we created two models, one based on VGG16 and the other on DenseNet. The frozen layers of these pre-trained models, i.e. whose weights are not updated during learning, followed by fully connected layers with 512, 256 and 128 neurons, concluding with a decision layer for five classes (Figure 2).

Ensemble Learning is a technique that combines the knowledge of several models instead of just one. Each model learns in a slightly different way and may learn characteristics or hidden patterns that are unknown to the other models. Then, their results are combined to give a final answer. This makes prediction more stable and accurate. These models can be combined by averaging or maximizing the predictions or using a vote to choose the best answer. This method helps reduce errors. It works particularly well when the data is complex. For example, when classifying medical images, Ensemble Learning often gives better results. There are several techniques, such as bagging, boosting and stacking. All aim to get the best out of each model. In our approach we have chosen to combine our models based on VGG16 and DenseNet using the average predictions.

3.4. Metrics

For measuring enhancement in images, we used the RMS Contrast (Root Mean Square Contrast) and Edge Intensity metrics, which respectively assess overall contrast and edge sharpness, two key indicators of visual improvement. The equations of these metrics are:

$$CRMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_i - \bar{I}\right)^2} \tag{2}$$

- **CRMS** : The RMS contrast,
- **N** : The total number of pixels in the image,
- $\boldsymbol{I_i}$: The intensity of pixel *i*,
- $\bar{\boldsymbol{I}}$: The mean intensity of the image.

$$EI = \frac{1}{N}\sum_{i=1}^{N} M_i \tag{3}$$

- $\boldsymbol{EI}$: The Edge Intensity,
- **N** : The total number of pixels,
- $\boldsymbol{M_i}$: The magnitude of the gradient at pixel *i*, often computed as:

$$M_i = \sqrt{\left(G_x^i\right)^2 + \left(G_y^i\right)^2} \tag{4}$$

$\boldsymbol{G_x^i}$ and $\boldsymbol{G_y^i}$ are the horizontal and vertical gradient values at pixel *i*, typically obtained using the Sobel operator.

The performance of our classification models have been evaluated using the metrics: Accuracy, Precision, Recall and f1-score. The equations of these metrics are:

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F1\text{–}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

- **TP**: True Positive
- **TN**: True Negative
- **FP**: False Positive
- **FN**: False Negative

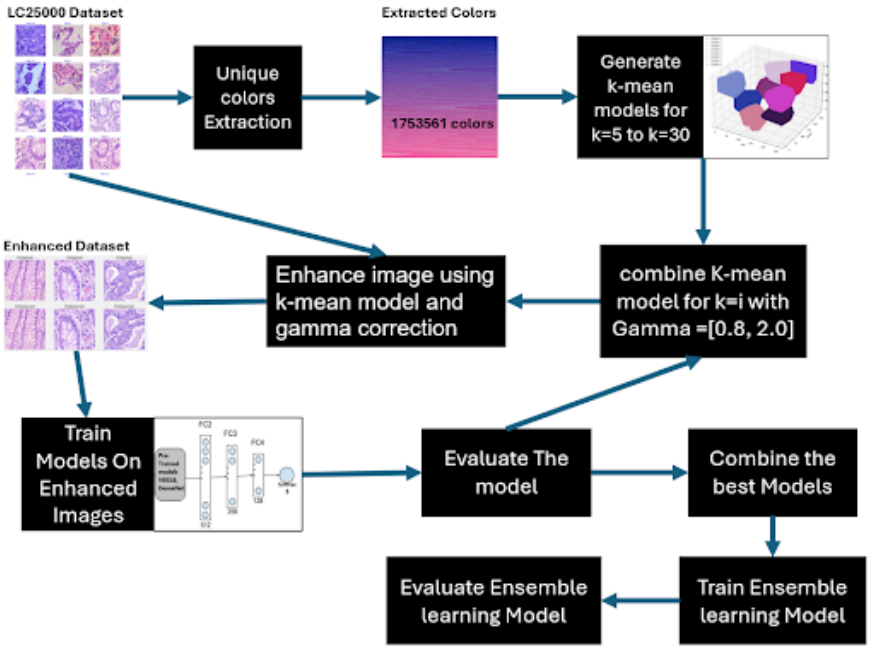

Figure 3. Our Methodology (the value "i" is changed in each iteration)

3.5. Proposed Methodology

In this work, we have adopted an iterative methodology that has been repeated several times until reaching the best models (Figure 3). First, we started with the preprocessing phase, in which we generated a combination of K-means and gamma correction function models with different parameters. Secondly, we trained our models based on transfer learning on the images enhanced by the first phase, and finally we combined the best models into a single model using the ensemble learning approach.

The first phase of our methodology was the selection of the best K-means models and gamma correction parameters. To achieve this, we have extracted the list of all the colors used in the images of the LC25000 dataset. Our objective in this stage is to reduce the number of colors used to facilitate the learning process of the classification models. The number of colors used, RGB values, is 1,753,561 colors (Figure 4).

We have created a dataset of extracted colors, without redundancy. This dataset was then used to select the best K value for K-means and the best gamma value for gamma correction function. The selection of K and gamma values cannot be automatic, as the clustering of colors does not depend on their similarities, but rather on the impact of this clustering on the classification after the enhancement of the images. For this purpose, our selection of optimal hyperparameters for the K-means method and gamma correction was carried out using an empirical search strategy [31]. We have chosen K values between 5 and 30 and created 25 K-means models. Each K-means model was combined with gamma values ranging from 0.2 to 0.8 with an iteration of 0.1. Although we have tested values outside these ranges, the performance was degraded. Each combination of K-means and gamma values is used to enhance the LC25000 dataset images before being used to train our transfer learning models based on VGG16 and DenseNet.

We have trained our models with all possible combinations and compared the performance achieved in order to select the best values for the K and gamma parameters.

Although we adopted an empirical approach, we then tried to understand why these hyperparameter values yielded the best performance. For example, comparing RMS contrast and edge intensity across different gamma values (Table 1) shows that RMS contrast and edge intensity decrease monotonically as the gamma value increases from 0.8 to 2.0. Too low gamma values (e.g., 0.8-0.9) cause excessive brightening of images, which

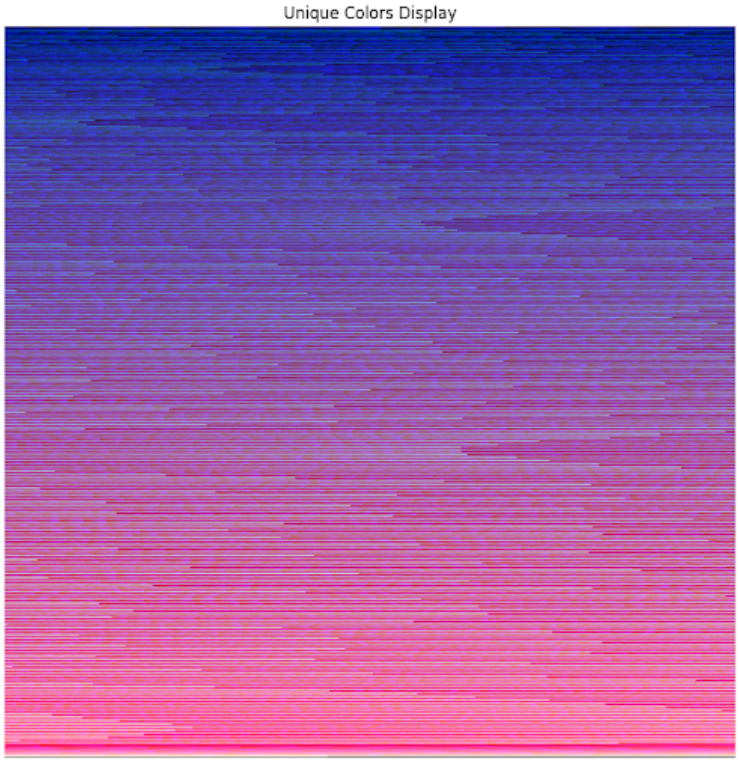

Figure 4. The Histopathological images of the LC25000 dataset

greatly increases contrast but risks saturating certain tissue structures. In contrast, too high gamma values (¿1.3) gradually darken images, reducing edge sharpness and potentially obscuring essential diagnostic details. The intermediate value gamma = 1.2 therefore represents an optimal compromise: it moderately improves brightness while maintaining good contour definition. Similarly, applying K-means clustering with K = 9 reduces color redundancy while retaining critical histopathological structures.

The best-performing models are combined in an ensemble learning model. Using the K-means models and the corresponding gamma correction function, we then trained and tested this model.

This methodology enabled us to improve classification performance thanks to the preprocessing phase and the ensemble learning approach. All the classification models used in this research were trained and tested on the LC25000 dataset before and after enhancement.

4. Results and Discussion

Training the K-means model in combination with the pre-trained VGG16 and DenseNet models enabled us to select the best hyperparameters and achieve optimal performance. The best hyperparameters achieved are K=9 and Gamma= 1.2. As a result, the number of colors used in the LC25000 dataset was reduced from 1753561 to 9 colors (Figure 5). Subsequently, this reduction will enhance the learning process of the classification models. The figure (Figure 6) shows a sample of the images before and after enhancement. The values K=9 and Gamma=1.2 were selected based on experimental results.

To validate and confirm the results obtained in this study and ensure that the results were not obtained by chance, we repeated the training process multiple times over several weeks using Kaggle Notebooks, a cloud-based computing environment that supports reproducible and collaborative research. This platform provides up to

Table 1. *RMS Contrast and Edge Intensity across different gamma values.*

Gamma	RMS Contrast	Edge Intensity
0.8	45.371389	105.217801
0.9	41.645878	96.542653
1.0	38.262816	88.496497
1.1	35.713858	82.430930
1.2	33.263105	76.977432
1.3	31.170363	72.881338
1.4	29.372071	69.842792
1.5	27.828180	67.613003
1.6	26.496518	65.977336
1.7	25.353136	64.790966
1.8	24.385819	63.973273
1.9	23.564390	63.415546
2.0	22.885659	63.089634

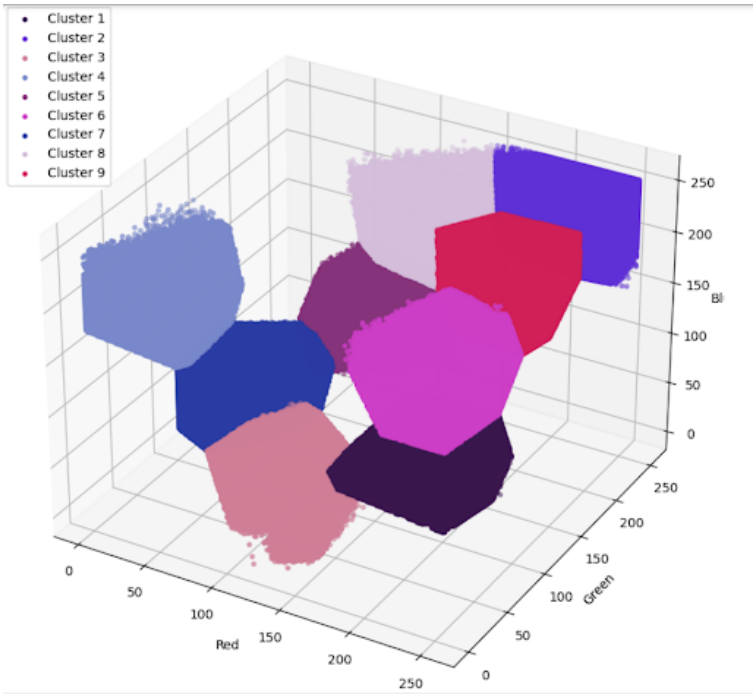

Figure 5. Colors clustering using K-means

32 hours per week of access to a machine equipped with 16 GB of memory and dual NVIDIA T4 GPUs. Model training was configured to stop early if accuracy did not improve for several consecutive epochs, or otherwise to terminate after a maximum of 40 epochs.

Models based on VGG16 and DenseNet were trained several times with and without the use of the preprocessing approach. Both initialized with ImageNet pre-trained weights, by averaging their prediction probabilities with equal weights. This strategy leverages complementary feature representations, reduces model variance, and maintains

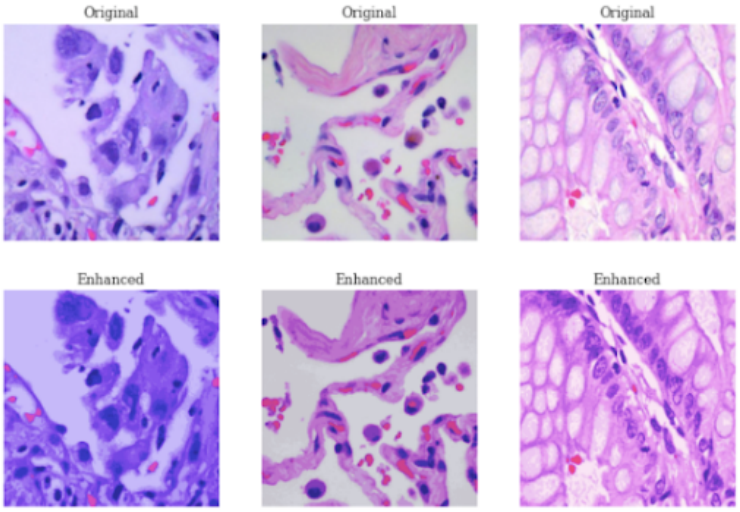

Figure 6. Colors clustering using K-means

low computational cost. The training of the models after image enhancement was carried out in several iterations, combining values of K ranging from k=5 to k=30 and values of Gamma varying between 0.8 and 2.0.

Preprocessing the images with the hyperparameters values of K=9 and gamma=1.2 illustrates an improvement in visual image quality. As evidenced in (Figure 7) , the RMS contrast and edge intensity increased. This indicates improved delineation of internal structures, which can facilitate the extraction of relevant features for automatic classification.

After obtaining the best results with K=9 and gamma=1.2, the training process was repeated more than 10 times to avoid adopting random results. At each repetition, the results were almost identical with negligible variance. This stability testifies to the great robustness of our approach and the absence of any significant random effect.

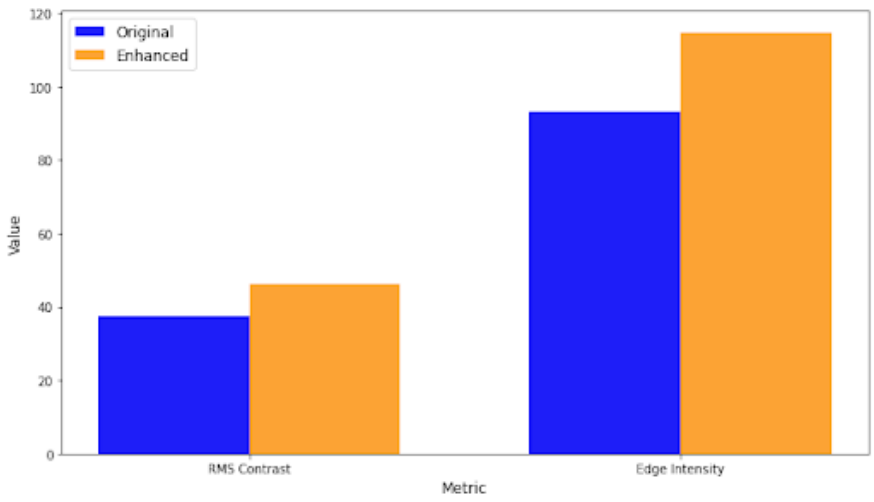

Figure 7. Comparison of RMS Contrast and Edge Intensity Between Original and Enhanced Histopathological Images

Training the models with K=9 and Gamma=1.2 resulted in the best performance for the VGG16 and DenseNet classification models, at 99.68% and 99.84% respectively. On the other hand, training these models without image

enhancement limited their accuracies to 99.44% and 99.64 respectively (Table 2). The model training phase was repeated several times to ensure the reliability of the results.

Training the VGG16-based model on the enhanced images required 112 minutes and 43 seconds, reaching its best performance at epoch 17 (Figure 8). The analysis of the confusion matrix on the test dataset indicates that the model made a single misclassification in the colon_aca class, 5 errors in lung_aca and a single error in lung_scc (Figure 9). The enhancement and prediction of a single image using this model take approximately 495 milliseconds.

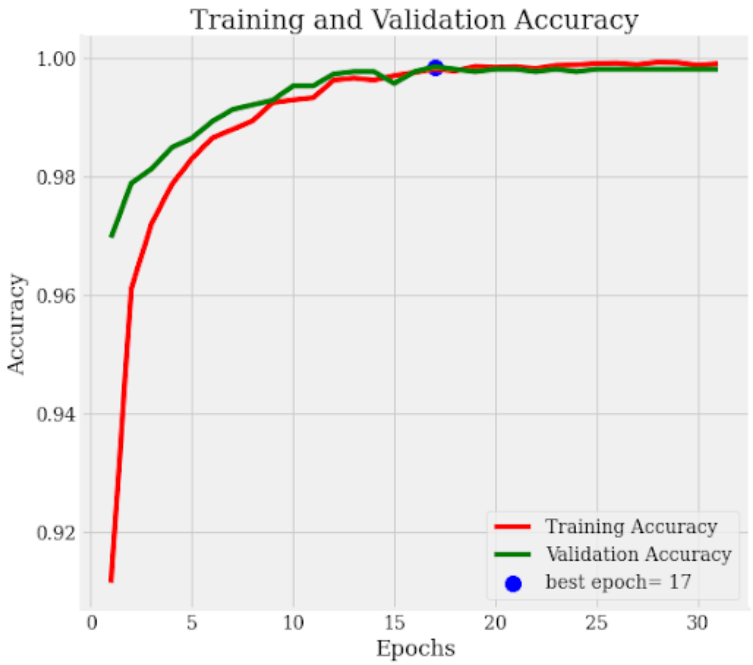

Figure 8. The learning evolution of VGG16-based model

The training of the DenseNet based model on the enhanced images required 69 minutes and 26 seconds, reaching its best value in training at epoch 23 (Figure 10). The analysis of the resultts of classification on test dataset, the confusion matrix shows that the model has made 5 misclassifications in the lung_aca class, and only one in lung_scc (Figure 11. This model requires approximately 518 milliseconds to enhance and classify a single image.

Our aim in this paper was to evaluate this approach in ensemble learning models. To do this, we combined the VGG16 and DenseNet based models, which have been achieved the best results, into a new ensemble learning model and then trained this model with and without the preprocessing approach. After training, the models are evaluated using the test and validation dataset and the complete dataset.

After training the ensemble learning models several times without the image enhancement step, the model achieved an accuracy of 99.96% in the validation dataset, 99.88% in the test dataset, and 99.98% in the full dataset. The average of the recall, precision, and F1-score values in the test dataset is 0.9988 (Table 3). the model training (Figure 12) lasted 3 minutes and 7 seconds. The best accuracy was achieved at the 4th epoch. The enhancement and the classification of single image using this ensemble model requires approximately 500 milliseconds.

By analyzing the results of classification on the test dataset, (Figure 13), we observe that the model performs well for all classes except for the classification of three images of the lung_aca class that have been classified as lung_scc.

Using enhanced images, with our K-means model and gamma correction function, to train our ensemble learning model improved performance considerably. we have achieved an accuracy of 99.96% in the test and validation datasets and an accuracy of 99.99% in the complete dataset. The average of the Recall Precision and F1-score values

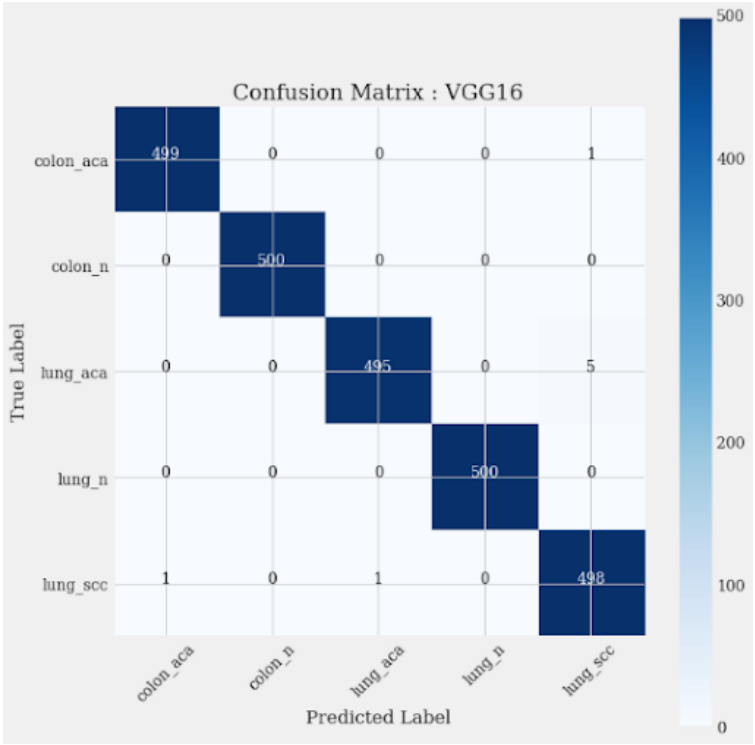

Figure 9. VGG16-based model Confusion Matrix

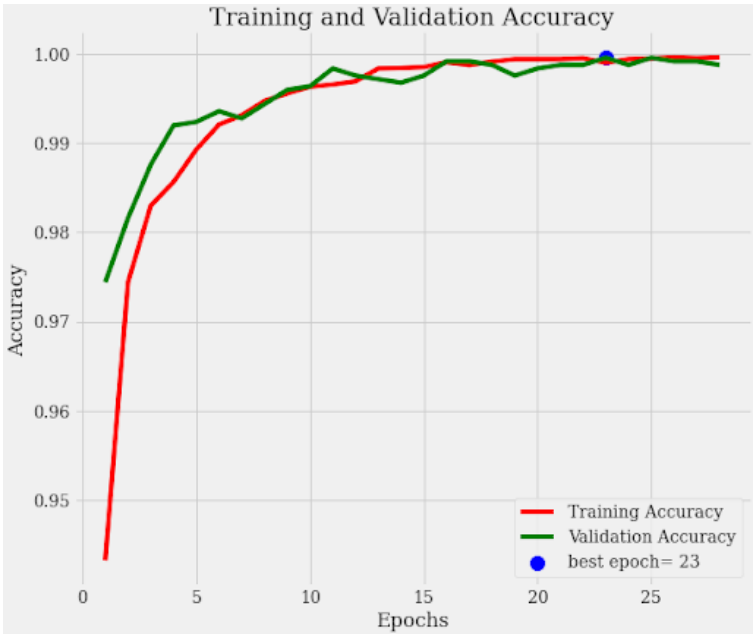

Figure 10. The learning evolution of DenseNet based model

in the test dataset is 0.9996 and the best accuracy was achieved in the 3rd epoch (Figure 14). The enhancement and the classification of single image using this ensemble model requires approximately 773 milliseconds.

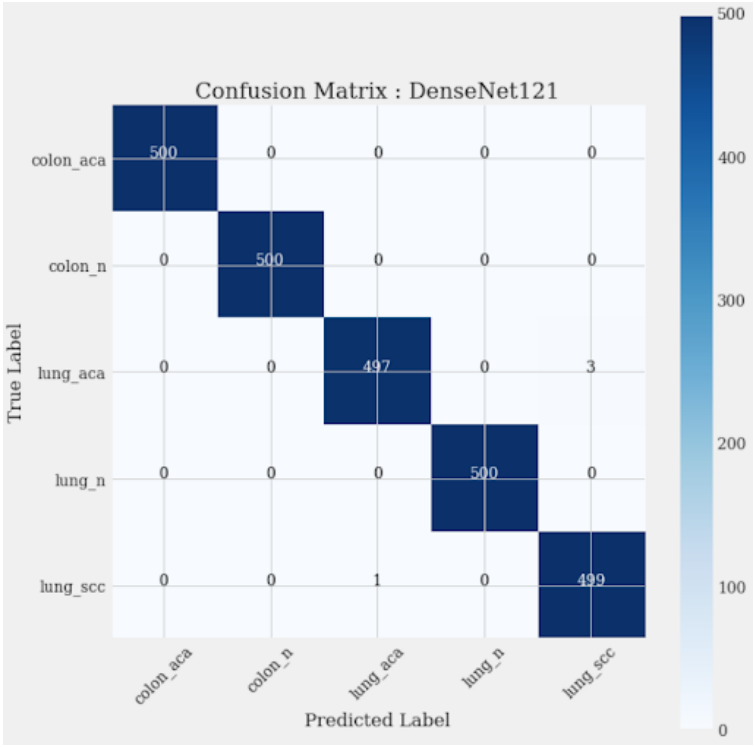

Figure 11. DenseNet based model Confusion Matrix

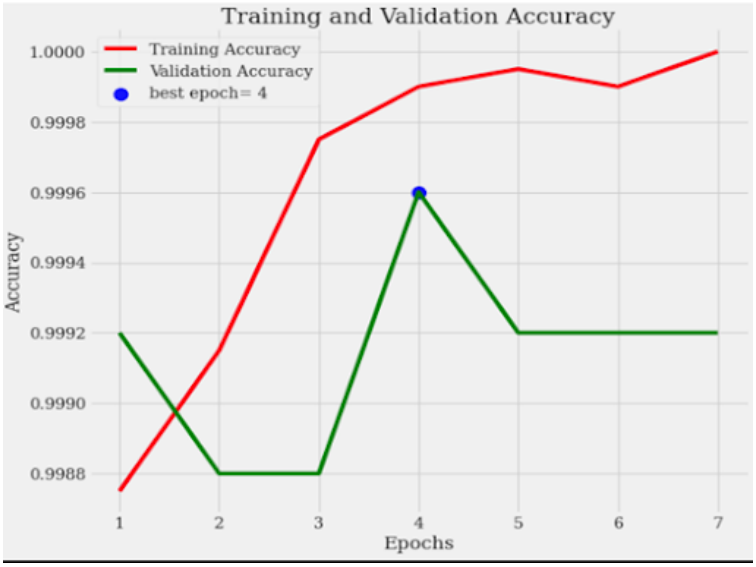

Figure 12. Ensemble learning model learning evolution (without enhancement)

The confusion matrix (Figure 15) shows that the preprocessing phase has reduced misclassification rates for the lung_aca class, which is classified as lung_scc for only one image.

The findings of this study highlight that the preprocessing stage—integrating gamma correction with K-means clustering—plays a significant role in enhancing the accuracy of histopathological image classification. Reducing

Table 2. *Precision, Recall and F1-Score before enhancement*

Class	Precision	Recall	F1-score
colon_aca	1.0000	1.0000	1.0000
colon_n	1.0000	1.0000	1.0000
lung_aca	1.0000	0.9940	0.9970
lung_n	1.0000	1.0000	1.0000
lung_scc	0.9940	1.0000	0.9970

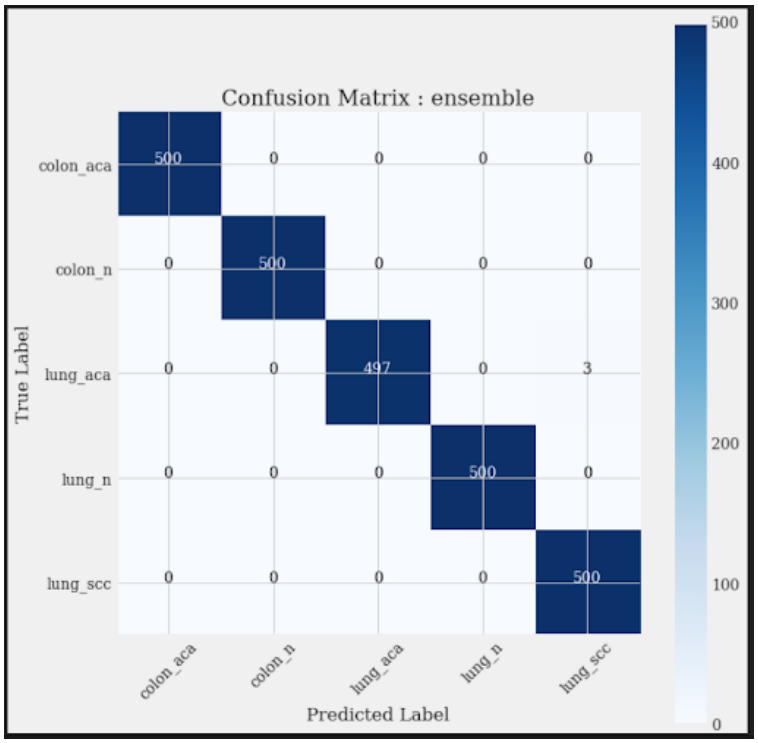

Figure 13. Confusion Matrix on test dataset (without enhancement)

Table 3. *Precision, Recall and F1-Score after enhancement*

Class	Precision	Recall	F1-score
colon_aca	1.000	1.000	1.000
colon_n	1.000	1.000	1.000
lung_aca	1.000	0.998	0.999
lung_n	1.000	1.000	1.000
lung_scc	0.998	1.000	0.999

the number of colors a classification model has to learn, by using K-means clustering and adjusting image contrast with a gamma correction of 1.2. This experimental choice reduced the number of colors from over a million to just nine. This enhances contrast and contours, as shown by measured increases in RMS contrast and edge intensity. This simplification of the images did not detract from the quality of information; on the contrary, it contributed to a

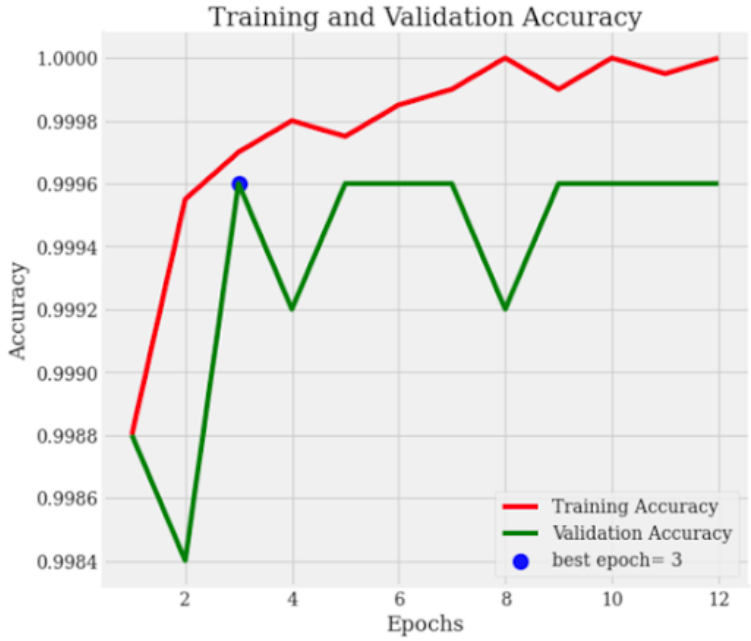

Figure 14. Ensemble learning model learning evolution (with enhancement)

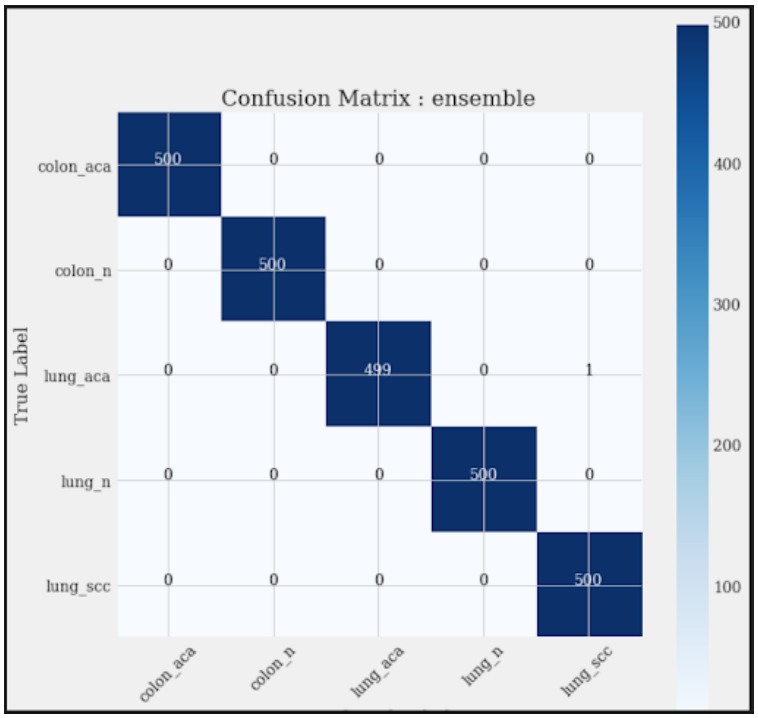

Figure 15. Confusion Matrix on test dataset (with enhancement)

better distinction of internal structures, which is crucial for the identification of cancer cells. That has enabled the VGG16 and DenseNet models to achieve high accuracy levels of 99.68% and 99.84% respectively. This performance exceeds the results obtained without preprocessing, where accuracies are only 99.44% and 99.64%.

Combining our preprocessing approach with the ensemble learning models further improved performance, achieving 99.96% accuracy for the DenseNet model and 99.88% for VGG16. These gains can be attributed to an effective simplification of the visual features of the images, making the boundaries between classes more distinct for the classification algorithms.

The integration of our VGG16 and DenseNet base models into model base using the ensemble learning approach enabled us to achieve a maximum accuracy of 99.96%, surpassing several methods previously reported in the literature. This performance is attributable to the complementary nature of the models used and the quality of the pre-processed images, which enabled better generalization on the test and validation sets. The analysis of the confusion matrices also revealed a reduction in the number of errors, particularly in the classification of images in the lung_aca class, often confused with lung_scc without preprocessing. After improvement, these confusions became marginal.

Finally, comparisons with previous work (Table 4) confirm that our approach is among the most effective for classifying images from the LC25000 dataset, offering the advantages of being reproducible, simple to implement, and improving both model performance and image visual quality.

Table 4. Comparison of classification accuracy with existing approaches

References	**Model/Approach**	**Accuracy**
Merabet et al. [32]	CNN + Inception-V3, ResNet50 + Inception-V3	99.27%, 99.20%
Gowthamy et al. [33]	ResNet-50 + InceptionV3 + DenseNet + KELM	98.90%
Omar et al. [34]	Ensemble (MobileNet V1, Inception V3, VGG16)	99.44%
Al-Ofary and Ilhan [35]	DenseNet201, ResNet101, EfficientNet-B0 (ensemble)	99.88%
Abd El-Aziz et al. [36]	ResNet-101V2 + NASNetMobile + EfficientNet-B0 (fusion-based model)	99.94%
S. Savaş and O. Güler [37]	DETL_V1, combines DenseNet121, InceptionV3, MobileNet, ResNet50, ResNet101, VGG16, and Xception	99.78%
O. Singh and K. K. Singh [38]	Random Forest (RF) + SVM + Logistic Regression (LR)	99.00%
Sethy et al. [39]	AlexNet, wavelet transform (DWT), and SVMs	99.30%
Ochoa-Ornelas et al. [40]	XGBoost + LightGBM + CatBoost	94.80%
Singh et al. [41]	EfficientNetB3 + InceptionNetV2 + ResNet50 + VGG16	99.00%
Our Model	Without Enhancement	99.88%
	With Enhancement	**99.96%**

5. Conclusion

Lung cancer remains a high-risk, high-mortality disease. Automating the analysis and diagnosis process is the most appropriate solution for early detection and increasing the chances of cure. In this paper, we have attempted to propose a new model based on AI technologies for the classification of histopathological images of these types of cancer. Our approach is based on image enhancement using unsupervised learning through the K-means model and the gamma correction function. To classify images after enhancement, we first used the transfer learning approach by developing two models based on VGG16 and DenseNet. These models were trained on the L25000 dataset before and after image enhancement. The accuracies of 99.44% and 99.64% were respectively achieved,respectively, by these models before enhancement and 99.68% and 99.84% after enhancement. The best models achieved were subsequently combined in an ensemble learning model, which achieved an accuracy of 99.96%. Our results demonstrate that unsupervised ML-based image enhancement significantly improves classification performance and highlight the importance of transfer learning and ensemble learning in achieving high accuracy. Our approach can be further improved by using other datasets, testing additional transfer learning models, or integrating transformers, which have shown promising results in computer vision tasks.

REFERENCES

1. (2024) Cancer today. International Agency for Research on Cancer (IARC). Accessed: Oct. 25, 2024. [Online]. Available: https://gco.iarc.who.int/today/
2. J. S. Brown, S. R. Amend, R. H. Austin, R. A. Gatenby, E. U. Hammarlund, and K. J. Pienta, "Updating the definition of cancer," *Molecular Cancer Research*, vol. 21, no. 11, pp. 1142–1147, 2023.
3. E. S. Cheng, M. Weber, J. Steinberg, and X. Q. Yu, "Lung cancer risk in never-smokers: An overview of environmental and genetic factors," *Chinese Journal of Cancer Research*, vol. 33, no. 5, p. 548, 2021.
4. S. Foser, K. Maiese, S. R. Digumarthy, J. A. Puig-Butille, and C. Rebhan, "Looking to the future of early detection in cancer: liquid biopsies, imaging, and artificial intelligence," *Clinical Chemistry*, vol. 70, no. 1, pp. 27–32, 2024.
5. R. Tandon, S. Agrawal, N. P. S. Rathore, A. K. Mishra, and S. K. Jain, "A systematic review on deep learning-based automated cancer diagnosis models," *Journal of Cellular and Molecular Medicine*, vol. 28, no. 6, p. e18144, 2024.
6. N. Ghaffar Nia, E. Kaplanoglu, and A. Nasab, "Evaluation of artificial intelligence techniques in disease diagnosis and prediction," *Discover Artificial Intelligence*, vol. 3, no. 1, p. 5, 2023.
7. M. Swapna, Y. K. Sharma, and B. Prasadh, "Cnn architectures: Alex net, le net, vgg, google net, res net," *Int. J. Recent Technol. Eng*, vol. 8, no. 6, pp. 953–960, 2020.
8. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211–252, 2015.
9. A. Oubaalla, H. El Moubtahij, and N. El Akkad, "Medical image segmentation using deep learning: A survey," in *International Conference on Digital Technologies and Applications*. Springer, 2023, pp. 974–983.
10. H. Fouadi, H. El Moubtahij, H. Lamtougui, and A. Yahyaouy, "Sentiment analysis of arabic comments using machine learning and deep learning model," *Indian Journal of Computer Science and Engineering (IJCSE)*, vol. 13, no. 3, pp. 598–606, 2022.
11. R. Javed, T. Abbas, A. H. Khan, A. Daud, A. Bukhari, and R. Alharbey, "Deep learning for lungs cancer detection: a review," *Artificial Intelligence Review*, vol. 57, no. 8, p. 197, 2024.
12. A. Lagree, M. Mohebpour, N. Meti, K. Saednia, F.-I. Lu, E. Slodkowska, S. Gandhi, E. Rakovitch, A. Shenfield, A. Sadeghi-Naini *et al.*, "A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks," *Scientific Reports*, vol. 11, no. 1, p. 8025, 2021.
13. A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, "Lung and colon cancer histopathological image dataset (lc25000)," *arXiv preprint arXiv:1912.12142*, 2019.
14. S. Hadiyoso, S. Aulia, I. D. Irawati *et al.*, "Diagnosis of lung and colon cancer based on clinical pathology images using convolutional neural network and clahe framework," *Int. J. Appl. Sci. Eng*, vol. 20, no. 1, pp. 1–7, 2023.
15. M. Al-Jabbar, M. Alshahrani, E. M. Senan, and I. A. Ahmed, "Histopathological analysis for detecting lung and colon cancer malignancies using hybrid systems with fused features," *Bioengineering*, vol. 10, no. 3, p. 383, 2023.
16. J. Van der Laak, G. Litjens, and F. Ciompi, "Deep learning in histopathology: the path to the clinic," *Nature medicine*, vol. 27, no. 5, pp. 775–784, 2021.
17. A. A. Ahmed, M. Abouzid, and E. Kaczmarek, "Deep learning approaches in histopathology," *Cancers*, vol. 14, no. 21, p. 5264, 2022.
18. P. Patharia, P. K. Sethy, and A. Nanthaamornphong, "Advancements and challenges in the image-based diagnosis of lung and colon cancer: A comprehensive review," *Cancer Informatics*, vol. 23, p. 11769351241290608, 2024.
19. A. Echle, N. T. Rindtorff, T. J. Brinker, T. Luedde, A. T. Pearson, and J. N. Kather, "Deep learning in cancer pathology: a new generation of clinical biomarkers," *British journal of cancer*, vol. 124, no. 4, pp. 686–696, 2021.
20. M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. AlZain, "A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework," *Sensors*, vol. 21, no. 3, p. 748, 2021.
21. M. Toğaçar, "Disease type detection in lung and colon cancer images using the complement approach of inefficient sets," *Computers in Biology and Medicine*, vol. 137, p. 104827, 2021.
22. A. Hage Chehade, N. Abdallah, J.-M. Marion, M. Oueidat, and P. Chauvet, "Lung and colon cancer classification using medical imaging: A feature engineering approach," *Physical and Engineering Sciences in Medicine*, vol. 45, no. 3, pp. 729–746, 2022.
23. M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, K. F. Hasan, and M. A. Moni, "Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning," *Expert Systems with Applications*, vol. 205, p. 117695, 2022.
24. E. A.-R. Hamed, M. A.-M. Salem, N. L. Badr, and M. F. Tolba, "An efficient combination of convolutional neural network and lightgbm algorithm for lung cancer histopathology classification," *Diagnostics*, vol. 13, no. 15, p. 2469, 2023.
25. A. Farhadipour, "Lung and colon cancer detection with convolutional neural networks and adaptive histogram equalization," *Iran Journal of Computer Science*, vol. 7, no. 2, pp. 381–395, 2024.
26. V. Bishnoi and N. Goel, "A color-based deep-learning approach for tissue slide lung cancer classification," *Biomedical Signal Processing and Control*, vol. 86, p. 105151, 2023.
27. O. Attallah, "Lung and colon cancer classification using multiscale deep features integration of compact convolutional neural networks and feature selection," *Technologies*, vol. 13, no. 2, p. 54, 2025.
28. S.-O. Shim, M. H. Alkinani, L. Hussain, and W. Aziz, "Feature ranking importance from multimodal radiomic texture features using machine learning paradigm: A biomarker to predict the lung cancer," *Big Data Research*, vol. 29, p. 100331, 2022.
29. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
30. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
31. M. Herrmann, F. J. D. Lange, K. Eggensperger, G. Casalicchio, M. Wever, M. Feurer, D. Rügamer, E. Hüllermeier, A.-L. Boulesteix, and B. Bischl, "Position: Why we must rethink empirical research in machine learning," *arXiv preprint arXiv:2405.02200*, 2024.
32. A. Merabet, A. Saighi, M. A. Ferradji, and Z. Laboudi, "Enhancing colon cancer prediction in histopathology with integrated deep learning models: A comparative study on the lc25000 dataset," in *2024 6th International Conference on Pattern Analysis and Intelligent Systems (PAIS)*. IEEE, 2024, pp. 1–7.

33. J. Gowthamy and S. Ramesh, "A novel hybrid model for lung and colon cancer detection using pre-trained deep learning and kelm," *Expert Systems with Applications*, vol. 252, p. 124114, 2024.
34. L. T. Omar, J. M. Hussein, L. F. Omer, A. M. Qadir, and M. I. Ghareb, "Lung and colon cancer detection using weighted average ensemble transfer learning," in *2023 11th international symposium on digital forensics and security (ISDFS)*. IEEE, 2023, pp. 1–7.
35. S. Al-Ofary and H. O. Ilhan, "Decision level fusion of transfer learning based models for diagnosing lung and colon cancer," in *2023 Intelligent Methods, Systems, and Applications (IMSA)*. IEEE, 2023, pp. 187–192.
36. A. Abd El-Aziz, M. A. Mahmood, and S. Abd El-Ghany, "Advanced deep learning fusion model for early multi-classification of lung and colon cancer using histopathological images," *Diagnostics*, vol. 14, no. 20, p. 2274, 2024.
37. S. Savaş and O. Güler, "Ensemble learning based lung and colon cancer classification with pre-trained deep neural networks," *Health and Technology*, vol. 15, no. 1, pp. 105–117, 2025.
38. O. Singh and K. K. Singh, "An approach to classify lung and colon cancer of histopathology images using deep feature extraction and an ensemble method," *International journal of information technology*, vol. 15, no. 8, pp. 4149–4160, 2023.
39. P. K. Sethy, A. Geetha Devi, B. Padhan, S. K. Behera, S. Sreedhar, and K. Das, "Lung cancer histopathological image classification using wavelets and alexnet," *Journal of X-Ray Science and Technology*, vol. 31, no. 1, pp. 211–221, 2023.
40. R. Ochoa-Ornelas, A. Gudiño-Ochoa, and J. A. García-Rodríguez, "A hybrid deep learning and machine learning approach with mobile-efficientnet and grey wolf optimizer for lung and colon cancer histopathology classification," *Cancers*, vol. 16, no. 22, p. 3791, 2024.
41. O. Singh, K. K. Singh, S. Das, A. S. Akbari, and N. Abd Manap, "Classification of lung cancer from histopathology images using a deep ensemble classifier," in *2023 IEEE International Conference on Imaging Systems and Techniques (IST)*. IEEE, 2023, pp. 1–6.