

Histopathological image recognition of colon cancer

Marco Frattarola

2024-06-04

Contents

1	Introduction	2
2	Dataset	2
3	Image analysis	3
3.1	Color histogram	3
3.2	Color Moments	5
4	Texture feature extraction	5
4.1	Image preprocessing	5
4.2	GLCM	6
4.3	Classification	6
5	Detection of elliptical structures	7
5.1	Detection algorithm	8
5.2	Viewing examples	12
5.3	Alternative approach	12
5.4	Classification	12
5.5	Observations	13
6	Feature extraction with CNN	14
6.1	Feature extractor	15
6.2	Classification	16
7	Classification with CNN	17
7.1	Model architecture	17
7.2	Data augmentation	18
7.3	Training	18
7.4	Classification	19
8	Comparison between approaches used	20
9	Conclusions	21

1 Introduction

The objective of the thesis is to exploit and use artificial vision techniques for the recognition and classification of images of colon tissues, distinguishing between benign tissues and those affected by adenocarcinoma. Colon adenocarcinoma is one of the most common forms of colorectal cancer. The current cancer detection method is quite time-consuming and expensive. Pathologists must learn to recognize tissues by studying labeled histopathological images. As a result, a significant amount of resources and manual labor is wasted. Therefore, automating this analysis through computer vision techniques represents a significant opportunity to improve diagnostic accuracy and efficiency.

The thesis is divided into two main approaches:

- **Traditional Computer Vision Techniques:** These techniques include the use of classical image processing methods for feature extraction. These features can then be used to train classifiers such as Support Vector Machine (SVM) or Random Forest, in order to distinguish between benign and malignant tissues.
- **Neural Networks:** Convolutional neural networks (CNN) represent the state of the art in image recognition and are particularly effective in classifying medical images. Neural networks can automatically learn the relevant features of images, without the need to extract them manually.

The main objective of the paper is to compare the effectiveness and accuracy of the two approaches, evaluating their performance in terms of precision, accuracy and processing time.

2 Dataset

A dataset available on the Kaggle platform was used to develop the thesis. The dataset in question, accessible at the URL [Lung and Colon Cancer Histopathological Images \[AAB\]](#), contains 25,000 histopathological images belonging to 5 different classes. All images have a size of 768 x 768 pixels and are in *jpeg* format. The images were generated from a sample of validated and Health Insurance Portability and Accountability Act (HIPAA) compliant sources. The dataset is composed of:

- 750 total images of **lung tissue**:
 - 250 benign lungs
 - 250 lung adenocarcinomas
 - 250 squamous cell carcinomas of the lung
- 500 total images of **colon tissues**
 - 250 benign colon
 - 250 colon adenocarcinomas

The dataset was subsequently expanded by a factor **20** using *data augmentation* techniques ([Augmentor](#)), thus obtaining 25,000 images. For each histopathological image, 20 different images were generated.

The Dataset contains images belonging to 5 different classes each with 5,000 images:

- Benign lung tissue
- Lung adenocarcinoma
- Squamous cell lung cancer
- Adenocarcinoma of the colon
- Benign colon tissue

The objective of the paper is to classify images relating to colon tissues, consequently images relating to lung tissues were ignored. The dataset used, therefore, considering only the images relating to the colon tissues, contains 10,000 images belonging to two classes:

- Adenocarcinoma of the colon
- Benign colon tissue

3 Image analysis

3.1 Color histogram

A **color histogram** is a representation of the distribution of colors in an image. The histogram is obtained:

- by quantizing each color space channel into a number bin prefix
- counting how many pixels in the image have a certain color.

After calculating, for each of the two classes, the average color histogram with respect to all the images in the dataset, it can be noted that there is no substantial difference between the two. However, a higher frequency can be observed in the blue channel of the images relating to benign tissues, compared to non-benign ones where lower frequency values are observed. In figure 1, you can compare the color histograms relating to the two classes, calculated using 200 bins. In figure 2, however, you can compare the color histograms relating to the two classes, calculated using 50 bins.

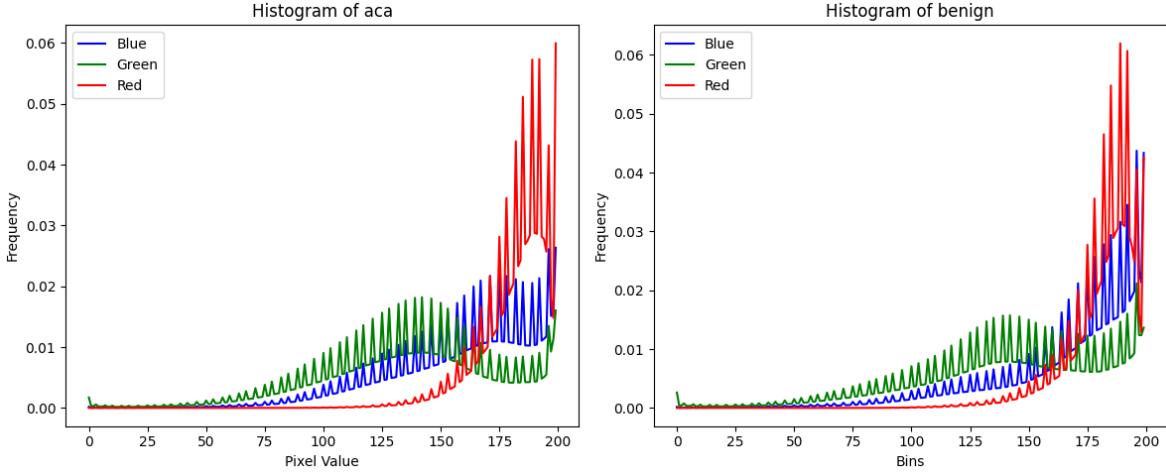


Figure 1: comparison between color histograms with 200 bins calculated on 500 images of the two classes.

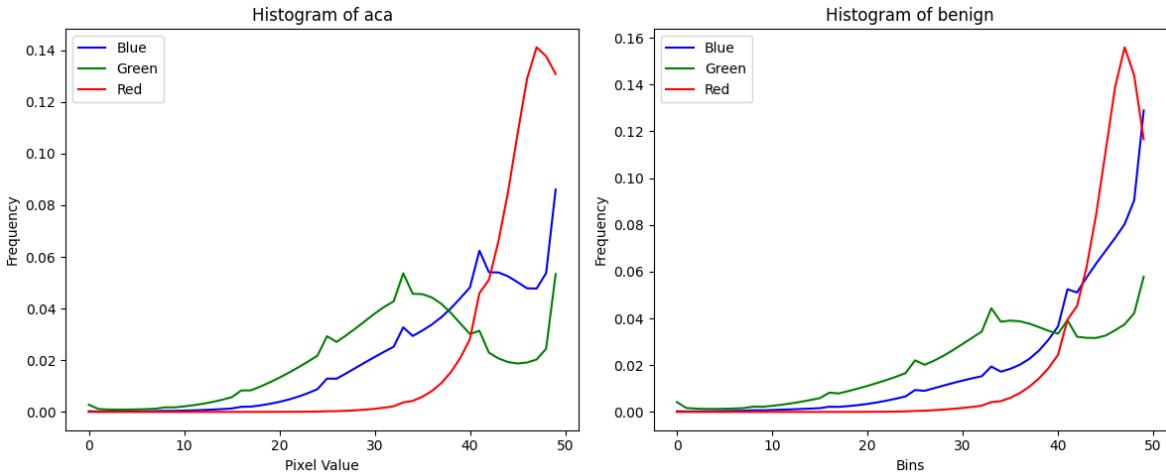


Figure 2: Comparison between color histograms with 50 bins calculated on 500 images of the two classes.

In order to compare images based on color histograms Two metrics were used:

- **Correlation:** In the context of histograms, quantifies the similarity between two histograms. Values close to 1 indicate high similarity. Values close to 0 indicate little or no similarity. It is calculated as:

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{\sum_I (H_1(I) - \bar{H}_1)^2 \sum_I (H_2(I) - \bar{H}_2)^2}} \quad (1)$$

- **Similarity:** measures the common area of the histograms. It quantifies the degree of overlap between two histograms. It is calculated as:

$$d(H_1, H_2) = \sum_I \min(H_1(I), H_2(I)) \quad (2)$$

In the 1 table you can see that the histograms of the images of the two classes have very high values in all channels, both in the correlation and in the intersection. The presence of such high values (close to 1) indicates that histograms of images of different classes are very similar to each other, making their classification difficult.

Channel	Correlation	Intersection
Red	0.9941	0.9640
Blue	0.9179	0.8263
Green	0.9182	0.8828

Table 1: Correlation and intersection values calculated between two histograms relating to the two classes, calculated on 500 images per class and 200 bins

3.2 Color Moments

color moments are widely used color descriptors for their simplicity and effectiveness. These descriptors capture information about the distribution of colors in an image. They allow describing the spatial distribution of colors and are invariant with respect to geometric and scale transformations of the image, which makes them robust in different contexts. They are based on the assumption that the distribution of colors in an image can be interpreted as a probability distribution, characterized by a certain number of moments that can be used to describe the image itself. 3 color moments were used, calculated globally across the entire image:

- **Average:** Average color in the image.

Calculated as: $\frac{1}{N} \sum_{i=1}^N x_i$

- **Standard deviation:** square root of the variance of the color distribution.

Calculated as: $\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \text{mean})^2}$

- **Skewness:** measures the skewness of the color distribution.

Calculated as: $\sqrt[3]{\frac{1}{N} \sum_{i=1}^N (x_i - \text{mean})^3}$

Mean, standard deviation and skewness were calculated on all images. Each color moment was calculated globally across the entire image and subsequently averaged over the color moments of images of the same class. The color moments values of each class were then compared using the Euclidean distance.

You can view the results obtained in the table 2.

Channel	Mean	Std	Skewness
Red	10.6891	1.9197	9.2136
Blue	0.4539	0.4518	2.2133
Green	7.6266	5.4526	15.0697

Table 2: Comparison between color moments of the two classes

It can be seen that the distances between the 3 different moments analyzed are rather small. The presence of such small distance values indicates that the color distribution is **similar**, between images of different classes. Larger values would have indicated, however, a higher dissimilarity.

4 Texture feature extraction

There are several techniques for extracting texture features. These techniques allow you to quantify and describe the intensity variations and patterns present in images, facilitating the distinction between benign and malignant tissues. In this paper, the extraction of texture features was carried out using the Gray Level Co-Occurrence Matrix (GLCM), which is a robust and widely used technique for texture analysis.

4.1 Image preprocessing

Image **preprocessing** plays a vital role in image analysis and processing. The objective is to reduce the noise present in the image, thus improving the quality of the images for the purposes of feature extraction. The *preprocessing* used is divided into the following steps:

1. **Conversion to Grayscale:** Initially, the image is converted to grayscale, facilitating subsequent feature extraction.
2. **Resizing:** The image is resized to a different size depending on the approach used.
3. **Histogram Equalization:** Histogram equalization is applied using the CLAHE (Contrast Limited Adaptive Histogram Equalization) algorithm, which improves the local contrast of the image. This process is essential to emphasize the characteristics of the tissues, improving the visibility of the darker areas that delimit the elliptical structures.

4. **Median Noise Filtering:** To reduce image noise without compromising important details, a median filter is used.
5. **Non-Local Denoising:** Finally, the image is subjected to a non-local denoising process to further improve the image quality.

In the figure 3 you can view the steps that are carried out during the preprocessing phase.

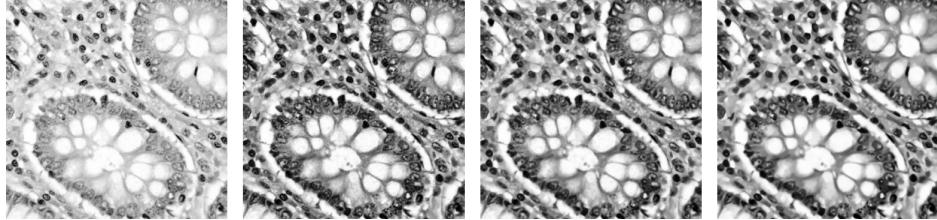


Figure 3: Steps performed during the preprocessing phase.

4.2 GLCM

GLCMs are two-dimensional histograms that they describe the co-occurrences of pairs of gray levels according to some specific directions. For each image, the GLCM was calculated, after preprocessing, considering 1, 2, 3 pixels as distances and as directions:

- 0°
- 45°
- 90°
- 135°

This matrix captures the spatial distribution of pixel intensities, reflecting the texture structure. Various features that describe the texture of the image were extracted from the GLCM:

- **Energy:** It represents the sum of the squares of the GLCM values, indicating the presence of repetitive patterns in the image, calculated as: $\sqrt{\sum_{i,j} M[i,j]^2}$
- **Entropy:** Measures the complexity of the weaving; higher values indicate a more complex and disordered texture, calculated as: $-\sum_{i,j} M[i,j] \cdot \log(M[i,j])$
- **Contrast:** Measures the difference in intensity between a pixel and its neighbors, reflecting the local variation in intensity, calculated as: $\sum_{i,j} (i-j)^2 \cdot M[i,j]$
- **Homogeneity:** Evaluate the uniformity of the texture; higher values indicate a more uniform texture, calculated as: $\sum_{i,j} \frac{M[i,j]}{1+|i-j|}$
- **Dissimilarity:** Measures the absolute difference in intensity between pixels; higher values indicate a greater variation in intensity, calculated as: $\sum_{i,j} |i-j| \cdot M[i,j]$

4.3 Classification

The 5 features described previously were extracted for each image in the dataset, thus obtaining a new feature dataset of size $10,000 \times 5$. This dataset was divided into train and test sets, and then used to train a Support Vector Machine (SVM) classifier. Performance was measured on the test set of 2000 images. In the table 3 you can observe the results obtained. The classifier manages to obtain an accuracy of 87.9% on the test set, thus failing to discriminate in an excellent manner between images of benign tissues or those affected by adenocarcinoma.

The confusion matrix is shown in figure 4. It can be seen that the classifier makes approximately double the errors in classifying non-benign images compared to benign ones. In fact, on the test set, 92% of the benign images were classified correctly, while only 84% of the non-benign ones were classified correctly.

Metric	Value
Accuracy	0.8790
Precision	0.9092
Precision	0.8370
F1 Score	0.8715
True Negatives	937
False Positives	82
False Negatives	160
True Positives	821

Table 3: Evaluation metrics of the SVM classifier trained on GLCM features, calculated on the test set.

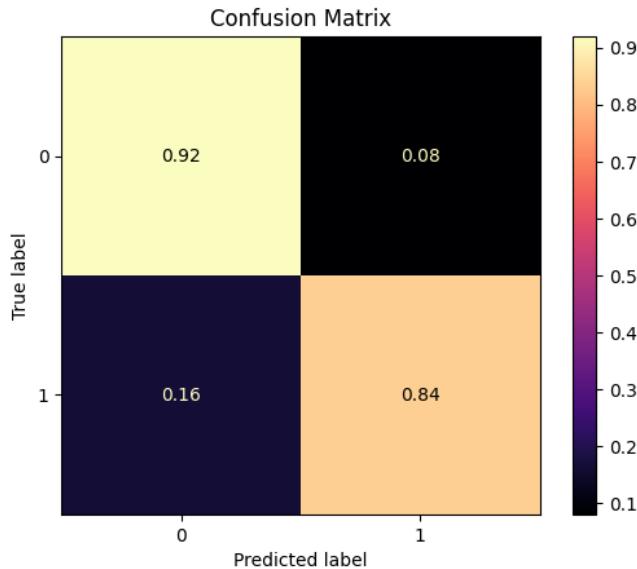


Figure 4: Confusion matrix of the values predicted by the SVM classifier trained on GLCM features computed over the entire image.

5 Detection of elliptical structures

In this section, we will describe the techniques that have been used with the aim of highlighting the **regions of interest** present in the image. In particular, the attempt was to highlight the **elliptical structures** present in the images of benign tissues. Observing the images of the dataset, one can in fact notice how images of benign tissues are characterized by the presence of elliptical structures with a rather delineated internal texture, unlike non-benign images, which present more discontinuous and irregular shapes. The objective is, therefore, to generate a binary mask which, once applied to the original image, preserves only the important areas of the image (elliptical shapes), obscuring (making black) the pixels belonging to regions not important than the image.

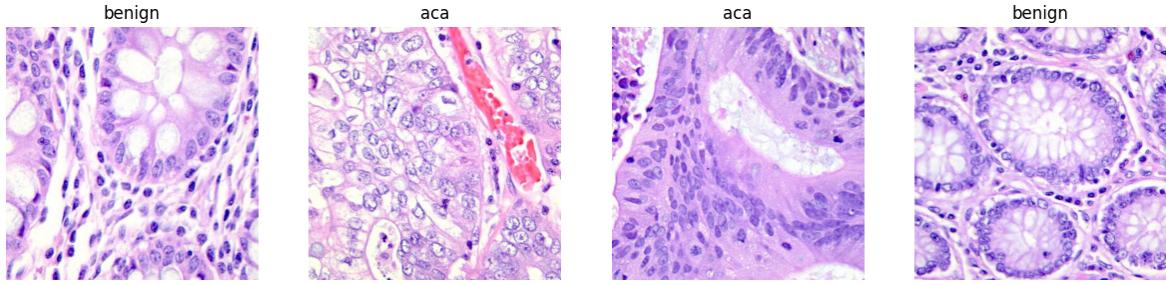
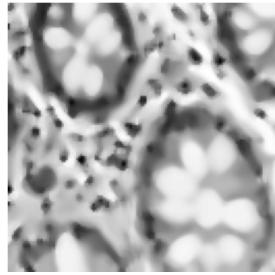


Figure 5: Example of benign (benign) and non-benign (aca) images.

5.1 Detection algorithm

The detection algorithm is divided into the following steps:

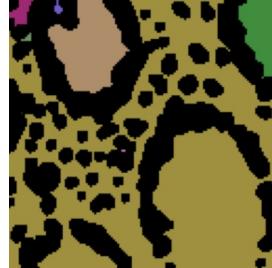
1. **Image preprocessing:** The preprocessing defined previously is performed. This time, however, the image is resized before preprocessing.



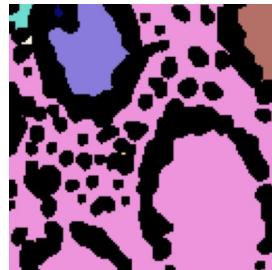
2. **Image binarization:** A binary threshold is applied to the preprocessed image using a threshold value. Pixels with values above the threshold are set to 255 (white), while those with values below are set to 0 (black).



3. **Morphological Erosion:** An initial morphological erosion is applied to the binarized image. The goal is to separate the inner part of the elliptical structures from their edges.



4. **Management of particular images:** Two parameters are used to recognize highly zoomed images and images where the edges are less evident: the percentage of pixels with values greater than 200 (light pixels) and the percentage of pixels with lower values to the threshold (dark pixels). If the percentage of light pixels is greater than or equal to 0.46 or the percentage of dark pixels is less than 0.16, an additional morphological opening operation is applied. Highly zoomed images, and those where the edges are very similar within the circular structures, typically require an additional opening operation.



5. Identification of elliptical structures:

- A morphological opening operation is performed first, followed by erosion. This way, the central parts of the fabrics, which are typically more "connected" than the rest of the image, should separate from the edges.

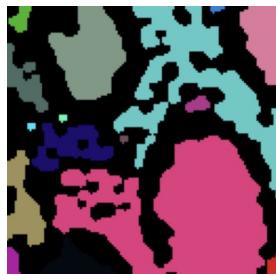
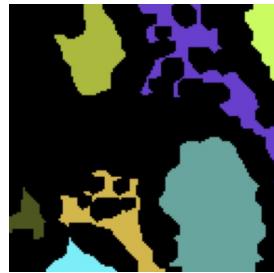


Figure 6: Opening



Figure 7: Erosion

- Small blobs of pixels are removed from the image. After the previous operations, less connected areas tend to separate, creating fragmented areas with small blobs of pixels. Since we are only interested in large blobs (inside the elliptical structures), all blobs smaller than the average 60% are removed.



- For each blob, the holes present inside it are closed, possibly also bordering the edges of the image. As a result of opening and erosion operations, holes may be created within the larger blobs. It may be convenient to close them to better highlight the shape of the blob.



- Pixel blobs that have a circularity below a certain threshold and therefore do not reflect a circular shape are removed. In particular, all blobs with circularity less than 0.2 were removed.



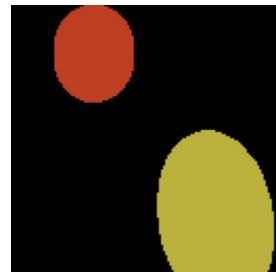
- Small blobs of pixels from the image are again removed. This time all blobs smaller than the average 80% are removed. The goal is to remove very small components that may not represent areas of interest that are normally approximately the same size.



- For each blob remaining in the image, an ellipse is drawn to reflect the shape of the blob.



- Finally, a morphological dilation is applied to restore the initial dimensions of the blob before erosion.



6. Mask Creation and Application:

- The binary mask, which was created previously, is applied to the original image. The bitwise AND operation is used to combine the original image with the mask, resulting in the final image.



5.2 Viewing examples

We visualize the intermediate steps of the algorithm applied on some images and subsequently its application on some images of benign tissues.

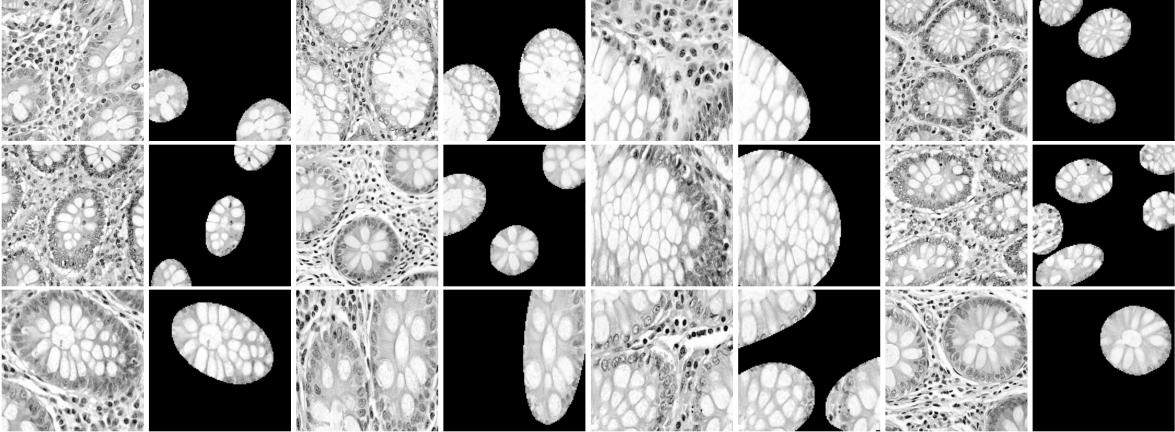


Figure 8: Images relating to benign tissues on which the detection algorithm has been applied.

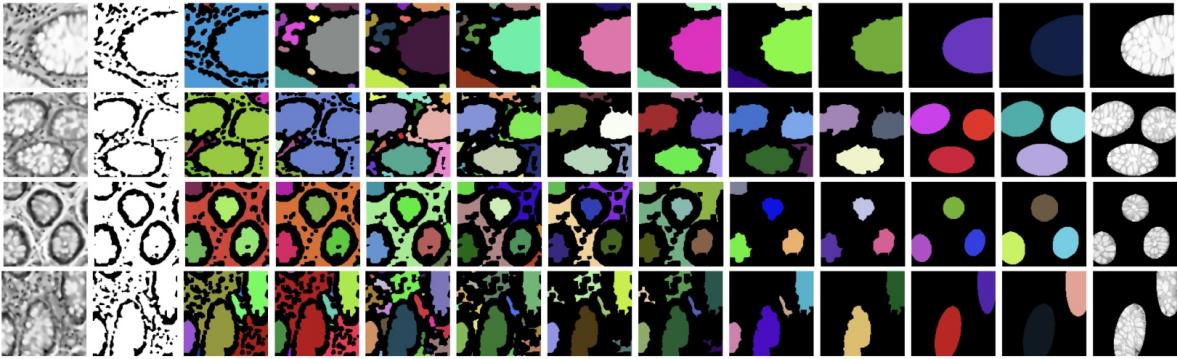


Figure 9: Step of the detection algorithm applied to images of benign tissues.

5.3 Alternative approach

Furthermore, a detection attempt was carried out using a different approach: unlike the previous algorithm, the idea was to highlight the **contours** of the elliptical shapes, rather than the internal part of them. As can be seen from the images, these contours are composed of darker pixels and surrounded by lighter areas of pixels that separate them from the rest of the tissue. The binarization of the image therefore occurs in the reverse manner to what was done previously, setting to 1 the pixels that have low intensity values (dark pixels). This approach, however, led to poor results, as it was rather difficult to distinguish the outlines of the elliptical shapes from the rest of the image.

5.4 Classification

An attempt at classification was carried out, using the images in which the areas of interest were highlighted, using the illustrated algorithm. For each image, therefore, after applying the binary mask that highlights the regions of interest, the previously defined GLCM features were extracted. Similar to what was done in the previous phase of GLCM feature extraction, a dataset containing the 5 features for each image was created. This dataset was subsequently used to train an SVM classifier. In the table 4 you can observe the results obtained on a test set of 2000 images.

Metric	Value
Accuracy	0.8290
Precision	0.8160
Precision	0.8410
F1 Score	0.8283
True Negatives	833
False Positives	186
False Negatives	156
True Positives	825

Table 4: Classifier evaluation metrics, calculated on the test set.

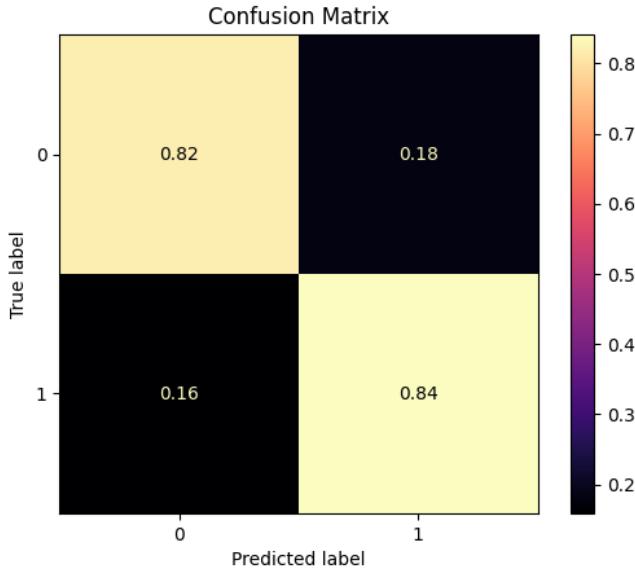


Figure 10: confusion matrix of the values predicted by the SVM classifier trained on GLCM features calculated over the entire image.

We can see how the results obtained are **worse**, compared to those obtained by classifying the images on the basis of the GLCM features, calculated on the entire image. Interestingly, the classifier makes many more errors on *benign* images than the previous approach. This behavior could be due to the fact that, in some images, the detection algorithm is unable to identify any region of interest, passing a completely black image to the classifier.

5.5 Observations

Although in images of benign tissues in which there are regular, circular elliptical structures the detection actually manages to highlight some areas of interest, this algorithm particularly suffers from variations in **scale**: in images with a strong *zoom*, the "circularity" of elliptical structures becomes less evident, making their identification more difficult. An attempt to mitigate this problem was made by trying to measure the percentage of very light pixels. If the percentage of light pixels is very high, it could be an image with a strong zoom, given that the internal area of the circular structures is characterized by light pixels. The same problem occurs with poorly zoomed images, where the circular structures are smaller and it is therefore more difficult to separate them from the rest of the image through morphological operations. In the figure 11 you can observe a very zoomed image in which the algorithm was not able to identify any area of interest.

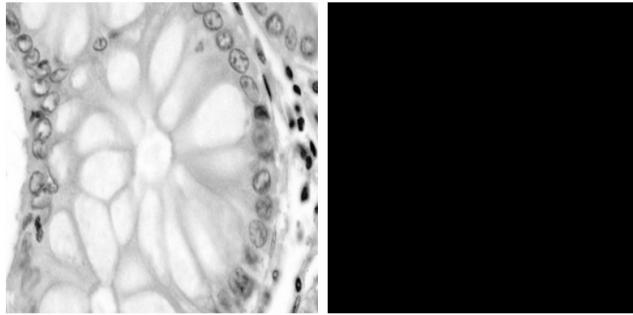


Figure 11: Example of an image in which the detection algorithm was unable to identify any region of interest.

Another problem with this algorithm is the fact that, even in images of non-benign tissues, elliptical regions of interest are highlighted where in reality there is no elliptical structure of those that characterize images of benign tissues. However, the regions of interest that are highlighted in images of non-benign tissues have a rather different texture compared to the regions of interest that are identified in benign tissues. The hope was that a classifier would then be able to learn these differences and be able to distinguish more easily and accurately between images of the two classes, thanks precisely to this difference in texture between the highlighted regions. From the results obtained, however, it can be seen that the classifier trained on GLCM features extracted over the entire image is more accurate, especially in recognizing benign images.

6 Feature extraction with CNN

The objective is to use *Transfer Learning*, therefore to exploit a convolutional neural network, pre-trained on an image-recognition task, in order to extract the significant features of the images. Since this is an already trained network, we rely on the fact that the network has already learned which are the most important features to consider. The extracted features can then be used to train a more traditional classifier such as SVM or Random Forest.

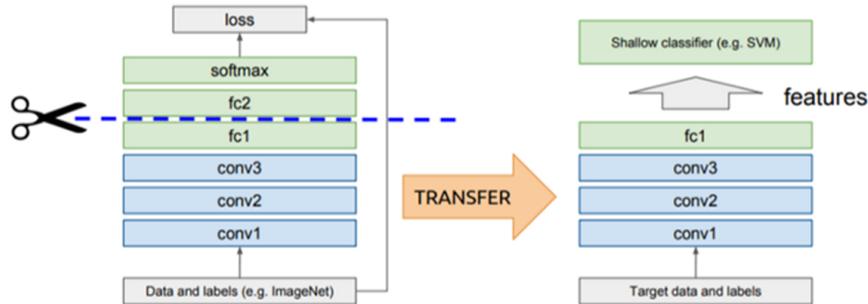


Figure 12: Diagram illustrating the transfer learning technique.

6.1 Feature extractor

In this work VGG16 [SZ15] was used as a feature extractor. VGG16 is a convolutional neural network pre-trained on the *ImageNet* dataset, which contains numerous images divided into 1000 classes. Since the goal is to use the network only as a feature extractor, the pre-trained VGG16 model was loaded, excluding the last fully-connected layers. In this way, the network extracts for each image a feature tensor of size $7 \times 7 \times 512$. The features extracted from VGG16 are very generic and can be used for various computer vision tasks, making the model versatile and robust.

For each image in the dataset, it is passed through the VGG16 network up to the last convolutional layer, thus obtaining a tensor of size $7 \times 7 \times 512$. This tensor represents the **features** extracted from the image.

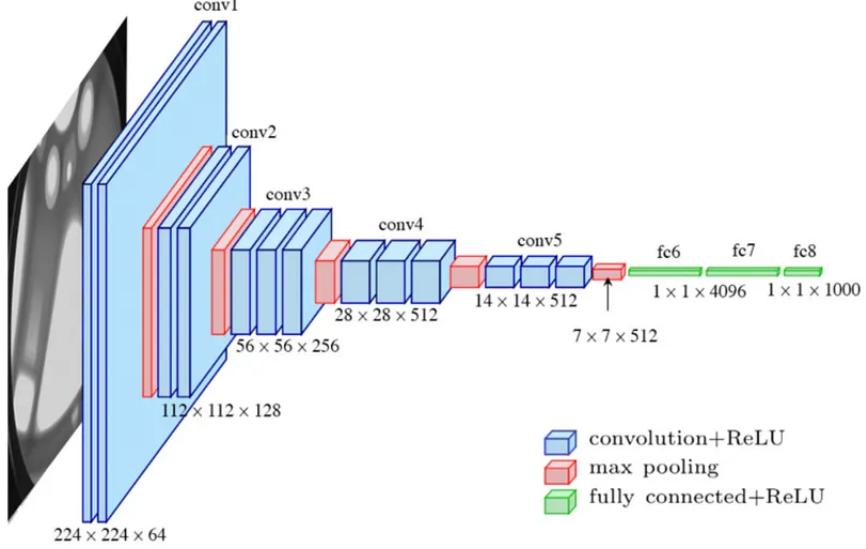


Figure 13: VGG16 architecture.

This feature tensor was then reduced in dimensionality using the **Principal Component Analysis** (PCA) technique. In fact, it would be inconvenient to train a classifier with a feature vector of dimensionality 25088.

6.2 Classification

After extracting features for each image in the dataset via the VGG16 network, dimensionality reduction was performed via *pca*, reducing the dataset to 10 and 20 dimensions. Therefore, two feature datasets of size $10,000 \times 10$ and $10,000 \times 20$ were generated. Each dataset was subsequently divided into train and test sets and then used to train a Support Vector Machine (SVM) classifier. As before, performance was measured on the test set of 2000 images.

Metric	10 features	20 features
Accuracy	0.957	0.924
Precision	0.9479	0.9550
Precision	0.9653	0.8869
F1 Score	0.9566	0.9197
True Negatives	967	978
False Positives	52	41
False Negatives	34	111
True Positives	947	870

Table 5: Evaluation metrics of the SVM classifiers, trained on features extracted using VGG16, calculated on the test set.

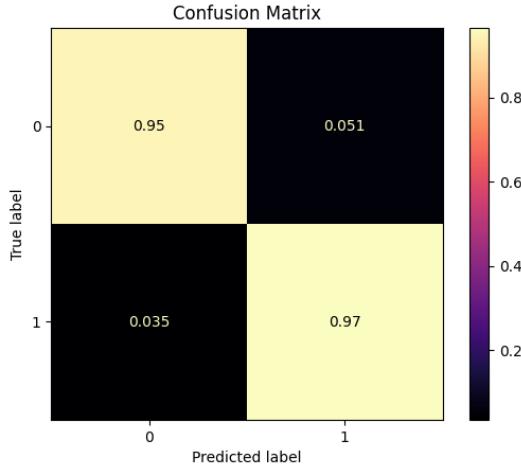


Figure 14: Confusion matrix of the values predicted by the SVM classifier trained on 10 features extracted using VGG16.

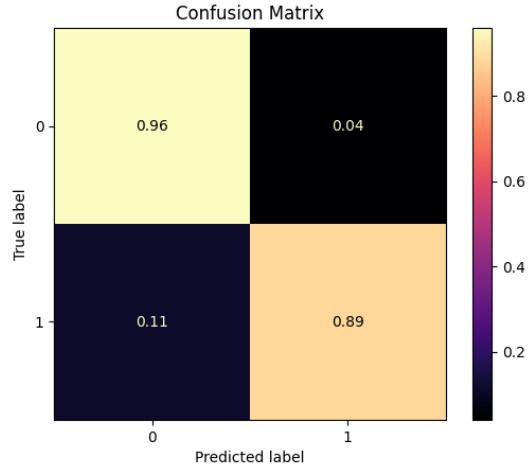


Figure 15: Confusion matrix of the values predicted by the SVM classifier trained on 20 features extracted using VGG16.

In the table 5 you can observe the results obtained. The classifier trained on 10 features manages to obtain an accuracy of 95.7% on the test set, thus obtaining decidedly **better** results compared to the classifier trained on GLCM features. The second classifier, trained on 20 features, also manages to obtain good results compared to the classifiers trained on GLCM, however it tends to **overfitting** on the training set. This leads to a worse generalization of the classifier which, as can also be seen in figure 14, leads to a worse accuracy on the test set.

7 Classification with CNN

The last approach used considers the use of a CNN as an image classifier and not only as a feature extractor. The architecture of the model used envisages, as previously, the use of VGG16 as a feature extractor, to which specific fully-connected layers are then added for the histopathological image classification task.

7.1 Model architecture

The basic model used is VGG16 which is used as a starting point. Its weights are **frozen** to avoid updating while training the new model. The following levels were then added to the basic model:

- **Global Average Pooling:** A *GlobalAveragePooling2D* layer is added to the base model. This layer averages over each filter 7×7 , reducing it to a single value. This reduces the spatial size of the features from $7 \times 7 \times 512$ to 1×512 .
- **Dense layer:** a fully-connected layer with 256 neurons and *ReLU* activation function.
- **Dropout 0.5:** a dropout layer with a rate of 0.5. This layer helps prevent *overfitting* during training by randomly turning off 50% of the neurons in the previous layer.
- **Dense layer:** an additional fully-connected layer with 128 neurons and *ReLU* activation function.
- **Output:** Finally, a final output layer with a number of neurons equal to the number of output classes. This layer uses the *softmax* activation function, which is ideal for multi-class classification tasks. In this case, only two output neurons were used, since the classification is performed on two classes.

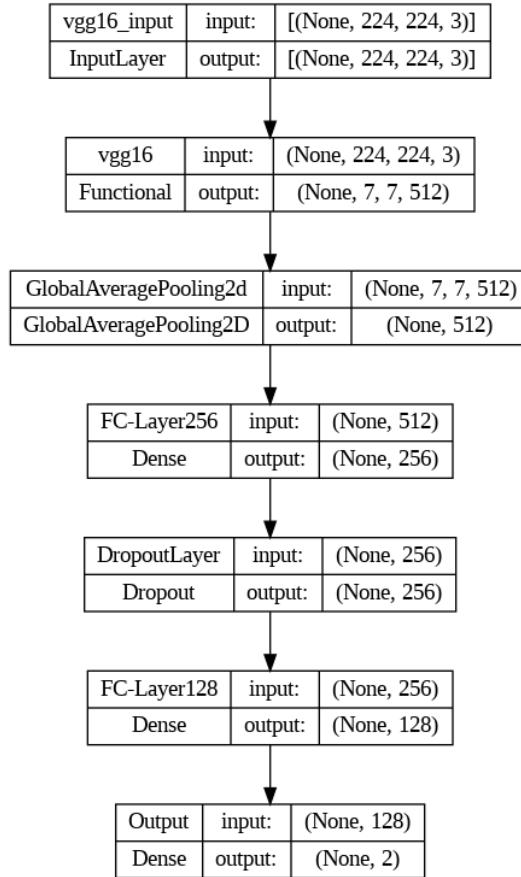


Figure 16: CNN model architecture.

7.2 Data augmentation

Data augmentation is a technique used to improve the performance of deep learning models. Artificially increasing the quantity and diversity of training data allows the network to generalize better and avoid the *overfitting* phenomenon.

In the context of the work, data augmentation applied to images introduces random variations such as rotations, shifts, distortions and zooms. These variations simulate the possible transformations that real images could undergo under different acquisition conditions. In this way, the CNN learns to recognize the characteristics of objects regardless of small variations in their position, orientation and scale.

In particular, the `ImageDataGenerator` class of the TensorFlow library was used, which allows you to define an image generator in *lazy* mode, avoiding loading all the images into memory.

This generator also allows us to apply the different augmentation techniques, described previously, to the images, so that at each **epoch** of training, the same image is seen by the network **in a different way**.

7.3 Training

The model was trained with the following configurations:

- **Loss:** **Binary cross-entropy** loss was used, suitable for binary classification problems.
- **Optimizer:** `Adam` was used as the optimizer
- **Weights:** The base model, VGG16, was used with weights already trained on ImageNet. The weights of the basic model were then frozen, to prevent their modification during the training phase of the last fully-connected layers.
- **Batch size:** 32 images per batch.
- **Epochs:** 15 epochs, using an *early stopping*.

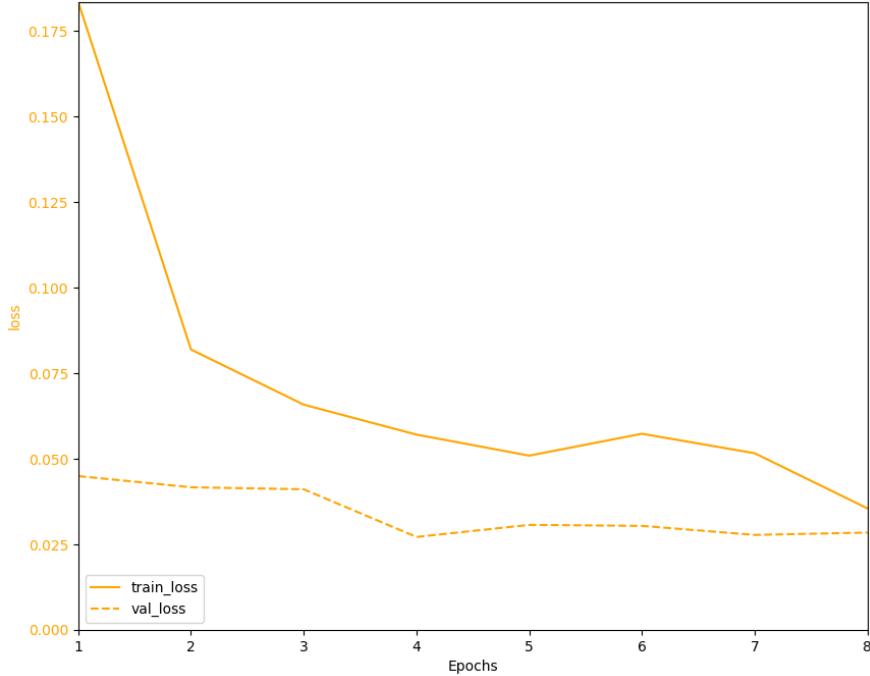


Figure 17: Loss trend during the training phase.

In figure 17 it is possible to observe the trend of the loss during the training phase. The model that was finally used is the one with the weights learned during the fourth training epoch, where the

val loss was lowest. Continuing with the epochs, in fact, the loss did not improve, interrupting the training after the eighth epoch.

7.4 Classification

As before, you can view the results obtained on the test set by using the trained model. We can see how the model is decidedly more accurate than those used previously, managing to correctly classify 98.6% of the images and significantly reducing the number of false positives and negatives.

Metric	Value
Accuracy	0.986
Precision	0.9799
Precision	0.9918
F1 Score	0.9858
True Negatives	999
False Positives	20
False Negatives	8
True Positives	973

Table 6: CNN evaluation metrics, computed on the test set.

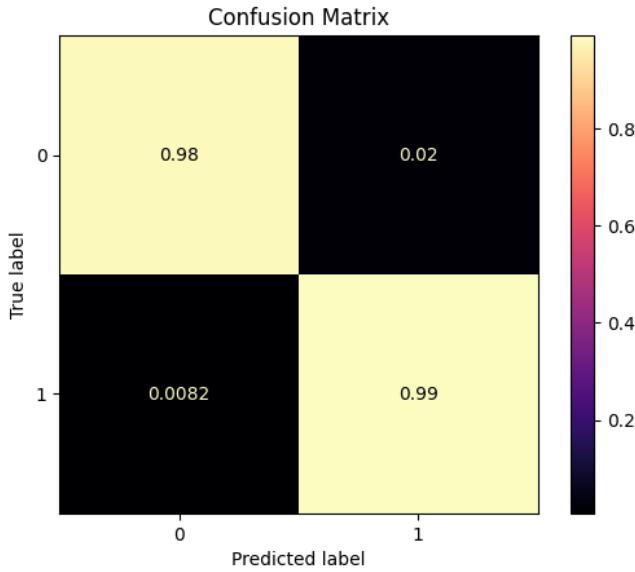


Figure 18: Confusion matrix of the values predicted by the CNN.

8 Comparison between approaches used

It is interesting to compare the different approaches used in order to analyze what the advantages and disadvantages of their use are. It is clear that the last approach used allows for significantly better results than the other approaches used. The execution time is reported in *ms* and calculated by taking the average of the classification times of 100 images. It can be seen that the use of the detection algorithm involves longer execution times compared to other approaches. Again, the last approach is the one that allows you to classify a new image more quickly. It is important to point out, however, that the training phase of the CNN in the last approach was rather long.

Metric	GLCM	GLCM + Detection	CNN-SVM 10 feature	CNN-SVM 20 feature	CNN
Accuracy	0.8790	0.8290	0.9570	0.9240	0.986
Precision	0.9092	0.8160	0.9479	0.9550	0.9799
Recall	0.8370	0.8410	0.9653	0.8869	0.9918
F1 Score	0.8715	0.8283	0.9566	0.9197	0.9858
True Negatives	937	833	967	978	999
False Positives	82	186	52	41	20
False Negatives	160	156	34	111	8
True Positives	821	825	947	870	973
Prediction time	135	799	131	276	108

Table 7: Comparison of evaluation metrics between all the approaches used.

9 Conclusions

In the paper, different approaches were explored for the recognition and classification of images of colon tissues, distinguishing between benign tissues and those affected by adenocarcinoma. The main approaches considered were traditional computer vision techniques and convolutional neural networks (CNN).

traditional computer vision techniques have involved the use of classical methods for feature extraction. These features were used to train classifiers such as Support Vector Machine (SVM). The obtained results demonstrated that these methods can be quite effective, but require careful feature selection to achieve good performance. Furthermore, the processing time was found to be relatively long, especially during the feature extraction phase.

Convolutional Neural Networks (CNNs) have shown superior results in histopathological image classification, thanks to their ability to automatically learn relevant features. CNNs achieved significantly higher accuracy than traditional methods. While the initial training time can be quite long, once trained, CNNs offer fast inference time.

The comparison between the two approaches has therefore highlighted that CNNs surpass traditional computer vision techniques in terms of accuracy and precision. However, traditional techniques can still be useful in situations where computational resources are limited or when you have a small dataset. Finally, the detection algorithm, although in a limited way, could help to identify areas of interest, which can then be used to classify the image.

In conclusion, the automation of colon tissue analysis using computer vision techniques represents a significant opportunity to improve diagnostic accuracy and efficiency. Convolutional neural networks, in particular, offer great potential for automatic classification of histopathology images, reducing the need for manual work by pathologists and speeding up the diagnostic process.

References

- [AAB] L. Brannon Thomas Catherine P. Wilson Lauren A. DeLand Stephen M. Mastorides Andrew A. Borkowski, Marilyn M. Bui. Lc25000 lung and colon histopathological image dataset.
- [SZ15] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.