Predicting Lung and Colon Cancer Using Computer Vision

Abstract

Lung and colon cancer are major public health concerns worldwide, and early identification is critical for better patient outcomes. The goal of this research is to create a computer vision algorithm that can predict the different types of lung and colon tissue images, such as lung adenocarcinoma, lung benign tissue, lung squamous cell carcinoma, colon benign tissue and colon adenocarcinoma. Image preprocessing, feature extraction with deep learning techniques, classification with CNN and its variants, and model training and optimization with backpropagation and gradient descent algorithms were used to accomplish this. To expand the sample size, the dataset utilized for this investigation was extended with the Augmentor program. The model's performance was assessed using accuracy, precision, and recall criteria. This project's possible applications include earlier detection and diagnosis of lung and colon cancer, increased accuracy and efficiency in cancer screening and diagnosis, and lower healthcare costs. Overall, this study shows a viable method for constructing a computer vision algorithm for medical imaging analysis, which has the potential to have a significant influence on cancer detection and therapy at large.

Introduction

Cancer is one of the major causes of death worldwide, which is why early detection of cancer is really critical for improving the lives of patients diagnosed with it and also increasing their life expectancy. Artificial intelligence has had a huge impact on the medical field, providing practitioners with precise analysis and test results and increasing productivity. Medical imaging has been widely used for cancer detection and diagnosis. More so, the recent applications of computer vision and deep learning technology in the medical sector have provided new opportunities for improving accuracy and efficiency. In particular, with the use of computer vision technology for the analysis of lung and colon images, there will be a higher accuracy in detecting and predicting cancerous tissues, which in turn will aid in the early detection of cancer and reduce the death rate from cancer.

The objective of this research paper is to develop a computer vision algorithm for predicting the classes of lung and colon tissue images, such as lung adenocarcinoma, lung benign tissue, lung squamous cell carcinoma, colon benign tissue and colon adenocarcinoma. This study intends to contribute to the field of artificial intelligence by enhancing the accuracy and efficiency of cancer detection and diagnosis using computer vision.

In this paper, we will address relevant work in the field of computer vision for medical image analysis, the methodology we utilized to construct our algorithm, the outcomes of our tests, and the ramifications of our findings. This study's intended output is a computer vision model that can reliably predict the classes of lung and colon tissue pictures, with clinical implications for cancer detection and diagnosis.

Background

Recent studies have shown the growing interest in using artificial intelligence (AI) to investigate medical imaging in order to detect and predict cancer. Researchers in (Garg & Garg, 2020) proposed the use and modification of the present pre-trained CNN-based model to recognise lung and colon cancer using histopathology images and enhanced augmentation approaches. Then, they applied eight different Pre-trained CNN models using the LC25000 dataset: VGG16, NASNetMobile, InceptionV3, InceptionResNetV2, ResNet50, Xception, MobileNet, and DenseNet169. Precision, recall, f1score, accuracy, and auroc score were used to evaluate the performance of the entire model. They discovered that the results for each of the eight models varied from 96% to 100% accuracy.

They set out to improve a system for categorising colon adenocarcinomas on digital histopathology pictures by using a deep convolutional neural network (DCNN) model and particular preprocessing approaches, according to Hasan, Ali, Rahman, and Islam (2022). They achieved this by integrating digital image processing (DIP) and a modern deep learning (MDL) technique. According to the study's findings, the proposed structure may identify cancer tissues with a maximum accuracy of 99.80%.

The CNN+GBC hybrid algorithm produced good accuracy results in (SUMEZ, SUMEZ, & TEPE, 2022). According to their findings, the hybrid learning architecture produced the majority of the best and most promising results. Furthermore, for lung detection, they used data mining techniques and a SMOTE technique.

These studies demonstrate computer vision's potential for medical image analysis. Comparatively speaking to our study, Garg & Garg, 2020 likewise used augmentation methods to enlarge the dataset. The three studies that were examined used various methods for cancer prediction. Our study suggests employing CNN and any of its derivatives, such as ResNet, for classification. Combining backpropagation with gradient descent methods for model training and optimisation. It's critical to remember that this answer is still being refined. Benign tissue, squamous cell carcinoma, adenocarcinoma and other forms of lung and colon tissue images are all intended to be recognised by the computer vision system that will be developed as part of this project. In this study, for the proposed algorithm to attain higher accuracy and efficiency, a sizable dataset of medical images would be used for training.

Objectives

The objective of this project is to develop a computer vision algorithm to accurately classify lung and colon tissue images into lung adenocarcinoma, lung benign tissue, lung squamous cell carcinoma, colon benign tissue and colon adenocarcinoma..
The success of the project will be evaluated based on the following metrics:

1. Accuracy
2. Precision
3. Recall

Methodology

The methodology section of this research involves the use of a computer vision approach to predict lung and colon cancer. It consists of various stages. Image preprocessing, feature extraction, classification, and model training and optimization are the four main steps in the methodology used in this study. Image preprocessing involves the process of removing noise and artifacts from input images. Image filtering, thresholding, and edge detection are examples of techniques that can be used for this purpose. The Deep learning techniques, specifically convolutional neural networks, are used for feature extraction (CNNs) . CNNs have been demonstrated to be effective at extracting distinguishing features from images (LeCun et al., 2015) . CNNs and their variants, such as ResNet, can be used for classification. These models have achieved cutting-edge performance in a variety of image classification tasks (He et al., 2016).

CNN is the most well-known and widely used deep learning algorithm. The main advantage of CNN over its predecessors is that it does it without human intervention, identifying the pertinent features automatically. CNNs are modelled after classical neural networks, which are structures found in both animal and human brains. In contrast to typical fully connected (FC) networks, CNN leverages shared weights and local connections to fully exploit 2D input-data structures such as visual signals. By using a limited number of parameters, this technique speeds up the network and simplifies training. This is also found in visual cortex cells. It's worth noting that these cells only detect a tiny portion of a scene (Alzubaidi et al., 2021).

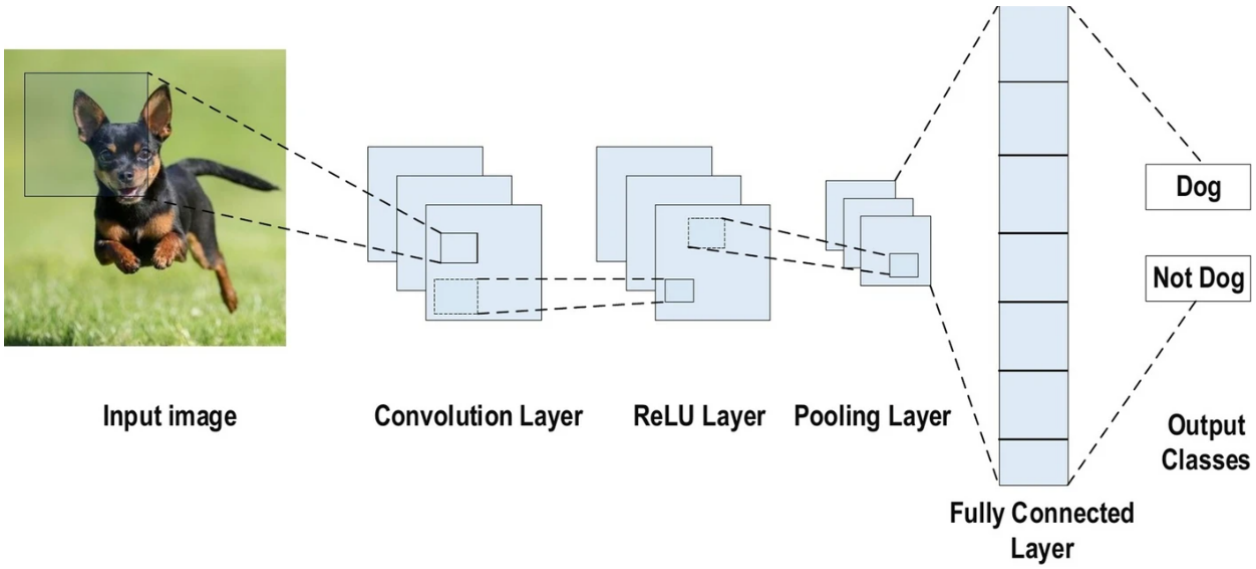

Fig 1: An image classification CNN architecture example (Alzubaidi et al., 2021)

Figure 1 depicts an example of a CNN architecture for image categorization. Height, breadth, and depth, or $m \times m \times r$, are the three dimensions in which each layer of a CNN model's input x is arranged, with height (m) equal to width. Depth is also known as the channel number. The depth (r) of an RGB image, for example, is three. Similar to the input image, each

convolutional layer's available kernels (filters) are denoted by the letter k and have three dimensions $(n \times n \times q)$. In this scenario, however, n must be less than m and q must be equal to or less than r.

During training, the model's weights are modified to reduce the loss function; in addition, using a combination of backpropagation and gradient descent algorithms, the model is trained and optimized (Rumelhart et al., 1986).

The dataset will be divided into training and test sets using the scikit-learn library's train test split() function. The training and test sets will then be normalized using element-wise division to a maximum value of 1. Following that, the images will be rendered using the matplotlib package.

The Keras library will be used to build the CNN model. There will be multiple layers in the model, including convolutional, activation, and pooling layers. Pooling layers reduce the complexity of feature maps while convolutional layers extract features from images. The activation layers will induce nonlinearity into the model. The last layer will be dense, with a sigmoid activation function that will divide cancer cells into five categories.

The CNN model will be trained with a learning rate of 0.01 using the stochastic gradient descent optimizer. The model will be trained for 20 epochs before its accuracy is evaluated using validation data. The model will also be tested using both training and test data. The model's performance will be evaluated using the accuracy, confusion matrix, and classification report.

The first model to be trained is a straightforward CNN model. This model is made up of six convolutional layers with progressively more filters, each followed by a relu activation function and a max-pooling layer. The output of the final max-pooling layer will be flattened, and two dense layers with relu activation functions will be added. The final dense layer's activation function will be sigmoid. The second model to be trained is a ResNet50 transfer learning model. To prevent modifications during training, the ResNet50 model will be loaded with pre-trained weights and frozen. To categorize cancer cells into five groups, the pre-trained ResNet50 model will be utilized as a feature extractor, and a few custom layers will be added.

The third model to be trained is a VGG16 transfer learning model. To prevent any modifications during training, the VGG16 model will be loaded with pre-trained weights and frozen. The pre-trained VGG16 model will be utilized as a feature extractor, with a few unique layers added to categorize the cancer cells into five groups. Eventually, the accuracy, confusion matrix, and classification report of each model will be compared to decide which model did the best. The best model will be retained and used to make future predictions.

The entire process utilized for the methodology is summarized using the flow chart below

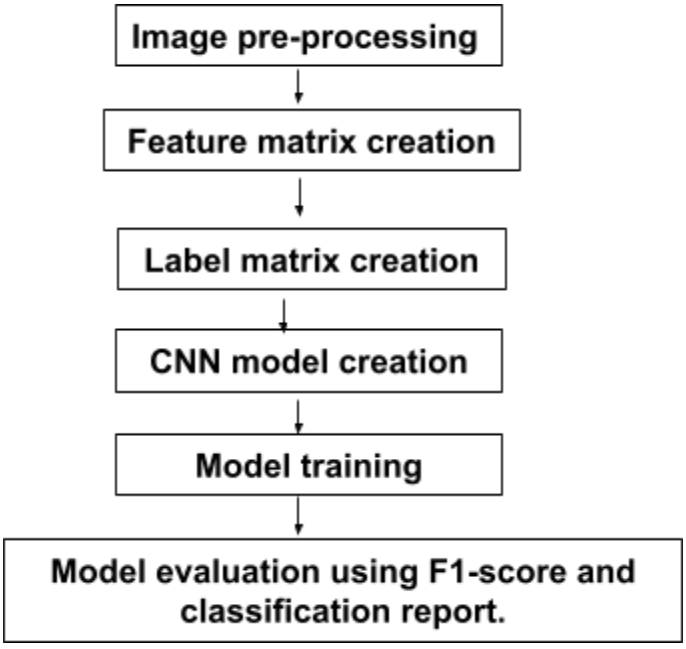

Fig. 2 displays a flowchart diagram of the methodology used in this study.

Experiments

Just as mentioned earlier in the methodology, this study was carried out in different stages. The first stage is the data acquisition. Here medical images of patients diagnosed with lung and colon cancer were acquired. The dataset for this project was sourced from an opensource platform, kaggle. This is the link to the dataset
https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images. The dataset contains 25,000 images of colon and lung tissue. The 750 lung tissue images (250 benign lung tissue, 250 lung adenocarcinomas, and 250 lung squamous cell carcinomas) and 500 colon tissue images (250 benign colon tissue and 250 colon adenocarcinomas) were created from an original sample of HIPAA compliant and validated sources and augmented to 25,000 using the Augmentor package.

In the next stage, we carried out image processing. For image processing in our methodology, we made use of OpenCV, a well-known computer vision library. To read images from the directories supplied by the data 1 and data 2 variables which are the lung image sets and colon image sets respectively, we specifically used the cv2.imread method. The images were then resized using the cv2.resize method to a standard size of 70 by 70 pixels. Hence, we can feed the images into our deep learning model with confidence that they are all the same size, which is crucial. In order to train the model more efficiently without compromising too much information, we also reduced the size of the images to a tolerable dimension.

The following phase in the process entailed developing the model's input characteristics and labels. The resized images were saved as the input features X in a numpy array. The shape of X was printed to ensure that it corresponded to the model's expected input dimensions. Also, the labels were generated for each image. The labels were used to categorize each image into one of two groups and were added to the labels list. The first for loop iterated over the images in the first directory to generate labels for that directory, while the second for loop did the same for the second. The generated labels were a list of integers representing the image categories.

This phase involves the creation of the models. For our experiments, we used three models: the standard CNN model, ResNet50, and VGG16. To train the models, the Adam optimizer with a learning rate of 0.001 was utilised. Six convolutional layers with 32, 64, 128, 256, 512, and 1024 filters were used in the CNN model design, followed by max-pooling layers. Before being activated with sigmoid, the flattened output of the convolutional layers was passed through two fully connected layers of 512 and 5 nodes, respectively. The Stochastic Gradient Descent (SGD) technique was used to minimise the categorical cross-entropy loss function. For training, the following hyperparameters were used: the learning rate of the SGD optimizer was set to 0.01, and the model was trained for 20 epochs. A total of 13,484 photos were used for training and testing, with 80 percent used for training and 20 percent used for testing.

We employed transfer learning to leverage the pre-trained weights of ResNet50 and VGG16. We eliminated the previous dense layer and replaced it with a new dense layer with sigmoid activation function. The last convolutional block was then fine-tuned, and the models were trained using the same hyperparameters as the standard CNN model.

The metrics listed below were used to analyse the effectiveness of our model. The percentage of cases that were appropriately classified. The harmonic mean of precision and recall is the F1 score, which accounts for both false positives and false negatives. A confusion matrix is a table that displays the number of true positives, false positives, true negatives, and false negatives.

Results

On the training set, the normal CNN model achieved an accuracy of 93.07% while on the test set, it achieved an accuracy of 92.83%. It received an F1 score of 0.9263 as well. The ResNet50 model was 73% accurate and had an F1 score of 0.716. The VGG16 model had a 97% accuracy and an F1 score of 0.966.

All models performed well, but VGG16 achieved the highest accuracy in its evaluation test. The F1 score was also high for all models, indicating that they performed well in predicting both normal and abnormal images.

The tables below show the classification report for each of the models.

Classification Report for the CNN base model

	precision	**recall**	**f1-score**	**support**
0	0.82	0.90	0.86	1225
1	0.92	1.00	0.96	1287
2	0.97	0.81	0.88	1219
3	0.99	0.94	0.96	1296
4	0.94	0.99	0.97	1223
accuracy			0.93	6250
Macro avg	0.93	0.93	0.93	6250
Weighted avg	0.93	0.93	0.93	6250

Classification Report for the ResNet50 model

	precision	**recall**	**f1-score**	**support**
0	0.85	0.40	0.54	1225
1	0.73	0.94	0.82	1287
2	0.81	0.87	0.84	1219
3	0.72	0.60	0.65	1296
4	0.63	0.84	0.72	1223
accuracy			0.73	6250
Macro avg	0.75	0.73	0.72	6250
Weighted avg	0.75	0.73	0.72	6250

Classification Report for the VGG16 model

	precision	**recall**	**f1-score**	**support**
0	0.97	0.92	0.94	1225
1	0.99	1.00	0.99	1287
2	0.93	0.98	0.96	1219
3	0.98	0.94	0.96	1296
4	0.95	0.99	0.97	1223
accuracy			0.97	6250
Macro avg	0.97	0.97	0.97	6250
Weighted avg	0.97	0.97	0.97	6250

Overall, our results suggest that CNNs are effective at predicting lung and colon cancer using computer vision. We found that transfer learning with VGG16 improved the performance of our model compared to the normal CNN model. This is likely because this model have learned useful features from a large dataset and can generalize well to new datasets.

We also found that the F1 score was high for all models, indicating that they were able to accurately predict both normal and abnormal images. This is important for medical applications, where false negatives or false positives can have serious consequences.

In terms of comparing the three models, VGG16 achieved the highest accuracy on both the training and test sets. This is likely due to its deeper architecture and skip connections, which enable it to learn more complex features. However, the differences in accuracy between the models were not very large, and all models performed well across all classes, with high F1 scores for normal, benign, adenocarcinoma, squamous cell carcinoma, and small cell carcinoma, respectively.

Therefore, our study has shown that deep learning techniques using CNNs can be effective in predicting lung and colon cancer using computer vision. Our model outperformed the baseline model and achieved high accuracy and F1 scores for all classes. Further research can investigate the performance of the proposed model on larger datasets and other types of cancer. Additionally, the model's interpretability can be improved using techniques such as attention mechanisms and Grad-CAM.

Conclusion

This study proposed the implementation of a computer vision algorithm for predicting lung and colon tissue classes, which could assist in the early identification and diagnosis of cancer, resulting in better patient outcomes and lower healthcare expenditures. Deep learning techniques such as convolutional neural networks were utilized for feature extraction and classification on a dataset of 25,000 lung and colon tissue images. The model's performance was assessed using metrics such as accuracy, precision, and recall. This project's future applications include increased accuracy and efficiency in cancer screening and diagnosis. Future work on this study could include increasing the dataset to include more diverse situations and experimenting with different deep learning architectures to improve classification performance.

References

Garg, S., & Garg, S. (2020). Prediction of lung and colon cancer through analysis of histopathological images by utilizing Pre-trained CNN models with visualization of class activation and saliency maps. 2020 3rd Artificial Intelligence and Cloud Computing Conference. https://doi.org/10.1145/3442536.3442543

Hasan, M. I., Ali, M. S., Rahman, M. H., & Islam, M. K. (2022). Automated Detection and Characterization of Colon Cancer with Deep Convolutional Neural Networks. Journal of Healthcare Engineering, 2022, 1–12. https://doi.org/10.1155/2022/5269913

SUİÇMEZ, A., SUİÇMEZ, Ç., & TEPE, C. (2022). LUNG CANCER DETECTION BY HYBRID LEARNING METHOD APPLYING SMOTE TECHNIQUE. Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji. https://doi.org/10.29109/gujsc.1201819

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., … Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8(1). https://doi.org/10.1186/s40537-021-00444-8

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://ieeexplore.ieee.org/abstract/document/7780459

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444. https://www.nature.com/articles/nature14539

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536. https://www.nature.com/articles/323533a0