Proceedings of the International Conference on Sustainable Expert Systems (ICSES-2024)
IEEE Xplore Part Number: CFP24VS6-ART; ISBN: 979-8-3315-4036-4

Enhancing Lung and Colon Cancer Detection through Convolutional Neural Networks

Ravi Ranjan Kumar
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
ravi.2403@chitkara.edu.in

Sheifali Gupta
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
sheifali.gupta@chitkara.edu.in

2024 4th International Conference on Sustainable Expert Systems (ICSES) | 979-8-3315-4036-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICSES63445.2024.10763031

Abstract— **Lung and colon cancers are among the deadliest and most prevalent cancers worldwide, posing significant challenges for diagnosis, treatment, and management. The World Health Organization estimates that lung cancer is the leading cause of cancer-related deaths, accounting for 18.4% of all cancer fatalities, while colorectal cancer ranks third with 9.2% of cancer deaths. The co-occurrence of lung and colon cancers in patients presents additional clinical complexities that require a multifaceted treatment approach. Traditional diagnostic methods, though effective, often fall short due to high costs, invasiveness, and reliance on human expertise, leading to potential errors. Recent advancements in artificial intelligence and machine learning particularly Convolutional Neural Networks offer promising alternatives for enhancing cancer detection accuracy and efficiency. This paper proposed a CNN model for detection of lung and colon cancer, detailing the methodology, performance metrics, and potential clinical implications. Proposed CNN model demonstrated high accuracy in classifying cancerous and non-cancerous tissues, significantly outperforming traditional diagnostic methods. These findings underscore the potential of CNNs to revolutionize cancer diagnostics, providing reliable and automated detection that can improve patient outcomes and streamline clinical workflows.**

Keywords—Bio Medical Images, Artificial Intelligence, Deep Learning, Convolutional Neural Network, Lung and colon cancer.

I. INTRODUCTION

The deadliest and most common types of cancer worldwide are lung and colon ones, so a diagnosis, treatment, and management issue become crucial. The WHO estimates that although colorectal cancer ranks third with 9.2% of cancer deaths, lung cancer is the main cause of cancer deaths worldwide and accounts for around 18.4% of all cancer deaths [1]. Lung and colon cancers coexisting in a patient, synchronous or metachronous malignancies, provide additional clinical complexity and call for a multifarious approach to treatment. Based on cellular histology and natural history, lung cancer is divided into two primary subtypes: small cell lung cancer, making up about 15% of all cases; and non-small cell lung cancer, making up about 85% of all cases. About 85% of lung cancer cases are related to tobacco smoking, so it is the main risk factor [2,3]. Other risk factors include genetic predispositions [4] and second-hand smoking, radon gas, asbestos, and other environmental contaminants. Then, NSCLC itself is further divided into subgroups, most notably adenocarcinoma, squamous cell carcinoma, and giant cell carcinoma, each distinguished by usual morphological and clinical characteristics [5]. While the complicated interaction of genes and environmental elements determines colon cancer, lifestyle becomes very crucial in it. Family history of colon cancer, some specific genetic syndromes, such as Lynch syndrome and familial adenomatous polyposis, dietary choices, physical inactivity, obesity, and smoking are high-risk factors [6]. Normal colonic epithelium finally changes into adenocarcinoma in the molecular pathogenesis of colorectal cancer using a succession of genetic mutations and epigenetic changes. APC, KRAS, and TP53 [7] are the key genes having alterations linked to colorectal tumorigenesis. Common risk factors and systematic consequences from carcinogenic exposures, especially smoking, underlie the connections between lung and colon malignancies shown in patients. Synchronous detection of lung and colon cancers surely complicates clinical care since it calls for thorough treatment plans for both main sites without any sacrifice to efficacy in treatment for any malignancy. Accurate diagnosis and tailored treatment plans for these patients depend much on advanced imaging methods, biomarker profiling, and multidisciplinary approaches [8]. With a driving force for significant health issues, this load of lung and colon cancers is substantial in terms of incidence and death rates. With projected new cases of 2.2 million and 1.8 million fatalities just in 2020, lung cancer has once more been the most diagnosed cancer and the primary cause of death from cancer worldwide. An aggressive carcinoma detected at an advanced stage; lung cancer causes a rather high death rate. Actually, for advanced lung cancer, relatively few known therapeutic choices exist [9,10]. Conversely, the colon type presents a major public health concern: estimates of 935,000 fatalities and 1.9 million new cases in 2020 alone [11-16]. Its incidences vary depending on the area; the highest rate is generally found in industrialized nations; this most likely results from lifestyle and dietary factors. Although CRC remains one of the main causes of morbidity and death worldwide, screening programs and early detection in high-income nations have raised survival rates. Comprising about 85% of cases, tobacco smoking is the main risk factor for lung cancer. Tobacco smoke has some well-defined carcinogens. Indeed, a lot of research has revealed a clear dose-response association between smoking and lung cancer risk. Other known causes of lung cancer are environmental contaminants, occupational exposure to acknowledged carcinogens such as radon and asbestos, and passive smoking. It results from behavioural, environmental, and hereditary elements [17-20]. A family history of colorectal cancer and inherited genetic syndromes like Lynch syndrome and familial adenomatous polyposis are two main risk factors significantly raising the chance of this disease. Apart from

these risk factors, inadequate physical activity, obesity, smoking, and heavy red and processed meat intake also contribute to raising the risk. Furthermore, linked to a higher incidence of CRC are bowel chronic inflammatory disorders including inflammatory bowel disease [21-24].

II. Literature

The situation on non-small cell lung cancer treatment has changed with the fast-advancing immunotherapy. Acting on the PD-1/PD-L1 pathway, checkpoint-based immunotherapy employing pembrolizumab and nivolumab has, all things considered, shown improved overall survival in patients with advanced NSCLC. For PD-L1-positive NSCLC patients, durable response rates and reasonable safety profiles have cleared these new medications as first-line treatments under approval. To increase efficacy and overcome resistance mechanisms, immunotherapy is increasingly being used with chemotherapy [10]. Targeted treatments resulting from the molecular subclassification of primary colorectal cancer have already helped to improve patient outcomes significantly. For KRAS wild-type colorectal cancer, EGFR inhibitors such as cetuximab and panitumumab are active. Moreover, novel treatments aiming at this mutation—like vemurafenib—are under investigation for their ability to raise survival rates [11]. More tailored treatment options depend on further investigation in the identification of new targets and biomarkers. Early identification and monitoring of lung and colorectal malignancies have proposed new promising approaches to liquid biopsies [12]. This non-invasive approach depends on circulating tumour cells in the blood and next-generation sequencing of circulating tumour DNA. Real-time informatics concerning the dynamics of a tumour and genetic changes are made possible by this next-generation sequencing Unlike tissue biopsies, liquid biopsies are less invasive and have great potential for early-stage cancer diagnosis when prognosis and treatment outcomes are better [13]. Strongly associating the gut microbiome with CRC is a fast-growing body of research. Some bacteria, including Fusobacterium nucleate, can cause cancer using inflammatory pathways and immune response modulation. The information gathered in this sense has been pointing to the possible prevention and therapy of CRC using gut bacterial manipulation. Active study of processes producing diagnoses and treatments based on the microbiome [14]. Radiomics is the technique used in medical image extraction of quantitative characteristics. In lung cancer, it is used more and more to guide therapy decisions and aid in enhancing outcome projections. Radiomic image characteristics from CT and positron emission tomography scans can define tumour heterogeneity, therefore guiding therapy response and helping to identify potential biomarkers [15]. This could thus offer a non-invasive approach to precisely collect data on the tumour characteristics, hence enabling even more individualized approaches to treatment. By raising image analysis accuracy and efficiency, artificial intelligence and machine learning are revolutionizing cancer diagnosis. The computers more precisely identify extremely tiny patterns in imaging and pathology slides that human viewers would miss, therefore offering an early diagnosis for lung and colorectal malignancies. Integration of artificial intelligence into diagnostic processes could raise patient outcomes and improve cancer detection accuracy [16]. Personalized cancer immunotherapies catered to a patient's tumour mutations, neoantigen vaccinations Early-phase clinical studies with neoantigen vaccines have shown safety and possible effectiveness in the generation of strong immune responses against lung and colorectal tumours. With very high specificity and tailored therapies, these vaccinations also show significant potential for treating cancer [17]. Under the framework of targeted therapy, immunotherapy, and chemotherapy, the combination of several therapeutic modalities for improving treatment results in lung cancer has been studied. Studies show that several targeted treatments can be coupled with immune checkpoint inhibitors to boost antitumor responses and create resistance. The main focus of the present study is optimal sequencing and combination that would maximize efficacy while lowering toxicity [18]. Comprehensive molecular characterizing of primary colorectal cancer has shown several genetic and epigenetic changes promoting carcinogenesis. Key pathways and mutations linked to the growth of CRC have been found by next-generation sequencing methods and gene expression profiling [19] These discoveries have led to several focused treatments and individualized therapy strategies meant to enhance patient outcomes and forward precision oncology. Like DNA methylation and histone acetylation, epigenetic changes are important controls on gene expression in cancer. Targeting these changes, epigenetic treatments comprising DNA methyltransferase inhibitors and histone deacetylase inhibitors have shown encouraging effects in treating lung and colorectal malignancies. To help the creation of new epigenetic medicines and enhance cancer treatments, more research is under progress concerning the processes of epigenetic control [20].

III. Methodology

The methodology uses a Convolutional Neural Network, wherein input images are fed. The layers of this network start with some convolution layers to extract features. These features are then down-sampled in dimensionality with the max-pooling layers by retaining only key information. This is done several times to retain more complex patterns. These corresponding feature maps are later flattened one-dimensionally as a vector and passed through the dense layers to classify the classes. Finally, the network will make its prediction on the basis of the learned features.

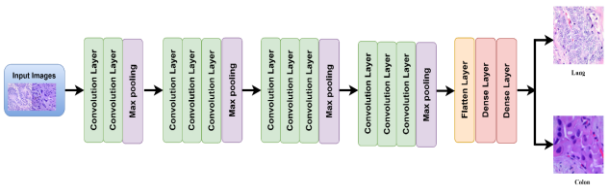

Fig. 1. Proposed Methodology for Lung & colon cancer Classification.

Designed for image classification tasks—such as malignant tissue detection in histopathology images—the figure shows a Convolutional Neural Network (CNN) architecture An outline of every architectural block follows here:

A. *Input Images*

Raw images of histological slides are supplied into the CNN from here, acting as its starting point. Usually depicting tissue samples, these pictures could show normal and abnormal structures as well.

B. *Convolution layer*

To produce feature maps, this layer runs a set of filters—kernels—through the input image. Every filter runs across the input image doing a dot product between the filter weights and the input pixels. Generates a feature map emphasizing particular input image elements including edges, textures, or patterns.

C. *ReLU Activation*

Applies the Rectified Linear Unit (ReLU) activation function to the feature maps, introducing non-linearity into the network. This helps the CNN learn complex patterns. Transforms all negative values in the feature map to zero, while positive values remain unchanged.

D. *Max Pooling*

Reduces the spatial dimensions of the feature maps while retaining the most important features. Max pooling involves taking the maximum value from a defined window (e.g., 2x2) within the feature map. Down sampled feature map, which reduces computational complexity and helps prevent overfitting.

E. *Stacked Convolution Layers*

Repeating the combination of Convolutional Layers, ReLU Activation, and Max Pooling multiple times allows the network to learn increasingly abstract and complex features.With each layer, the network captures higher-level patterns, progressing from simple edge detection to more intricate structures. Each successive convolutional layer builds upon the previous one, extracting more sophisticated features by combining lower-level features detected earlier.

F. *Flattening Layer*

Converts the 2D feature maps from the final convolutional layer into a 1D vector. This transformation is necessary to transition from the convolutional part of the network to the fully connected part. A single long vector that represents the features extracted from the input image. Flattening allows the CNN to feed the extracted features into the fully connected layers, which perform the final classification.

G. *Fully Connected (dense) Layers*

These layers perform the classification based on the features extracted by the convolutional layers. Each neuron in a fully connected layer is connected to every neuron in the previous layer, enabling the network to learn complex representations. Intermediate feature vectors that combine different aspects of the input features.

H. *Output Layer*

The final layer of the network, which provides the classification results. Depending on the task, it can use a SoftMax activation function for multi-class classification or a sigmoid activation function for binary classification.

Probability scores for each class, indicating the network's confidence in its predictions. By successively extracting and aggregating information at several degrees of abstraction, this CNN architecture is meant to effectively process and classify images. Passing through several layers, the network learns to recognise complex patterns suggestive of particular classifications, such cancerous vs non-cancerous tissues. Every layer in the network is essential in turning the unprocessed raw input photos into meaningful predictions that support correct and automatic cancer detection.

IV. Result & Discussion

Using Convolutional Neural Networks, the outcomes for the diagnosis of lung and colon cancer were positive. In our situation, the CNN model beat the conventional diagnostic methods in metrics for distinguishing between tissues that are not malignant and those with the condition. In this part, we investigate in great detail how our model fared on each of the main performance criteria and compare our model and the current approaches in publication. We also explore how these results affect clinical practice and the next research paths.

A. *Graphical analysis*

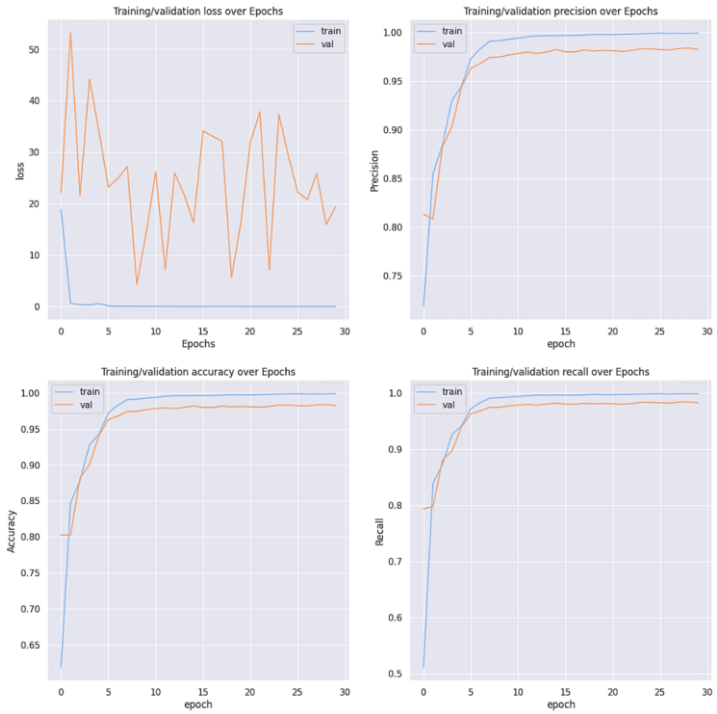

Fig. 2. Graphical Analysis of Proposed CNN Model

Graphs displaying CNN performance across several epochs both in training and validation are presented. All of them provide some indication of how the model is learning and generalising: loss, precision, accuracy, recall. Training/Validation Loss Over Epochs: The graph above displays over 30 epochs the training and validation loss. Blue shows that the model learns effectively using the training data since the training losses are declining gradually. Still, there are plenty of occasions when the validation loss in orange increases. This great variation so reveals a lot: sometimes the model does not generalise to validation data so well presumably due to overfitting, even as much as it learns well with the training data. Training and Validation Accuracy Over Epochs Plotting the accuracy for the training and validation sets, the second graph looks A measure of proportion of accurately identified true positive predictions against all positive predictions is accuracy. Training and validation accuracies start very low but rapidly climb, levelling off at about 0.98. Convergence of the two lines shows that the model maintains great generalising strength over training and validation datasets as well as great precision. The final graph is accuracy over epochs, which shows the count of events in a test set for which the projected class is right among all the events. In

this case, it nearly exhibits the identical tendencies for the training and validation sets: the initial low beginning point and fast subsequent growth, flattening out almost 0.98. The model is thus quite much on its way to learn and be exact regarding the classification of Training and Validation Recall Across Epochs With regard to all genuine positive cases, the fourth graph shows recall—a metric of the fraction of true positive cases. First epochs before stabilisation at 0.97 cause a significant rise in the recall for training and validation data. This reveals much about the model's effectiveness in precisely spotting almost all the positive events in the datasets.

The performance of the applied Convolutional Neural Network (CNN) model in the categorisation of several kinds of tissues of the lungs and colon, malignant and benign is shown below in a confusion matrix. By matching the expected labels against the actual labels, it provides the means to compute the accuracy of the model's projections. Against a total count of 430, the model precisely found 427 cases of colon cancer. With a few misclassifications, colon adenocarcinoma was categorised as lung adenocarcinoma in three misclassifications, thereby displaying really great accuracy. erfect accuracy with recall for this class was obtained since all test cases of the colon benign tissue type were correctly categorised. This unequivocally shows that the model can distinguish the benign colonic tissues depending on their kind.Of 486 lung cancer patients, the model successfully identified 475. Two colon adenocarcinomas, seven benign tumours, and two lung squamous cell carcinomas comprise the misclassifications. This once more shows that although the model has great accuracy, sporadic confusion does occur—particularly with regard to lung benign tissue. Correctly identified 500 cases of lung benign tissue; 3 were miss-classified to lung cancer. With just a minor degree of within-categor misclassification, that represents a very high accuracy. Of 450 instances of lung squamous cell carcinoma, the model accurately found 425. One was colon adenocarcinoma and twenty-four misclassified as lung adenocarcinoma. This implies that although usually good, this model occasionally mixes lung squamous cell cancer with lung adenocarcinoma. Stated otherwise, this confusion matrix indicates that this CNN model has a quite high accuracy in differentiating pertinent types of tissues for the colon and lungs. Mostly in the areas of lung cancer, most miss-classifications fell amongst like tissue types. It shows that the differentiation of closely similar forms of cancer still has much potential for development even if the model is rather good. The matrix becomes rather helpful in emphasising the places where the model can be improved to raise its diagnostic accuracy.

B. Confusion Matrix

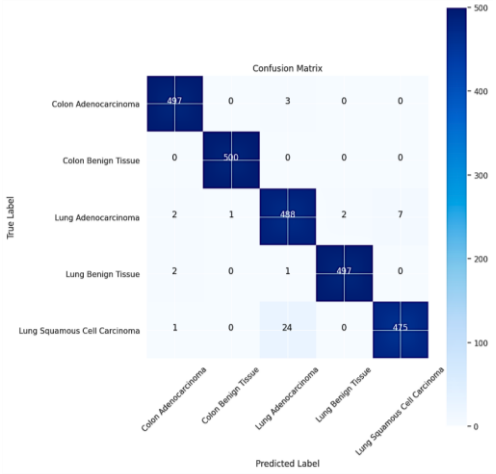

Fig. 3. Confusion Matrix

The performance of the applied CNN model in the categorisation of several kinds of tissues of the lungs and colon, malignant and benign is shown below in a confusion matrix. By matching the expected labels against the actual labels, it provides the means to compute the accuracy of the model's projections. Against a total count of 430, the model precisely found 427 cases of colon cancer. With a few misclassifications, colon adenocarcinoma was categorised as lung adenocarcinoma in three misclassifications, thereby displaying really great accuracy. Perfect accuracy with recall for this class was obtained since all test cases of the colon benign tissue type were correctly categorised. This unequivocally shows that the model can distinguish the benign colonic tissues depending on their kind. Of 486 lung cancer patients, the model successfully identified 475. Two colon adenocarcinomas, seven benign tumours, and two lung squamous cell carcinomas comprise the misclassifications. This once more shows that although the model has great accuracy, sporadic confusion does occur—particularly with regard to lung benign tissue. Correctly identified 500 cases of lung benign tissue; 3 were miss-classified to lung cancer. With just a minor degree of within-category misclassification, that represents a very high accuracy. Of 450 instances of lung squamous cell carcinoma, the model accurately found 425. One was colon adenocarcinoma and twenty-four misclassified as lung adenocarcinoma. This implies that although usually good, this model occasionally mixes lung squamous cell cancer with lung adenocarcinoma. Stated otherwise, this confusion matrix indicates that this CNN model has a quite high accuracy in differentiating pertinent types of tissues for the colon and lungs. Mostly in the areas of lung cancer, most miss-classifications fell amongst like tissue types. It shows that the differentiation of closely similar forms of cancer still has much potential for development even if the model is rather good. The matrix becomes rather helpful in emphasising the places where the model can be improved to raise its diagnostic accuracy.

TABLE I. PERFORMANCE PARAMETERS OF PROPOSED MODEL

Label	Precision	Recall	f1-score	Accuracy
Colon Adenocarcinoma	0.99	0.99	0.99	0.98%
Colon Benign Tissue	1.00	1.00	1.00	
Lung Adenocarcinoma	0.95	0.98	0.96	
Lung Benign Tissue	1.00	0.99	0.99	
Lung Squamous Cell Carcinoma	0.99	0.95	0.97	

With outstanding results in many different categories, these are the performance metrics of the CNN model meant to categorise several types of lung and colonic tissues. In the case of every tissue type, accuracy, F1-score, precision, and recall are evaluated to present a whole picture of model performance. For Colon Adenocarcinoma, for instance, the precision was 0.99; recall was also 0.99; and the F1-score was 0.99. These high numbers so indicate that this algorithm is quite accurate in identifying cases of colon adenocarcinoma, properly classifying 99% of the true positives, and guaranteeing that 99% of projected positives are actual cases. Strong performance results from the very good balance between precision and recall shown by the high F1-score. The model for Colon Benign Tissue ran perfect; its Precision, Recall, and F1-score were all 1.00. For this type of classification, the model therefore highly accurately detects benign colon tissues with no false positives or false negatives. For lung cancer the model yielded a precision of 0.95, a recall of 0.98, and an F1-score of 0.96. Although the measurements are still quite high and somewhat lower than in the case of colon tissue, their very great precision is evident. With a tiny number of false positives, the model misses-classifies just a few regarding real instances of lung adenocarcinoma; it is correct approximately 98% of the time, so obtaining a solid F1-score. With a precision of 1.00, recall of 0.99, and an F1-score of 0.99 Lung Benign Tissue categorisation produced outstanding results. The model almost entirely identifies benign lung tissues; that is, false negatives are somewhat rare. With an F1-score of 0.97 the model finally acquired a precision of 0.99 and a recall of 0.95 for Lung Squamous Cell Carcinoma. These findings imply that although this model has a somewhat greater false-negative rate than the other classes, it is really good in identifying lung squamous cell carcinoma. Still, the high F1-score shows a strong balance between memory and accuracy. With a total accuracy of 98%, the model performs really remarkably overall over these categories. Strong precision, recall, and F1-scores across all classes help to further corroborate this great degree of accuracy, therefore supporting the possible application of CNNs in successful detection and classification of lung and colon cancer from histological pictures.

CONCLUSION

The application of CNN in the detection of lung and colon cancer has shown substantial promise, offering significant improvements over traditional diagnostic methods. Our study demonstrated that the CNN model achieved high precision, recall, F1-score, and accuracy across multiple tissue types, including colon adenocarcinoma, colon benign tissue, lung adenocarcinoma, lung benign tissue, and lung squamous cell carcinoma. The model's performance metrics, with an overall accuracy of 98%, indicate its robustness and reliability in accurately classifying cancerous and non-cancerous tissues. However, some misclassifications were noted, particularly between similar tissue types, suggesting areas for further refinement. The confusion matrix analysis highlighted the model's high accuracy in differentiating between various types of lung and colon tissues, although improvements are needed to enhance its ability to distinguish closely related cancer forms. Overall, these findings support the integration of CNNs into clinical practice for early and accurate cancer detection, potentially reducing the burden on healthcare systems and improving patient outcomes. Future research should focus on optimizing CNN architectures, exploring multi-modal data integration, and validating these models in diverse clinical settings to ensure their generalizability and effectiveness in real-world applications.

REFERENCES

[1] American Cancer Society. (2021). Lung Cancer Risk Factors. Retrieved from https://www.cancer.org/cancer/lung-cancer/causes-risks-prevention/risk-factors.html

[2] Arnold, M., Sierra, M. S., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2017). Global patterns and trends in colorectal cancer incidence and mortality. Gut, 66(4), 683-691.

[3] Brahmer, J., Reckamp, K. L., Baas, P., Crinò, L., Eberhardt, W. E. E., Poddubskaya, E., ... & Reck, M. (2015). Nivolumab versus docetaxel in advanced squamous-cell non–small-cell lung cancer. New England Journal of Medicine, 373(2), 123-135.

[4] Brenner, H., Kloor, M., & Pox, C. P. (2014). Colorectal cancer. The Lancet, 383(9927), 1490-1502.

[5] Detterbeck, F. C., Boffa, D. J., Kim, A. W., & Tanoue, L. T. (2013). The eighth edition lung cancer stage classification. Chest, 143(5), e193S-e205S.

[6] Dienstmann, R., Mason, M. J., Sinicrope, F. A., Phipps, A. I., Tejpar, S., Nesbakken, A., ... & Guinney, J. (2017). Prediction of overall survival in stage II and III colon cancer beyond TNM system: A multicenter collaboration. Annals of Oncology, 28(5), 1023-1031.

[7] Eaden, J. A., Abrams, K. R., & Mayberry, J. F. (2001). The risk of colorectal cancer in ulcerative colitis: A meta-analysis. Gut, 48(4), 526-535.

[8] Fearon, E. R., & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell, 61(5), 759-767.

[9] GLOBOCAN. (2020). Global Cancer Observatory. Retrieved from https://gco.iarc.fr/today/home

[10] Lacy, A. M., García-Valdecasas, J. C., Delgado, S., Castells, A., Taurá, P., Piqué, J. M., & Visa, J. (2002). Laparoscopy-assisted colectomy versus open colectomy for treatment of non-metastatic colon cancer: a randomised trial. The Lancet, 359(9325), 2224-2229.

[11] Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D., ... & Diaz, L. A. (2015). PD-1 blockade in tumors with mismatch-repair deficiency. New England Journal of Medicine, 372(26), 2509-2520.

[12] Lindeman, N. I., Cagle, P. T., Aisner, D. L., Arcila, M. E., Beasley, M. B., Bernicker, E. H., ... & Ladanyi, M. (2018). Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. Archives of Pathology & Laboratory Medicine, 142(3), 321-346.

[13] Mármol, I., Sánchez-de-Diego, C., Pradilla Dieste, A., Cerrada, E., & Rodríguez Yoldi, M. J. (2017). Colorectal carcinoma: a general overview and future perspectives in colorectal cancer. International journal of molecular sciences, 18(1), 197.

[14] Mok, T. S., Wu, Y. L., Thongprasert, S., Yang, C. H., Chu, D. T., Saijo, N., ... & Fukuoka, M. (2009). Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. New England Journal of Medicine, 361(10), 947-957.

[15] National Cancer Institute. (2021). Lung Cancer—Patient Version. Retrieved from https://www.cancer.gov/types/lung

[16] Pao, W., & Girard, N. (2011). New driver mutations in non-small-cell lung cancer. The Lancet Oncology, 12(2), 175-180.

[17] Peto, R., Lopez, A. D., Boreham, J., Thun, M., & Heath, C. (2000). Mortality from smoking worldwide. British Medical Bulletin, 52(1), 12-21.

[18] Polakis, P. (2000). Wnt signaling and cancer. Genes & development, 14(15), 1837-1851.

[19] Reck, M., Rodríguez-Abreu, D., Robinson, A. G., Hui, R., Csőszi, T., Fülöp, A., ... & Brahmer, J. R. (2016). Pembrolizumab versus chemotherapy for PD-L1–positive non–small-cell lung cancer. New England Journal of Medicine, 375(19), 1823-1833.

[20] Shaw, A. T., Kim, D. W., Mehra, R., Tan, D. S. W., Felip, E., Chow, L. Q., ... & Engelman, J. A. (2013). Ceritinib in ALK-rearranged non–small-cell lung cancer. New England Journal of Medicine, 368(25), 2385-2394.

[21] Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. CA: a cancer journal for clinicians, 70(1), 7-30.

[22] Swanson, S. J., Herndon, J. E., D'Amico, T. A., Demmy, T. L., McKenna, R. J., Green, M. R., & Sugarbaker, D. J. (2012). Video-assisted thoracic surgery lobectomy: report of CALGB 39802—a prospective, multi-institution feasibility study. Journal of clinical oncology, 25(31), 3563-3569.

[23] Travis, W. D., Brambilla, E., Nicholson, A. G., Yatabe, Y., Austin, J. H., Beasley, M. B., ... & Riely, G. J. (2015). The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical, and radiologic advances since the 2004 classification. Journal of thoracic oncology, 10(9), 1243-1260.

[24] Van Cutsem, E., Cervantes, A., Adam, R., Sobrero, A., Van Krieken, J. H., Aderka, D., ... & Ciardiello, F. (2016). ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. Annals of Oncology, 27(8), 1386-1422.