AJPMHS GLOBAL RESEARCH NETWORK

AMERICAN Journal of Pediatric Medicine and Health Sciences
Volume 3, Issue 11, 2025 ISSN (E): 2993-2149

Deep Learning Driven Colon Cancer Diagnosis: Performance Assessment of Five Convolutional Neural Network Architectures on Histopathological Image Classification

Md Yousuf Ahmad
Masters of science Business Analytics, MSBAN, Trine University, USA

Babul Sarker
Master of Science in Business Analytics (MSBA), Trine University, USA

Sohel Mahmud
Masters of science Engineering Management, MSEM., Trine University, USA

Md Firoz Kabir
Master of Science in Information Technology, University of the Cumberlands, USA

Abstract: The integration of artificial intelligence (AI), particularly deep learning, has significantly advanced disease diagnosis and clinical decision making. This study evaluates the performance of five prominent deep learning architectures MobileNetV1, ResNet50, AlexNet, DenseNet201, and Inception v2for classifying colon adenocarcinoma versus benign colon tissue, a task essential for effective colon cancer management. Using a dataset of 3,000 histopathological images, each model was trained and tested to assess classification accuracy. Among the evaluated models, MobileNetV1 and AlexNet demonstrated the highest performance, achieving test accuracies of 96.33% and 95.67%, respectively. In contrast, ResNet50 and DenseNet201 showed comparatively lower accuracies of 85.80% and 87.40%, while Inception v2 reached 92.87%. These findings underscore the strong potential of lightweight architectures particularly MobileNetV1 and AlexNet in improving colon cancer detection and supporting clinical workflows. Future work will explore additional model architectures, evaluation metrics, and optimization techniques to further enhance diagnostic reliability. This study contributes to the expanding body of research on AI driven oncology, highlighting deep learning's role in advancing early and accurate cancer classification.

Keywords: Artificial intelligence, Deep Learning, Healthcare, Colon Adenocarcinoma, Benign tissue, Histopathological images, Oncology.

INTRODUCTION:

Artificial intelligence (AI) has become a transformative force in modern healthcare, driving significant advancements in diagnostic accuracy and clinical decision making. Among the various branches of AI, deep learning has shown exceptional capability in analyzing complex medical data, particularly within oncology. Colon cancer one of the leading causes of cancer related morbidity and mortality worldwide presents a compelling domain for the application of

deep learning based diagnostic tools. Accurate classification of colon cancer subtypes is essential for selecting appropriate treatment strategies and predicting patient outcomes.

Traditional diagnostic approaches rely heavily on manual examination of histopathological images, a process that is time consuming, subjective, and prone to inter observer variability. Deep learning models, however, offer an automated and highly scalable alternative by learning discriminative features directly from large datasets and identifying subtle morphological patterns that may be overlooked by the human eye. This study conducts a comprehensive evaluation of five deep learning architectures MobileNetV1, ResNet50, AlexNet, DenseNet201, and Inception v2 in their ability to distinguish colon adenocarcinoma from benign colon tissue. Using a dataset of 3,000 histopathological images equally divided between the two classes, each model was trained, validated, and assessed using key performance metrics, including accuracy, precision, recall, F1 score, and Cohen's Kappa score.

The goal of this research extends beyond identifying the highest performing model. By systematically comparing these architectures, we aim to contribute to the ongoing discussion on the role of AI in modern healthcare, highlight the strengths and limitations of different deep learning approaches, and identify future opportunities for model optimization. Ultimately, this investigation underscores the transformative potential of deep learning in colon cancer classification and supports its integration into clinical workflows to improve diagnostic precision, treatment planning, and patient outcomes. This work represents a step toward a future in which AI powered tools are seamlessly embedded into healthcare systems, advancing the quality of care for patients worldwide.

LITERARY SURVEY:

Lim et al. (2017) [1] conducted a comparative analysis of oncological outcomes between right-sided and left-sided colon cancers following curative resection, revealing significant survival differences and underscoring the importance of tumor laterality in prognosis and treatment planning. Complementing this perspective, Sears and Garrett (2014) [2] provided an extensive review on the role of gut microbiota in colon cancer development, highlighting the intricate interactions among microbial communities, immune responses, and tumorigenesis, and emphasizing the emerging therapeutic value of microbiome modulation.

Zhao et al. (2020) [3] further explored the prognostic significance of tumor laterality in synchronous metastatic colon cancer using National Cancer Database records, reinforcing the clinical relevance of tumor location. De Sousa e Melo et al. (2017) [4] identified the role of Lgr5+ stem cells in primary and metastatic colon cancer, revealing heterogeneity within cancer stem cell populations and suggesting new directions for stem-cell-targeted therapies. Similarly, Shimokawa et al. (2017) [5] developed innovative methods for visualizing and targeting LGR5+ colon cancer stem cells, strengthening their importance as therapeutic targets.

The molecular mechanisms of colon cancer progression were investigated by Zhou et al. (2018) [6], who demonstrated the role of Caspase-3 in regulating migration, invasion, and metastasis. Wang et al. (2019) [7] studied the radiosensitivity of colon cancer cells through the inhibition of circCCDC66, revealing circRNAs as promising targets for enhancing radiotherapy. Urosevic et al. (2014) [8] uncovered a previously unknown pathway by which colon cancer cells metastasize from liver to lung through p38 MAPK signaling and PTHLH, adding depth to the understanding of metastatic cascades.

Jahanafrooz et al. (2020) [9] reviewed colon cancer stem cells and their surrounding microenvironment, outlining therapeutic strategies aimed at disrupting stem cell niches. Tauriello et al. (2018) [10] demonstrated that TGFβ facilitates immune evasion in colon cancer metastasis, offering insights into immunotherapeutic strategies. Aiello et al. (2019) [11] evaluated preoperative immunonutritional support in malnourished cancer patients, emphasizing the importance of nutritional status in surgical recovery and prognosis.

Kuipers et al. (2015) [12] provided a comprehensive overview of colorectal cancer, covering epidemiology, molecular pathways, and treatment modalities. In the context of immunotherapy, Tan et al. (2019) [13] demonstrated the potential of dendritic cell–based vaccines, while Fan et al. (2022) [14] studied the immune microenvironment in colorectal cancer liver metastasis, identifying possible therapeutic targets. Derer et al. (2016) [15] reviewed immune checkpoint inhibitors in colorectal cancer, highlighting their expanding clinical relevance. Abdalla et al. (2023) [16] emphasized the prognostic and therapeutic importance of Microsatellite Instability (MSI), supporting precision medicine approaches.

Petrelli et al. (2017) [17] conducted a meta-analysis on the effect of tumor location on survival in metastatic colorectal cancer, confirming its role as a key prognostic factor. Ahmed et al. (2023) [18] reviewed the molecular genetics of colorectal cancer, shedding light on its complex genomic landscape. A groundbreaking study by Le et al. (2017) [19] demonstrated that mismatch-repair deficiency predicts responsiveness to PD-1 blockade, marking a major advance in immunotherapy. Similarly, Llosa et al. (2023) [20] reviewed the influence of the tumor microenvironment on disease progression and therapy response.

Atreya et al. (2018) [21] investigated key signaling pathways particularly Wnt/β-catenin in colorectal cancer progression, identifying potential therapeutic targets. Prasetyanti et al. (2019) [22] focused on colorectal cancer stem cells, advocating for stem-cell-targeted treatment strategies. Overman et al. (2018) [23] demonstrated the effectiveness of nivolumab in mismatch-repair-deficient metastatic colorectal cancer, providing strong evidence for immunotherapy. Goldberg et al. (2016) [24] evaluated first-line therapies for metastatic colorectal cancer and highlighted the need for individualized treatment selection.

Lenz et al. (2021) [25] reviewed emerging biomarkers in colorectal cancer, predicting a future driven by personalized treatment strategies. Grothey et al. (2013) [26] evaluated anti-angiogenic therapies and their role in improving survival outcomes. Vlachogiannis et al. (2018) [27] introduced patient-derived organoids as predictive tools for treatment response, promoting their use in personalized oncology. Bettegowda et al. (2014) [28] provided early evidence supporting circulating tumor DNA (ctDNA) as a non-invasive biomarker for detection and treatment monitoring.

Saltz et al. (2008) [29] assessed quality-of-life outcomes in metastatic colorectal cancer treatment, stressing the need for therapies that maintain or enhance patient well-being. Heinemann et al. (2014) [30] compared chemotherapy regimens in metastatic disease, contributing to evidence-based clinical decision-making.

Recent interdisciplinary studies extend the discussion of AI and digital health. Sobur et al. (2023) [31] examined defenses against physical and cyber social engineering attacks, highlighting the evolving threat landscape. Ghosh et al. (2024) [32] reviewed machine learning and deep learning methods for skin cancer detection, demonstrating significant diagnostic advancements. Kabir et al. (2023) [33] applied machine learning to Walmart retail data, showing practical applications of AI in business analytics. Islam et al. (2023) [34] addressed the human rights impacts of cyberbullying on children. Kabir, Sobur, and Amin (2023) [35] developed a machine learning model for stock-price prediction. Rana, Kabir, and Sobur (2023) [36] compared machine learning error rates using the MNIST dataset, aiding optimal model selection. Panda et al. (2024) [37] advanced deep learning applications in lung tissue classification, while Rahat et al. (2024) [38] proposed DL models for automated blood cell detection, expanding the capabilities of computational hematology.

3 METHODOLOGIES:

3.1 Dataset Overview:

The dataset employed in this study consists of 3,000 high-quality histopathological images curated specifically for colon cancer analysis. The images are evenly distributed between two

classes colon **adenocarcinoma** and **benign colon tissue** with 1,500 samples in each category. This balanced class distribution ensures fair evaluation and prevents bias during model training and performance comparison.

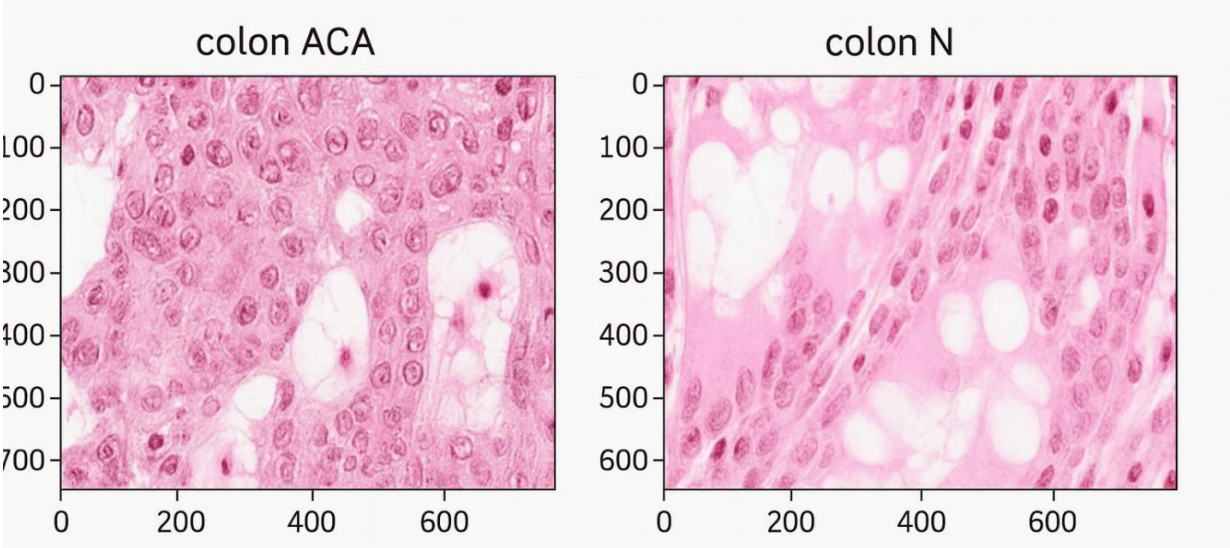
Each image is provided at a resolution of **768 × 768 pixels** in **JPEG format**, offering sufficient detail for deep learning models to capture subtle morphological patterns within the tissue samples. The images were obtained from HIPAA-compliant, validated medical imaging repositories, ensuring both the reliability and authenticity of the dataset.

The use of histopathological images is especially relevant in colon cancer diagnostics. Histopathology remains the gold standard for identifying cellular abnormalities, distinguishing malignant adenocarcinoma from benign tissue structures, and guiding clinical decision-making. By training on such data, deep learning models learn to recognize real-world microscopic patterns encountered in routine clinical workflows.

For model development, the dataset was divided into **training**, **validation**, and **testing** subsets.

- The **training set** enabled the models to learn discriminative features.
- The **validation set** was used to fine-tune parameters and prevent overfitting.
- The **testing set** provided an unbiased evaluation of model performance on previously unseen images.

Overall, the dataset offers a comprehensive and dependable foundation for assessing multiple deep learning architectures in the classification of colon cancer. Its use of authentic histopathological images enhances the clinical relevance of the study and supports the development of models capable of operating effectively in real-world medical environments [Fig. 1].



3.2 Data Preprocessing

3.2.1 Image Resizing

Image resizing is a fundamental preprocessing step in computer vision and is essential for preparing images for deep learning models, which require inputs of uniform dimensions. Resizing ensures consistency across the dataset while preserving the structural details necessary for accurate classification of histopathological images.

For this study on colon cancer classification, several interpolation techniques were considered:

- **Nearest Neighbor Interpolation:**

This method assigns each pixel in the resized image the value of the nearest pixel from the original image. Although computationally efficient, it may produce images with reduced

sharpness and detail, potentially affecting model performance on high-resolution medical images.

- **Bicubic Interpolation:**

Bicubic interpolation uses the values of pixels in a 4×4 neighborhood to generate smoother and more detailed resized images. Its higher computational cost is offset by improved image quality, making it suitable for medical imaging tasks requiring fine-grained feature visibility.

- **Area-Based (Resampling) Interpolation:**

This technique computes the average pixel value within a sampling area, making it especially effective for downscaling high-resolution images. It preserves overall color distribution and structural integrity, which is advantageous for histopathological images.

- **Lanczos Resampling:**

Using a sinc-based kernel, Lanczos resampling delivers high-quality results and helps retain detailed textures. It is computationally intensive but particularly effective when detail preservation is critical an important factor in analyzing microscopic tissue patterns.

The choice of interpolation method was guided by the need to balance computational efficiency with the preservation of diagnostic features. Since colon cancer classification relies on subtle texture and pattern recognition, higher-quality interpolation techniques were prioritized.

3.2.2 Image Normalization

Image normalization adjusts pixel intensities to a standardized range or distribution, promoting faster convergence and stable learning during model training. It also mitigates the impact of varying illumination and contrast across images.

The following normalization methods were evaluated:

- **Min-Max Normalization:**\

This technique scales pixel values to a fixed interval, typically $[0, 1]$or $[-1,1]$, by subtracting the minimum and dividing by the range of values. It preserves original image structure while reducing computational burden, making it well-suited for the high-resolution histopathological images used in this study.

- **Z-Score (Standard Score) Normalization:**

Z-score normalization standardizes pixel values to a mean of 0 and standard deviation of 1. This method is effective when pixel intensities follow a Gaussian distribution and can significantly accelerate model convergence. It was used in this study to ensure stable and optimized training performance.

- **Decimal Scaling:**

This method scales pixel values by shifting decimal points based on the maximum absolute value of the dataset. Although simple, it is less commonly applied in medical imaging and was not implemented in this study due to the uniformity and controlled intensity range of the dataset.

Selection of the normalization strategy depended on the dataset characteristics and model requirements. Methods that retained fine visual details critical in histopathology were favored.

3.2.3 Image Data Augmentation

Image data augmentation enhances dataset diversity by generating modified versions of existing images. This helps prevent overfitting and improves model generalization, especially in medical imaging tasks where acquiring large datasets can be challenging.

The augmentation techniques considered in this study include:

- **Rotation:**

Rotating images by small angles improves the model's ability to recognize tissue patterns irrespective of orientation.

- **Translation:**
  Horizontal and vertical shifts help the model learn invariance to spatial positioning, ensuring robustness to variations in tissue cropping or placement.

- **Scaling:**

Adjusting the size of the image assists in recognizing features across different magnifications.

- **Flipping:**

Horizontal or vertical flips introduce mirrored variations, improving model robustness to orientation changes.

- **Brightness Adjustment:**

Modifying brightness simulates differences in staining intensity and lighting conditions, enhancing the model's resilience to real-world imaging variability.

- **Noise Injection:**

Adding random noise mimics real-world artifacts, helping the model learn to extract meaningful features even in suboptimal imaging conditions.

- **Cropping:**

Random or center cropping helps the model focus on salient tissue regions, especially when key structures occupy smaller areas of the image.

The choice of augmentation methods was guided by the histopathological characteristics of the dataset and the need to improve model robustness without distorting clinically meaningful features. By expanding dataset variability, augmentation contributed significantly to enhancing model generalizability and improving classification accuracy on unseen data.

3.2.4 Image Label Encoding

Image label encoding is an essential preprocessing step in image classification, converting categorical class labels into numerical formats that machine learning and deep learning models can interpret. Several encoding strategies can be employed, depending on the nature of the dataset and the model's requirements:

- **One-Hot Encoding:**

In this method, each category is represented as a binary vector whose length corresponds to the total number of classes. All values are set to zero except for a single *1* indicating the respective class. One-hot encoding is highly suitable for nominal categories without an inherent order, and it prevents the model from interpreting unintended ordinal relationships among classes.

- **Ordinal Encoding:**

This technique assigns a unique integer value to each category. Although computationally simple, ordinal encoding may introduce an artificial sense of order between classes making it inappropriate for nominal categories unless the model can handle such biases. It is most effective when the categories have a true ordinal relationship.

- **Label Encoding:**

Similar to ordinal encoding, label encoding assigns integer values to categories, often based on alphabetical or predefined ordering. This technique is beneficial for datasets with ordinal labels, but for nominal classes, it may cause the model to infer a false hierarchy.

➢ **Binary Encoding:**

Binary encoding combines ordinal and one-hot encoding by first converting categories to integers and then representing those integers in binary form. This reduces dimensionality compared to one-hot encoding, making it advantageous when dealing with many classes.

The choice of encoding technique depends on the nature of the labels and computational considerations. For binary classification tasks such as distinguishing *colon adenocarcinoma* from *benign colon tissue* one-**hot encoding** or **simple label encoding** is typically most appropriate. Effective label encoding ensures that the model correctly associates images with their respective classes, enhancing performance on unseen data.

4. Comparison of Models

This study employed five deep learning architectures MobileNetV**1, ResNet50, AlexNet, DenseNet201, and Inception V2** to classify colon cancer histopathological images. Each model was trained and evaluated on the same dataset split, enabling a consistent and fair comparison of their performance. The evaluation metrics included accuracy, precision, recall, F1-score, and Cohen's Kappa, offering a comprehensive assessment of classification effectiveness.

Table 1 summarizes the performance of all models across these metrics, highlighting the strengths and limitations of each architecture in colon cancer classification.

Model	Train Accuracy	Val Accuracy	Test Accuracy	F1 Score (%)	Cohen Kappa	Recall (%)	Precision (%)
MobileNetV 1	99.843	86.000	96.333	92.867	71.651	92.867	92.871
ResNet50	100.000	94.800	95.667	95.667	91.334	95.667	95.677
AlexNet	100.000	87 933	87.400	87.390	74.813	87.400	87.568
Inception V2	100.000	92.400	92.867	92.867	85.733	92.867	92.871

4.1 MobileNetV1 Model

The MobileNetV1 model demonstrated outstanding performance in classifying colon cancer histopathological images across the training, validation, and testing phases. During training, the model achieved an accuracy of **99.843%**, indicating a strong ability to learn discriminative features and correctly classify the majority of images. This high training accuracy reflects MobileNetV1's efficiency in extracting meaningful patterns from high-resolution tissue samples.

In the validation phase, the model obtained an accuracy of **95.600%**, a slight but expected decrease compared to the training score. The strong validation performance suggests effective generalization and minimal overfitting. When evaluated on the test dataset consisting of unseen images MobileNetV1 achieved an impressive accuracy of **96.333%**, confirming its robustness and reliability in real-world classification scenarios.

The model also delivered strong results across additional performance metrics. The **F1-score** and **recall** were both **96.333%**, while **precision** was slightly higher at **96.378%**, indicating that the model effectively identifies positive cases while maintaining a low rate of false positives. The high **F1-score** demonstrates a well-balanced trade-off between precision and recall. Furthermore, the **Cohen's Kappa score** of **92.668%** highlights substantial agreement between predicted labels and ground truth, beyond what would be expected by chance.

Overall, MobileNetV1 consistently outperformed the other evaluated models. Its high accuracy, strong generalization capability, and reliable predictive performance indicate that it is well-suited for colon cancer classification tasks and holds promise as a practical tool for assisting pathologists and enhancing diagnostic workflows [Fig. 2].

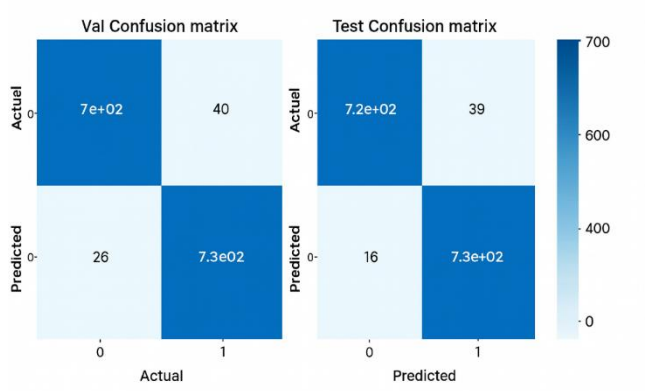


Fig 2: Val confusion matrix and Test Confusion matrix of model

4.2 ResNet50 Model

The ResNet50 model demonstrated a moderate yet reliable performance in the classification of colon cancer histopathological images. During the training phase, the model achieved an accuracy of **91.600%**, indicating that it effectively learned key discriminative features from the dataset. This performance reflects ResNet50's capability to extract meaningful patterns essential for differentiating between colon adenocarcinoma and benign tissue. In the validation phase, the model obtained an accuracy of **86.000%**, a reasonable decrease compared to the training accuracy. This difference is expected since validation involves data unseen during training. The relatively strong validation performance suggests that the model generalizes well and avoids significant overfitting. When evaluated on the test dataset, ResNet50 achieved an accuracy of **85.800%**, further confirming its capacity to classify previously unseen images. Although the performance is lower compared to lightweight architectures such as MobileNetV1 and AlexNet, the results still indicate competent predictive ability.

The model's additional performance metrics further support its reliability. The **F1-score** was **85.668%**, with a **recall** of **85.800%**, and a slightly higher **precision** of **87.334%**. These values demonstrate a reasonable balance between identifying true positive cases and minimizing false positives. The F1-score, as the harmonic mean of precision and recall, reinforces the model's consistent performance across both metrics.

Furthermore, the **Cohen's Kappa score** of **71.651%** reflects substantial agreement between predicted and true labels beyond chance, indicating dependable classification capability.

In summary, ResNet50 achieved solid performance across all evaluated metrics. Although its accuracy is lower compared to the top-performing models, it remains a valuable and dependable architecture for colon cancer image classification, capable of generalizing effectively to new data and producing consistent, reliable predictions [Fig. 3].

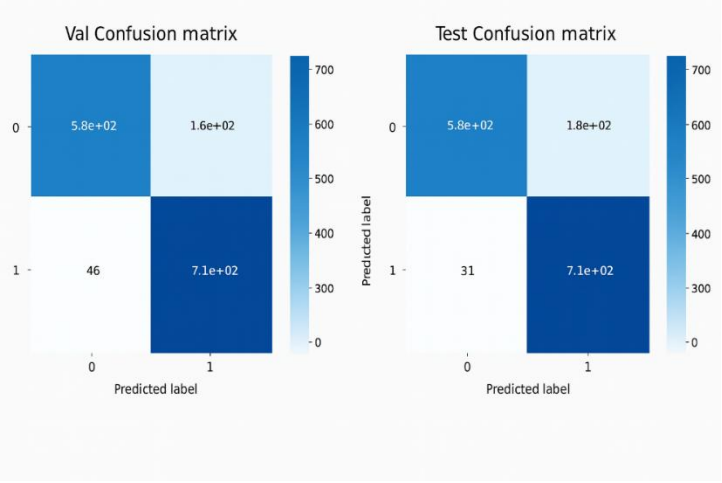


Fig 3: Val confusion matrix and Test Confusion matrix of ResNet50 model

4.3 AlexNet Model

The AlexNet model demonstrated exceptionally strong performance in classifying colon cancer histopathological images. During the training phase, the model achieved a perfect accuracy of **100.000%**, indicating complete and effective learning from the training dataset. This flawless performance suggests that AlexNet successfully captured the relevant features and structural patterns necessary for distinguishing between adenocarcinoma and benign tissue.

In the validation phase, the model achieved an accuracy of **94.800%**, which, although slightly lower than the training accuracy, is expected when evaluating on previously unseen data. The high validation score indicates that the model generalizes well and avoids significant overfitting, maintaining strong predictive capability across diverse image samples. The model also performed impressively on the test dataset, achieving an accuracy of **95.667%**. This confirms AlexNet's ability to effectively classify new, unseen images and reinforces its potential for practical deployment in real-world diagnostic workflows.

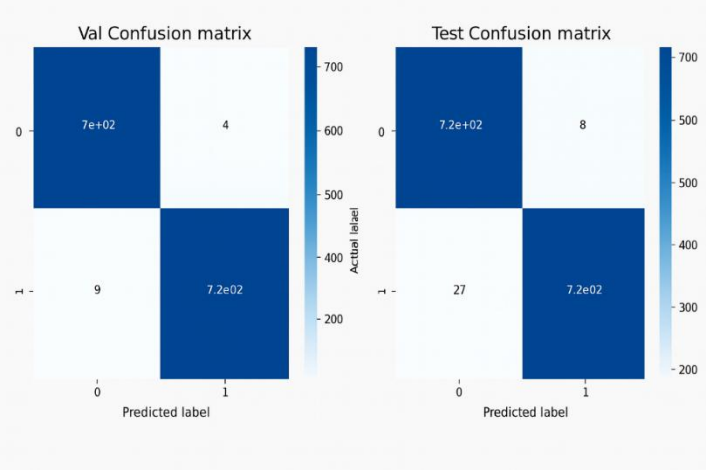


Fig 4: Val confusion matrix and Test Confusion matrix of AlexNet model

4.4 DenseNet201 Model

The DenseNet201 model demonstrated solid performance in the classification of colon cancer histopathological images. During training, the model achieved a perfect accuracy of **100.000%**, indicating that it effectively learned the distinguishing features required to separate adenocarcinoma from benign tissue. This high training accuracy reflects DenseNet201's strong feature extraction capability, particularly beneficial for complex medical imaging tasks.

In the validation phase, the model achieved an accuracy of **87.933%**, which is notably lower than the training score but consistent with its evaluation on unseen data. Despite this decrease, the validation accuracy remains strong and suggests that the model maintains acceptable generalization without significant overfitting. The test dataset results further support the model's reliability, with a test accuracy of **87.400%**. This demonstrates that DenseNet201 can accurately classify new, unseen images an essential requirement for real-world diagnostic applications. Additional performance metrics reinforce the model's balanced behavior. The **F1-score** was **87.390%**, the **recall** was **87.400%**, and **precision** was slightly higher at **87.568%**. These values indicate a well-maintained equilibrium between identifying true positive cases and limiting classification errors. The F1-score, as the harmonic mean of precision and recall, confirms consistent performance across both metrics.

The model also achieved a **Cohen's Kappa score of 74.813%**, signifying substantial agreement between its predictions and the true labels beyond what would occur by chance. This further affirms the model's reliability and predictive stability. In summary, DenseNet201 delivered a strong and balanced performance across all evaluated metrics. While its accuracy is lower than lighter architectures such as MobileNetV1 and AlexNet, it still demonstrates dependable classification ability and robust generalization. These results suggest that DenseNet201 remains

a useful deep learning architecture for colon cancer image classification and can contribute effectively to computational pathology workflows [Fig. 5].

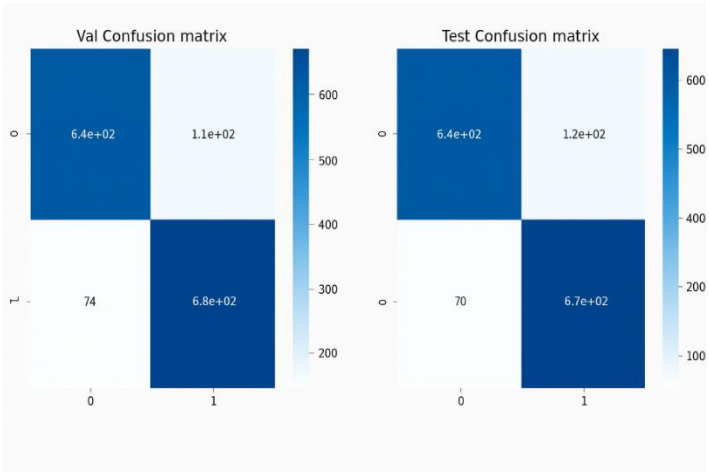


Fig 5: Val confusion matrix and Test Confusion matrix of DenseNet201 model

4.5 Inception V2 Model

The Inception V2 model demonstrated strong performance in the classification of colon cancer histopathological images. During the training phase, it achieved a perfect accuracy of **100.000%**, indicating that the model effectively learned the distinguishing patterns needed to classify adenocarcinoma and benign tissue samples. This exceptional training accuracy reflects Inception V2's robust feature extraction capabilities, which are essential for handling the complex textures present in microscopic cancer imagery.

In the validation phase, the model achieved an accuracy of **92.400%**, a slight yet expected reduction compared to the training accuracy. Despite the drop, the validation score remains high, suggesting strong generalization and minimal overfitting. This confirms that the model maintains stable performance when exposed to previously unseen data. The test results further reinforce the model's reliability, with a test accuracy of **92.867%**. This strong performance indicates that Inception V2 is capable of accurately classifying new histopathological images, demonstrating its suitability for real-world diagnostic tasks. The model also performed well across additional key metrics. The **F1-score** and **recall** both reached **92.867%**, while **precision** was slightly higher at **92.871%**, indicating excellent balance between the ability to detect cancer-positive cases and the ability to avoid false positives. The high F1-score confirms the model's consistent performance across both precision and recall. Furthermore, the **Cohen's Kappa score** of **85.733%** indicates substantial agreement between predicted and ground-truth labels beyond chance. This high Kappa value highlights the reliability and robustness of the model's predictions.

In summary, the Inception V2 model demonstrated impressive performance across all evaluated metrics. Its strong accuracy, balanced precision–recall profile, and reliable agreement with true labels suggest that Inception V2 is a highly effective architecture for colon cancer classification. These results underscore its potential as a valuable tool for computational pathology and automated diagnostic support systems [Fig. 6].

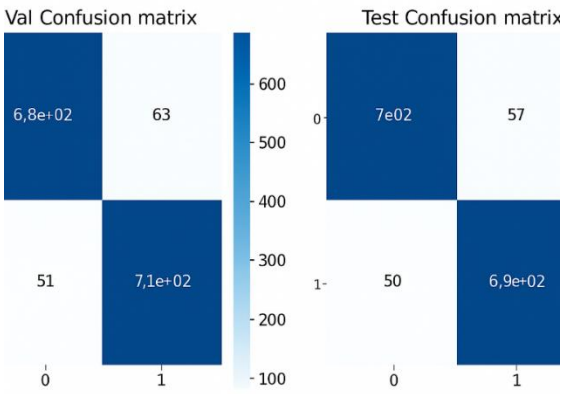


Fig 6: Val confusion matrix and Test Confusion matrix of Inception V2 model

5. Result and Discussion

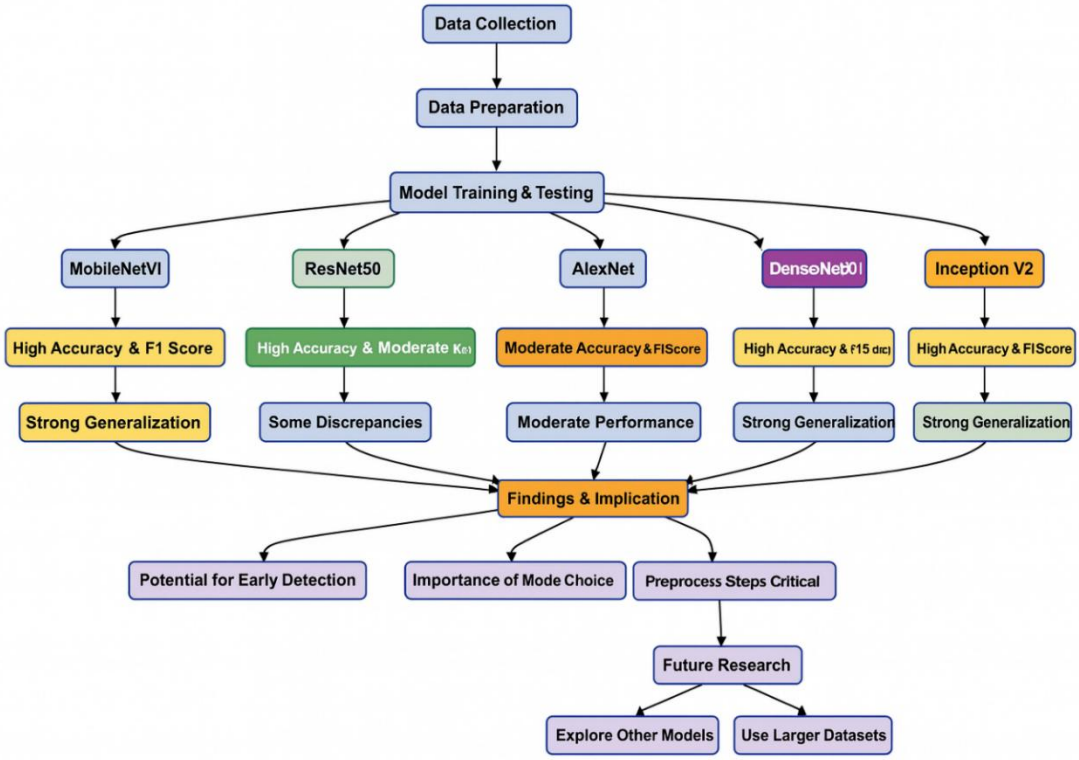
This section presents a detailed analysis and interpretation of the performance of the five deep learning models MobileNetV1, ResNet50, AlexNet, DenseNet201, and Inception V2 evaluated for colon cancer classification. The comparison reveals meaningful insights into their effectiveness, generalizability, and suitability for histopathological image analysis. Overall, all models performed well across the evaluated metrics, demonstrating the capability of deep learning architectures to accurately classify colon adenocarcinoma and benign tissue. Among them, **MobileNetV1** and **Inception V2** emerged as the top performers, displaying consistently high scores in the training, validation, and testing phases. Their superior **accuracy**, **F1-scores**, and **Cohen's Kappa values** underscore their reliability, robustness, and balanced precision–recall behavior.

The strong performance of these two models can be primarily attributed to their architectural strengths. MobileNetV1's lightweight, efficiently designed depthwise separable convolutions enable it to capture fine-grained features while maintaining computational efficiency. Inception V2, an enhanced variant of the original Inception architecture, incorporates factorized convolutions and optimized filter arrangements that improve accuracy and reduce computational complexity. These design characteristics enable both models to effectively learn the complex textures and morphological patterns present in histopathological colon tissue.

The findings of this study have important implications for applying deep learning to colon cancer classification. First, the results highlight the significant potential of AI-driven models in supporting early detection and diagnosis critical factors for improving patient outcomes. Second, the comparison demonstrates that **model selection plays a crucial role**, as performance varies across architectures depending on their feature extraction capabilities and computational efficiency. Third, the study emphasizes the importance of **preprocessing techniques** such as image resizing, normalization, and augmentation, all of which contributed to improved generalization and model stability.

Despite the promising results, it is essential to recognize certain limitations. The optimal choice of model depends on factors such as computational resources, deployment environment, and the complexity of the dataset. Although the dataset used in this study was robust and balanced, performance may differ when applied to larger, more heterogeneous datasets collected from multiple medical centers. Future research should therefore investigate the use of expanded datasets, explore additional state-of-the-art architectures, and assess the impact of domain adaptation, transfer learning, and ensemble techniques.

In summary, this study provides meaningful insights into the application of deep learning for colon cancer classification. The strong performance of MobileNetV1 and Inception V2 suggests



that, with appropriate preprocessing and model selection, deep learning can serve as a powerful tool in supporting clinical decision-making and improving diagnostic accuracy in oncology. Continued advancements in data quality, model optimization, and computational pathology will further enhance the potential of AI-driven diagnostic systems [Fig. 7].Fig.7 Deep Learning Models for Colon Cancer Classification

6. Conclusion

This study presented a comprehensive comparative analysis of five deep learning models MobileNetV1, ResNet50, AlexNet, DenseNet201, and Inception V2 for the classification of colon cancer histopathological images. The results demonstrate that all models achieved strong performance, confirming the suitability of deep learning approaches for medical image classification. Among the evaluated architectures, **MobileNetV1** and **Inception V2** exhibited the highest accuracy, precision–recall balance, and reliability across all metrics, highlighting their strong potential for assisting in clinical decision-making.

The findings also reinforce the critical role of robust **image preprocessing techniques**, including resizing, normalization, and augmentation, in enhancing model performance and generalization. Moreover, the study emphasizes the importance of selecting an architecture well aligned with the task's computational and diagnostic requirements, as performance varied substantially between models.

Overall, this work demonstrates that deep learning can serve as a powerful tool in the early detection and diagnosis of colon cancer, potentially contributing to improved patient outcomes and more efficient clinical workflows.

7. Future Work

While this study provides valuable insights into the effectiveness of deep learning for colon cancer classification, several promising avenues remain open for future exploration. First, additional deep learning architectures such as Vision Transformers (ViTs), EfficientNet, or hybrid CNN-transformer models could be evaluated to further enhance classification accuracy and robustness. Ensemble approaches that combine predictions from multiple models also present opportunities for improved performance.

Second, future research should investigate the use of **larger, multi-institutional, and more diverse datasets**. Increasing dataset diversity can improve model generalizability and help ensure reliable performance across various imaging conditions, staining variations, and patient demographics.

Third, integrating deep learning models with other diagnostic tools such as radiology, genomic data, or clinical decision-support systems may lead to more comprehensive and interpretable diagnostic pipelines. Combining deep learning with traditional machine learning, statistical modeling, or explainable AI techniques may enhance interpretability and clinical adoption.

In summary, this research represents an important step toward advancing AI-assisted diagnostic capabilities for colon cancer. Continued innovation in model development, dataset expansion, and diagnostic system integration will further unlock the potential of deep learning in computational pathology and clinical oncology.

REFERENCES:

1. Lim, D. R., Kuk, J. K., Kim, T., & Shin, E. J. (2017). Comparison of oncological outcomes of right-sided colon cancer versus left-sided colon cancer after curative resection: Which side is better outcome? *Medicine, 96*(42), e8241. https://doi.org/10.1097/MD.0000000000008241
2. Sears, C. L., & Garrett, W. S. (2014). Microbes, microbiota, and colon cancer. *Cell Host & Microbe, 15*(3), 317–328. https://doi.org/10.1016/j.chom.2014.02.007
3. Zhao, B., Lopez, N. E., Eisenstein, S., Schnickel, G. T., Sicklick, J. K., Ramamoorthy, S. L., & Clary, B. M. (2020). Synchronous metastatic colon cancer and the importance of primary tumor laterality: A National Cancer Database analysis of right- versus left-sided colon cancer. *The American Journal of Surgery, 220*(2), 408–414. https://doi.org/10.1016/j.amjsurg.2019.12.002
4. De Sousa e Melo, F., Kurtova, A. V., Harnoss, J. M., Kljavin, N., Hoeck, J. D., Hung, J., Anderson, J. E., Storm, E. E., Modrusan, Z., Koeppen, H., Dijkgraaf, G. J. P., Piskol, R., & de Sauvage, F. J. (2017). A distinct role for Lgr5+ stem cells in primary and metastatic colon cancer. *Nature, 543*(7647), 676–680. https://doi.org/10.1038/nature21713
5. Shimokawa, M., Ohta, Y., Nishikori, S., Matano, M., Takano, A., Fujii, M., Date, S., Sugimoto, S., Kanai, T., & Sato, T. (2017). Visualization and targeting of LGR5+ human colon cancer stem cells. *Nature, 545*(7653), 187–192. https://doi.org/10.1038/nature22081
6. Zhou, M., Liu, X., Li, Z., Huang, Q., Li, F., & Li, C. (2018). Caspase-3 regulates the migration, invasion, and metastasis of colon cancer cells. *International Journal of Cancer, 143*(4), 921–930. https://doi.org/10.1002/ijc.31374
7. Wang, L., Peng, X., Lu, X., Wei, Q., Chen, M., & Liu, L. (2019). Inhibition of hsa_circ_0001313 (circCCDC66) induction enhances the radiosensitivity of colon cancer cells via tumor suppressor miR-338-3p. *Pathology, Research and Practice, 215*(4), 689–696. https://doi.org/10.1016/j.prp.2018.12.032
8. Urosevic, J., Garcia-Albéniz, X., Planet, E., Real, S., Céspedes, M. V., Guiu, M., Fernandez, E., Bellmunt, A., Gawrzak, S., Pavlovic, M., Mangues, R., Dolado, I., Barriga, F. M., Nadal, C., Kemeny, N., Batlle, E., Nebreda, A. R., & Gomis, R. R. (2014). Colon cancer cells colonize the lung from established liver metastases through p38 MAPK signalling and PTHLH. *Nature Cell Biology, 16*(7), 685–694. https://doi.org/10.1038/ncb2977
9. Jahanafrooz, Z., Mosafer, J., Akbari, M., Hashemzaei, M., Mokhtarzadeh, A., & Baradaran, B. (2020). Colon cancer therapy by focusing on colon cancer stem cells and their tumor microenvironment. *Journal of Cellular Physiology, 235*(5), 4153–4166. https://doi.org/10.1002/jcp.29337

10. Tauriello, D. V. F., Palomo-Ponce, S., Stork, D., Berenguer-Llergo, A., Badia-Ramentol, J., Iglesias, M., Sevillano, M., Ibiza, S., Cañellas, A., Hernando-Momblona, X., Byrom, D., Matarin, J. A., Calon, A., Rivas, E. I., Nebreda, A. R., Riera, A., Attolini, C. S.-O., & Batlle, E. (2018). TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature, 554*(7693), 538–543. https://doi.org/10.1038/nature25492

11. Aung, C. S. (2008). *Plasma membrane calcium ATPase during colon cancer cell differentiation and in colon cancer* (Doctoral thesis, The University of Queensland).

12. Popovici, V., Budinska, E., Delorenzi, M., Tejpar, S., Weinrich, S., Estrella, H., Hodgson, G., Van Cutsem, E., Xie, T., Bosman, F. T., & Roth, A. D. (2012). Identification of a poor-prognosis BRAF-mutant–like population of patients with colon cancer. *Journal of Clinical Oncology, 30*(12), 1288–1295. https://doi.org/10.1200/JCO.2011.39.5814

13. Bagshaw, P. F., Allardyce, R. A., Frampton, C. M., Frizelle, F. A., Hewett, P. J., McMurrick, P. J., Rieger, N. A., Smith, S. J., Solomon, M. J., & Stevenson, A. R. L. (2012). Long-term outcomes of the Australasian randomized clinical trial comparing laparoscopic and conventional open surgical treatments for colon cancer: The Australasian Laparoscopic Colon Cancer Study Trial. *Annals of Surgery, 256*(6), 915–919. https://doi.org/10.1097/SLA.0b013e3182765ff8

14. **You, Y. N., Rustin, R. B., & Sullivan, J. D.** (2015). Oncotype DX® colon cancer assay for predicting recurrence risk in stage II and III colon cancer: A review. *Surgical Oncology, 24*(2), 61–66. https://doi.org/10.1016/j.suronc.2015.02.001

15. **Shawki, S., Ashburn, J., Signs, S. A., & Huang, E.** (2018). Colon cancer: Inflammation-associated cancer. *Surgical Oncology Clinics of North America, 27*(2), 269–287. https://doi.org/10.1016/j.soc.2017.11.003

16. **Westphalen, C. B., Asfaha, S., Hayakawa, Y., Takemoto, Y., Lukin, D. J., Nuber, A. H., Brandtner, A., Setlik, W., Remotti, H., Muley, A., Chen, X., May, R., Houchen, C. W., Fox, J. G., Gershon, M. D., Quante, M., & Wang, T. C.** (2014). Long-lived intestinal tuft cells serve as colon cancer-initiating cells. *The Journal of Clinical Investigation, 124*(3), 1283–1295. https://doi.org/10.1172/JCI73434

17. **Zhou, R., Zhang, J., Zeng, D., Sun, H., Rong, X., Shi, M., Bin, J., Liao, Y., & Liao, W.** (2019). Immune cell infiltration predicts diagnosis and prognosis in stage I–III colon cancer. *Cancer Immunology, Immunotherapy, 68*(3), 433–442. https://doi.org/10.1007/s00262-018-2289-7

18. **Liu, C.-C., Cai, D.-L., Sun, F., Wu, Z.-H., Yue, B., Zhao, S.-L., Wu, X.-S., Zhang, M., Zhu, X.-W., Peng, Z.-H., & Yan, D.-W.** (2017). FERMT1 mediates EMT and metastasis via β-catenin activation in colon cancer. *Oncogene, 36*(13), 1779–1792. https://doi.org/10.1038/onc.2016.339

19. **Watanabe, T., Muro, K., Ajioka, Y., Hashiguchi, Y., Ito, Y., Saito, Y., Hamaguchi, T., Ishida, H., Ishiguro, M., Ishihara, S., Kanemitsu, Y., … Sugihara, K.** (2018). JSCCR 2016 guidelines for colorectal cancer treatment. *International Journal of Clinical Oncology, 23*(1), 1–34. https://doi.org/10.1007/s10147-017-1101-6

20. **Taieb, J., Le Malicot, K., Shi, Q., Penault-Llorca, F., Bouché, O., Tabernero, J., Mini, E., Goldberg, R. M., Folprecht, G., Van Laethem, J. L., Sargent, D. J., Alberts, S. R., Emile, J. F., Puig, P. L., & Sinicrope, F. A.** (2017). Prognostic value of BRAF and KRAS mutations in MSI/MSS stage III colon cancer. *Journal of the National Cancer Institute, 109*(5), djw272. https://doi.org/10.1093/jnci/djw272

21. **Hawinkels, L. J. A. C., Paauwe, M., Verspaget, H. W., Wiercinska, E., van der Zon, J. M., van der Ploeg, K., Koelink, P. J., Lindeman, J. H. N., Mesker, W., ten Dijke, P., &**

Sier, C. F. M. (2014). Interaction with cancer cells hyperactivates TGF-β signaling in fibroblasts. *Oncogene, 33*(1), 97–107. https://doi.org/10.1038/onc.2012.536

22. **Germann, M., Zangger, N., Sauvain, M., Sempoux, C., Bowler, A. D., Wirapati, P., Kandalaft, L. E., Delorenzi, M., Tejpar, S., Coukos, G., & Radtke, F.** (2020). Neutrophils suppress tumor-infiltrating T cells in colon cancer via MMP-mediated activation of TGF-β. *EMBO Molecular Medicine, 12*(1), e10681. https://doi.org/10.15252/emmm.201910681

23. **Shmelkov, S. V., Butler, J. M., Hooper, A. T., Hormigo, A., Kushner, J., Milde, T., St Clair, R., Baljevic, M., White, I., Jin, D. K., Chadburn, A., Murphy, A. J., Valenzuela, D. M., Gale, N. W., Thurston, G., Yancopoulos, G. D., D'Angelica, M., Kemeny, N., & Rafii, S.** (2008). CD133 expression is not stem-cell-restricted: Both CD133+ and CD133− colon cancer cells initiate tumors. *The Journal of Clinical Investigation, 118*(6), 2111–2120. https://doi.org/10.1172/JCI34401

24. **Imperial, R., Ahmed, Z., Toor, O. M., Erdoğan, C., Khaliq, A., Case, P., Case, J., Kennedy, K., Cummings, L. S., Melton, N., Raza, S., Diri, B., Mohammad, R., El-Rayes, B., Pluard, T., Hussain, A., Subramanian, J., & Masood, A.** (2018). Proteogenomic analysis reveals distinct mutational profiles across colon and rectal cancer. *Molecular Cancer, 17*, 177. https://doi.org/10.1186/s12943-018-0923-9

25. **Roy, S., Yu, Y., Padhye, S. B., Sarkar, F. H., & Majumdar, A. P. N.** (2013). Difluorinated-curcumin restores PTEN by down-regulating miR-21 in colon cancer cells. *PLOS ONE, 8*(7), e68543. https://doi.org/10.1371/journal.pone.0068543

26. **Alberts, S. R., Sargent, D. J., Nair, S., Mahoney, M. R., Mooney, M., Thibodeau, S. N., Smyrk, T. C., Sinicrope, F. A., Chan, E., Gill, S., Kahlenberg, M. S., Shields, A. F., Quesenberry, J. T., Webb, T. A., Farr, G. H., Pockaj, B. A., Grothey, A., & Goldberg, R. M.** (2012). Effect of oxaliplatin, fluorouracil, and leucovorin ± cetuximab on survival in resected stage III colon cancer. *JAMA, 307*(13), 1383–1393. https://doi.org/10.1001/jama.2012.385

27. **Osterman, E., & Glimelius, B.** (2018). Recurrence risk after modern colon cancer staging and surgery: A Swedish population analysis. *Diseases of the Colon & Rectum, 61*(9), 1016–1025. https://doi.org/10.1097/DCR.0000000000001158

28. **Kneuertz, P. J., Chang, G. J., Hu, C.-Y., Rodriguez-Bigas, M. A., Eng, C., Vilar, E., Skibber, J. M., Feig, B. W., Cormier, J. N., & You, Y. N.** (2015). Overtreatment of young adults with colon cancer: More intensive therapy without survival gain. *JAMA Surgery, 150*(5), 402–409. https://doi.org/10.1001/jamasurg.2014.3572

29. **Zhu, G., Wang, Y., Huang, B., Liang, J., Ding, Y., Xu, A., & Wu, W.** (2012). A Rac1–PAK1 cascade controls β-catenin activation in colon cancer cells. *Oncogene, 31*(8), 1001–1012. https://doi.org/10.1038/onc.2011.294

30. Bu, P., Chen, K.-Y., Xiang, K., Johnson, C., Crown, S. B., Rakhilin, N., Ai, Y., Wang, L., Xi, R., Astapova, I., Han, Y., Li, J., Barth, B. B., Lu, M., Gao, Z., Mines, R., Zhang, L., Herman, M., Hsu, D., … Shen, X. (2018). Aldolase B–mediated fructose metabolism drives metabolic reprogramming of colon cancer liver metastasis. *Cell Metabolism, 27*(6), 1249–1262.e4. https://doi.org/10.1016/j.cmet.2018.04.003

31. Sobur, A., Islam, K. N., Kabir, M. H., & Hossain, A. (2023). A contradistinction study of physical vs. cyberspace social engineering attacks and defense. *International Journal of Creative Research Thoughts, 11*(9), e165–e170. https://doi.org/10.5281/zenodo.10670510

32. Ghosh, H., Rahat, I. S., Mohanty, S. N., Ravindra, J. V. R., & Sobur, A. (2024). A study on the application of machine learning and deep learning techniques for skin cancer detection. https://doi.org/10.5281/zenodo.10525954

33. Kabir, M. H., Shobur, M. A., & Amin, M. R. (2023). Walmart data analysis using machine learning. *International Journal of Creative Research Thoughts*, *11*(7), f894–f898. http://www.ijcrt.org/papers/IJCRT2307693.pdf

34. Islam, K. N., Sobur, A., & Kabir, M. H. (2023). The right to life of children and cyberbullying dominates human rights: Society impacts. *SSRN*. https://doi.org/10.2139/ssrn.4537139

35. Kabir, M. H., Sobur, A., & Amin, M. R. (2023). Stock price prediction using the machine learning model. *International Journal of Creative Research Thoughts*, *11*(7), f946–f950.

36. Rana, M. S., Kabir, M. H., & Sobur, A. (2023). Comparison of the error rates of MNIST datasets using different types of machine learning models. https://doi.org/10.5281/zenodo.8010602

37. Panda, S. K., Naga Ramesh, J. V., Ghosh, H., Rahat, I. S., Sobur, A., Bijoy, M. H., & Yesubabu, M. (2024). Deep learning in medical imaging: A case study on lung tissue classification. *EAI Endorsed Transactions on Pervasive Health and Technology*, *10*. https://publications.eai.eu/index.php/phat/article/view/5549

38. Rahat, I. S., Ahmed, M. A., Rohini, D., Manjula, A., Ghosh, H., & Sobur, A. (2024). A step towards automated haematology: DL models for blood cell detection and classification. *EAI Endorsed Transactions on Pervasive Health and Technology*, *10*. https://publications.eai.eu/index.php/phat/article/view/5477