https://doi.org/ 10.33472/AFJBS.6.Si2.2024.1445-1453

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

CANCER CELL SEGMENTATION AND DETECTION USING MACHINE LEARNING

M. Rudra Kumar, Professor,
Dept of IT, MGIT, mrudrakumar@gmail.com
G. Ananya Saravani, UG Student, MGIT gananyasaravani_csb203220@mgit.ac.in
Dinesh Prabhath Repalle, UG Student dprabhathrepalle_csb203218@mgit.ac.in
S. Krish Raj, UG Student, MGIT skrishraj_csb203251@mgit.ac.in

Article History

Volume 6,Issue Si2, 2024

Received:27 Mar 2024

Accepted : 28 Apr 2024

doi: 10.33472/AFJBS.6.Si2.2024.1445-1453

***Abstract**:*

The combination of computer science, medical imaging technology and biology has turned cancer research upside down. This collaborative project, which integrates expertise from many fields and uses advanced methods such as deep learning for accurate pixel-wise segmentation of cell images to allow other researchers working in cancer research to carry out precise identification at the level of individual cells. These subtle algorithms reveal how cancer cells behave, what form they take and the ways in which they interact--in particular giving us a lot of information on cancer at its most microscopic. Thanks to cutting-edge segmentation, researchers can determine how to treat cancer cells depending on their specific characteristics separately. This not only improves diagnostic accuracy and treatment efficacy but also better suits monitoring patients diagnosed with this disease. The synergy is multidisciplinary; the accumulation of scientific understanding will also help transform cancer diagnosis, treatment strategies and therapeutic monitoring to effect better patient outcomes.

Keywords: Cancer cell segmentation, Deep learning, Convolutional Neural Networks (CNN), Transfer learning, Flask web interface, Lung cancer, Breast cancer, Skin cancer, Medical imaging, Cell tracking, Predictive modeling, Multi-modal data fusion, Real-time decision support systems, Validation studies, Clinical trials, MobileNet, ResNetV2, DenseNet169,

1. INTRODUCTION

On the one hand, cancer is a complex and life-threatening disease which presents great difficulties not only in diagnosis but also treatment. Furthermore, it's little understood how this lethal condition develops. Cancer cell segmentation is a difficult process, but it plays an important role in analysing cancer cells' behaviour. This helps to improve diagnosis and the planning of treatment oncology research experts throughout the world. This project undertakes construction of an integrated system, which combines state-of-the art deep learning methods and a user interface enabling web-based interaction to help physicians segment cancer cells (lung, breast), the tracking as well. This project's central goal is to develop a powerful platform based on Convolutional Neural Network (CNN) and transfer learning, with an accessible web interface built using Flask. Each trained CNN is accurate in identifying and tracing cancer cells within medical images. Transfer learning shortens the learning process even further, so that effective cell identification can be achieved with relatively few datasets available for training. The result is a Flask-based interface that provides medical professionals and researchers with an easy to use platform upon which they can upload images, receive segmentation assistance and also track the movements of important contribution to progress on all fronts related with cancer research Ultimately, the project's aim is to provide an effective solution that can be easily applied in cancer research and can continue contributing its valuable contributions toward preventing or eradicating cancer. This would further advance current work on halting populations of cancer cells from growing through wider insights into what types of treatment are likely most successful for different kinds of patients suffering any given form a particular type of malignant.

By integrating the most advanced technology with deep-learning algorithms and user interface design, this project has a vision to make contributions in many outlets possible within oncology science; from helping increase diagnostic sensitivity, through designing personalized treatment strategies for cancer patients complete upthrougher reshaping our understanding of how these cells divide.

2. LITERATURE SURVEY

Analysing and classifying of the cancer cells in medical images is incredibly important finding cancer exactly and therefore make the right choices for efficient treatment, for example if it is lung cancer, breast cancer, or skin cancer. During the last couple of years, scientists concentrate on the implementation of the machine learning systems and deep learning algorithm designed for the analysis of the computer tomography scanning (CT) images and cancer cells in real time.

The combination of these constituent parts is intended to be a complete tool for studying cancer cell behaviour, helping in early diagnoses and assessing treatment options until even after the threat has been vanquished it may end up making an

microscopic sections. The goal is to carry out the job of spotting and differentiating cancer cells, we must achieve the effective and precise process.

Here, however, we touch upon the relevant papers on the most recent work pertaining to the detection and classification of cancer. It is an examination of varied approaches put forward on different evolutionary theories, where each one is critically assessed to examine their individual strengths and drawbacks. Therefore, our ultimate goal is to bring to light the essential aspects of clinical research progress and discoveries.

The paper covers the utilization of machine vision technology in the assortment of cancerous cells through microscopic pictures, an outlook to the traditional and its remodelling. This paper deals with overcoming those limitations associated with the traditional methods of diagnosis which include subjectiveness, resource competitions and dependency on the subjective assessments of pathologists. Such machine vision technologies have the capacity of providing the potentials of offering an aggregate improvement of the latter. It sequentially describes the machine vision detection system model with great clarity that includes those image preprocessing techniques like denoising, enhancement and colour normalization whose main role is to make good the edge of images. In addition, it presents the ways with which we can segment our cells for we can identify crucial contours, it also gives the methods for feature extraction covering the textures, shapes, and colour attributes, and also identifies the way to improve recognition outcomes by combining multiple features. Conclusively, the paper outlines the bottlenecks and future opportunities in machine vision for histopathologic image detection, which serve as a crucial step for cancer screening progression in the field of innovations. [1], [2], [3].

This research article describes the use of machine learning and imaging techniques to improve melanoma diagnosis. Knowing the difficulty of distinguishing benign and malignant diseases, different methods were used in the study. This method uses dermoscopic images from the ISIC dataset, starting with careful preprocessing to improve image quality. OTSU thresholding and other segmentation techniques separate organisms for further feature extraction (such as texture, shape, and colour). To solve the class bias problem, Synthetic Minority Oversampling Technique (SMOTE) is used, which creates a good balance between good and bad samples. This study improves the use of packaging methods in feature selection and reveals important features that enable accurate classification. This is a new way to help dermatologists detect skin cancer at an early stage, providing patients with earlier diagnosis and better outcomes [2], [4], [5].

The paper discusses the conceptualization and integration of an automatic computer technology for lung cancer detection in lung MRI images. This is evidence that early detection is the only way to provide timely and proper treatment that will stop the disease from worsening. The potential method is intending to simplify the identification and in specific the using MRI images and its still a challenge because of the complex cell structure. The study presents two methods for detection: classic image processing and advanced Convolutional Neural Network (CNN)-driven classification method, respectively. Besides, the paper outlines the Adaptive Hierarchical Heuristic Mathematical Model (AHHMM) as a multi-level approach by adjusting image segmentation, deep learning, as well as probability modeling for lung cancer prediction in CT scans and MRI. This strategy is designed to increase precision and correctness of lung cancer identification; potentially improving automated diagnostic tools for better patients care [3], [6], [7].

Cellular segmentation and tracking in microscopy images is a new approach. The main goal of this method was to tackle the problem of removing phase touching cells from microscopy images with low signal-to-noise ratios. Instead, cell borders are represented as distance maps. Thus, it makes training possible to learn both touching and adjacent cells. In order to achieve this, a U-Net network using dual-decoder pathway modified for predicting neighbour distances was used showing promising results despite challenges posed by cell segmentation over different types of microscopy.

Additionally, the paper describes Graph-based Cell Tracking with motion estimation module that solves the problem of missing segmentation masks across few frames for tracking purposes. In conclusion, there is a quantitative comparison and evaluation results provided by the paper which show that the proposed technique outperforms other methods in terms of handling merged cells,

annotation inconsistencies and capability to deal with different cell types when having limited training data set up or number available to it. This work asserts oneself on robust strategies in biomedical image analysis by presenting an application example for system. [4], [8], [9], [10]
Utilizing techniques like image preprocessing, segmentation, feature extraction, and machine learning, several of them reported solutions that are appropriate for the many types of cancer and imaging modalities. The shortcomings of both platforms designed for the diagnosis of various cancer kinds and imaging techniques, as well as traditional methods of detection, are highlighted.

3. METHODOLOGY

3.1 Proposed Work:

The proposed system tackles challenges which are closely related to lung, breast and skin cancer diagnosis provides a complete approach that is aimed at cell segmentation and detection. The designed algorithms (MobileNet , ResNet and DenseNet) that are reflective training enable spot-on analysis of the images. It's VR effectiveness lies in that it can adjust to various tissues shapes, it is able to integrate seamlessly to different imaging modes, and a simple user interface, thus. MobileNet has a computational advantage wrt dermatoscopic images, ResNet is superior in deep data generation concerning lung histopathology and DenseNet helps in efficient feature use for different types of breast cancer. The unified approach enables models to work together and data to flow smoothly without interruption every now and then or entry into patient care workflow.

3.2 System Architecture:

The architecture for the segmentation and detection of cancer cells in the system consists of three main parts. At first, an input block uses image correction techniques to make it possible to apply photos of lung, breast and skin cancer obtained from image sets. Subsequently, deep learning models such as MobileNet , ResNet and DenseNet were implemented using Jupyter Notebook for cell segmentation and tracking. This entailed designing segmentations models and integrating methods for detecting cell Lastly, the system will be Flask-based which is a framework that enables web development to facilitate user interaction with live inputs. The interface is connected with backend modules that process user inputs using the trained models to display segmented cells and track results. Fundamentally, this architecture comprehensively combines data handling, model development, user interactions through GUIs and presentation of the analytics on cancer cell monitoring.

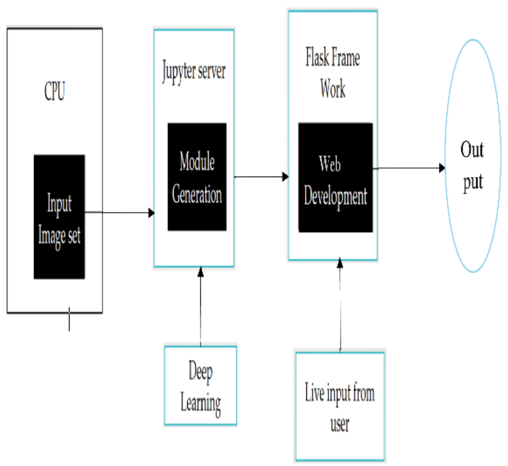

Fig 3.2: Proposed System Architecture

3.3 Dataset:

The dataset used for the project is made up of multiple sources to achieve comprehensive cancer types coverage, including lung, breast and skin cancers. Specifically:

Lung Cancer Dataset: The data set obtained from independent repositories such as Lung Image Database Consortium (LIDC) and Cancer Imaging Archive (TCIA) comprises high resolution CT scans showing different stages and types of lung cancer like adenocarcinoma, squamous cell carcinoma, small cell carcinoma among others.

Breast Cancer Dataset: This dataset was individually collected from healthcare centres and research institutions presenting mammographic pictures and histopathological slides of various forms of breast cancer such as DCIS, IDC, lobular carcinoma [11].

Skin Cancer Dataset: Open-access datasets that include a variety of melanoma and non-melanoma skin cancers consisting dermoscopy images and histopathological sections have been obtained from ISICs and dermatology research databases including basal cell carcinoma (BCC), squamous cell carcinoma (SCC).

For each instance there are ground truth masks which can be used as references when training deep learning models or validating them.

3.4 Image Processing:

To prepare the medical images for subsequent analysis, a series of preprocessing steps are applied:

Noise Reduction: Using the tools like Gaussian and median filtering for denoising medical imaging modalities so that noise artifacts are reduced resulting in smoother and cleaner images for soft as well as hard tissues segmentation.

Contrast Enhancement: The histogram equalization or adaptive histogram equalization acts as a tool for enhancing the visibility of subtle anatomical features and pathological structures, thereby creating a higher discrimination power of the segmentation models.

Normalization: Model standardization of image intensity distributions across different modalities and acquisition conditions to homogenize the influence and to ensure the consistency throughout the model training, optimization, and inference [12].

Region of Interest (ROI) Extraction: Localizing suspicious lesions or abnormal character regions from the image by means of region of interest techniques thereby, analysis can be targeted and computational cost is minimized.

These preprocessing techniques are developed for each cancer subtype to address the specific challenges and features belonging to each of these cancer types, therefore offering better segmentation and detection processes for lung, breast, and skin cancers.

3.5 Algorithms:

MobileNet: MobileNet stands out as a convolutional neural network architecture designed to excel on mobile and embedded devices, prioritizing efficiency and swift inference. Its key innovation lies in employing depth-wise separable convolutions, which split the standard convolution into separate depth-wise and point-wise convolutions. This strategy drastically cuts down computational costs while maintaining respectable accuracy levels. Through depth-wise convolutions, it disentangles the spatial and depth dimensions of the input tensor, leading to a smaller model size and faster inference speeds. MobileNet is particularly favoured for applications demanding real-time processing on devices with limited computational resources, including mobile phones and IoT devices.

ResNetV2: ResNetV2 is an improved version of the standard ResNet architecture, which aims to address the issues of overfitting associated with very deep models. It undertakes this by skip connections (identity shortcuts) between layers of the network to perform residual functions. These connections offer a solution in case of vanishing/exploding gradients when trained deeper networks. Hence, those layers can be skipped by the network and take learning as deeper by more than one layer can be done. Now, we are able to train networks that have as many as 150 layers. It has been proven that these deep nets can be applied to classification, recognition, and a number of other vision applications such as object detection, etc. In our project this algorithm is used for lung and breast cancer.

DenseNet169: The architecture DenseNet169 is one of the models that DenseNet family is renowned for because of its peculiar dense connectivity pattern. Unlike similar structures in which only the few layers mentioned before and after the current layer are connected, DenseNet allows the connections show up between all the layers in a feed-forward fashion. This indicates that networking is very fast so features becoming reused, gradient propagating quicker, there is no vanishing gradient and it is widely noticed in better neural networks. In the architectures of DenseNet which are known to be very effective when it comes to complexity efficiency, features from the layer layers below are always stacked together in feature extraction. In this instance, DenseNet169 can be used in different industries, for instance image recognition and medical imaging applications, and it performs well for features extraction tasks.

Xception: The Inception architecture variant known as Xception uses depth separable warps as its core building blocks. Instead of transforming both spatial and depth parameters at once as in standard curve layers, depth-wise separation transforms perform these two operations individually by first applying spatial diffraction followed by depth-wise transformation decoupling these methods these two thereby reduce the complexity of parameter estimation so that faster and scalable models are possible. However, Xception is able to achieve impressive performance in a variety of image segmentation parameters by choosing its small construction step which makes it attractive. In our project this algorithm is used for skin cancer.

4. EXPERIMENTAL RESULTS

The web application consists of three type cancer , where we can upload the file of the respective cancer.

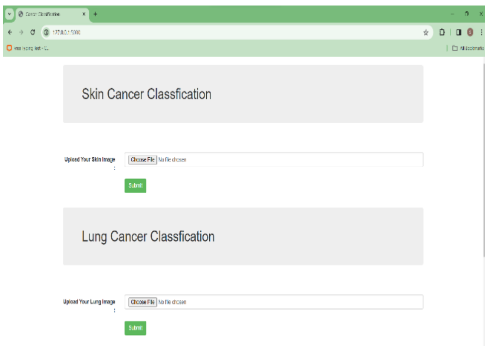

Fig 4.1: Web Application

Upon uploading the file we will get prediction. For instance uploading lung cancer file we can predict the type of lung cancer.

Fig 4.2: Lung Cancer Prediction

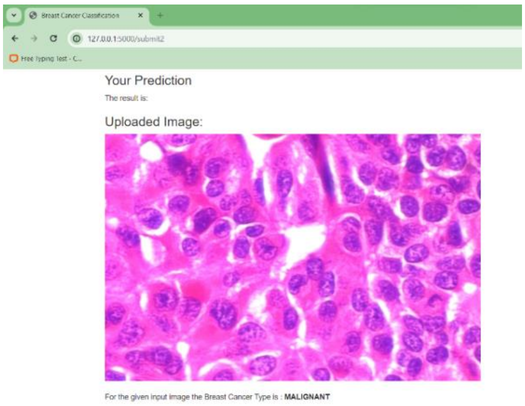

Fig 4.3: Breast Cancer Prediction

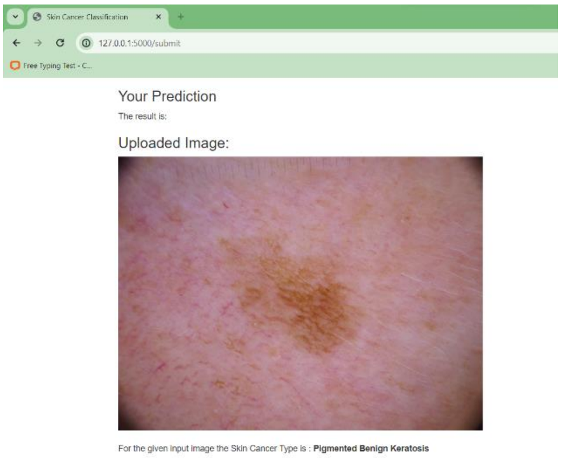

Fig 4.4: Skin Cancer Prediction

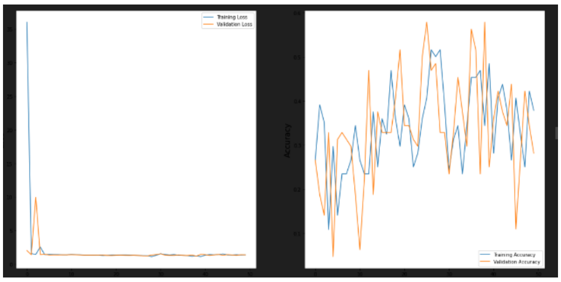

Fig 4.5: Accuracy Graph of Breast cancer

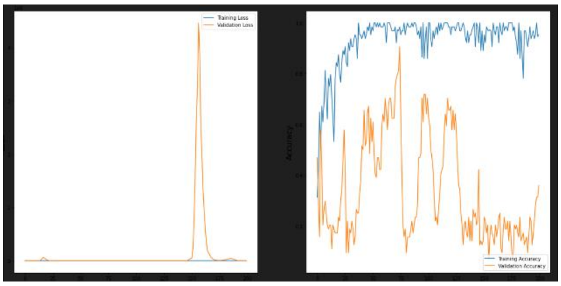

Fig 4.6: Accuracy Graph of lung cancer

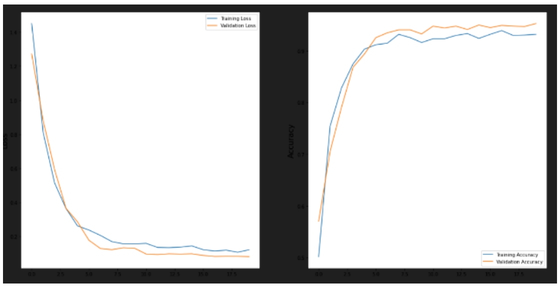

Fig 4.7: Accuracy Graph of skin cancer

The accuracy of our project is of 95%,where lung cancer's accuracy is 95.31%,breast cancer's accuracy is 89.06,and skin cancer's accuracy is 95.26%.

5. CONCLUSION

The project goal was to build a reliable system, where deep learning methods i.e. convolutional neural network architectures (CNN), transfer learning techniques, and flask web interface were used. This well established integrative system shows huge potential in the cancer research and medical imaging domain by giving sophisticated methods for the identification, segmentation, and detection of cancer cells of diverse types such as lung, breast, and SCC (squamous cell carcinoma).

On the basis of conventional CNN structures, the project was meant to help with feature extraction and learning from medical images, which, in turn, is critical for the proper recognition of cancer cells. The training techniques that used representations from pre-trained models such as MobileNet, ResNetV2 and DenseNet169 allowed the segmentation and tracking process to become faster as well as more robust, especially when dealing with datasets which were not very large.

The Flask-based web interface turned out to play an effective role in providing an intuitive platform for users to upload images causing the process of running the CNN models on the backend. This interface could be manipulated to achieve accurate segmentation and detection results.Such a fast and interactive system, which has a great potential, can be used by both medics, researchers, and all the stakeholders at large involved in cancer diagnosis, treatment planning, and monitoring.

In conclusion, this project combines the power of deep learning, transfer learning, and simple web usage into a user-friendly application with a significant role to play in cancer cell analysis. Such system helps in narrowing down the types of diagnosis, treatment, and behaviour of cancer cells. This promotes scientific research which can lead to improvements in diagnostic accuracy, treatment strategies, and general understanding about the nature of cancer, thereby contributing to the fight against cancer.

6.FUTURE SCOPE

Furthermore, this project may be carried forward in future by adding cell tracking to monitor changes in cancer cells continuously over time. Also, such predictive models could also be build to forecast cancer stages and progression with the help of imaging and molecular information. Multi-modal data fusion utilization should assist to make cancer cell examination more precise. Currently, clinicians during cancer diagnosis and treatment are aided by real-time decision support systems. Moreover, validation studies alongside clinical trials will be important in order to rate the usability and effectiveness of the developed technologies.

REFERENCES

[1] Wenbin He, Yongjie Han," Progress of Machine Vision in the Detection of Cancer Cells in

Histopathology",IEEE,vol.10,Mar. 2022

[2] J. Hayashida and R. Bise, "Cell tracking with deep learning for cell detection and motion estimation in low-frame-rate," in MICCAI. Springer,2019, pp. 397–405.

[3] Scherr, Tim, et al. "Cell segmentation and tracking using CNN-based distance predictions and a graph-based matching strategy." Plos One 15.12 (2020):e0243219.

[4] A. Javaid, M. Sadiq and F. Akram, "Skin Cancer Classification Using Image Processing and Machine Learning," *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*, Islamabad, Pakistan, 2021, pp. 439-444,

[5] Thulasi , M. S. ., B. . Sowjanya, K. . Sreenivasulu, and M. R. . Kumar. "Knowledge Attitude and Practices of Dental Students and Dental Practitioners Towards Artificial Intelligence". International Journal of Intelligent Systems and Applications in Engineering, vol. 10, no. 1s, Oct. 2022, pp. 248-53

[6] Madapuri, R.K., Mahesh, P.C.S. HBS-CRA: scaling impact of change request towards fault proneness: defining a heuristic and biases scale (HBS) of change request artifacts (CRA). Cluster Comput **22** (Suppl 5), 11591–11599 (2019). https://doi.org/10.1007/s10586-017-1424-0

[7] J. R. Dwaram and R. K. Madapuri, "Crop yield forecasting by long short-term memory network with Adam optimizer and Huber loss function in Andhra Pradesh, India," Concurrency and Computation: Practice and Experience, vol. 34, no. 27. Wiley, Sep. 18, 2022. doi: 10.1002/cpe.7310

[8] Swetha, A. ., M. S. . Lakshmi, and M. R. . Kumar. "Chronic Kidney Disease Diagnostic Approaches Using Efficient Artificial Intelligence Methods". International Journal of Intelligent Systems and Applications in Engineering, vol. 10, no. 1s, Oct. 2022, pp. 254.

[9] Rudra Kumar, M., Gunjan, V.K. (2022). Machine Learning Based Solutions for Human Resource Systems Management. In: Kumar, A., Mozar, S. (eds) ICCCE 2021. Lecture Notes in Electrical Engineering, vol 828. Springer, Singapore. https://doi.org/10.1007/978-981-16-7985-8_129

[10] Md. Mokhlesur Rahman; T. M. Shahriar Sazzad; Farhan Sedeed Ferdaus," Automate d Detection of Lung Cancer Using MRI Images" 2021 3rd International Conference on Sustainable Technologies for Industry.

[11] Yu, H., Zhou, Z., & Wang, Q. (2020). Deep Learning Assisted Predict of Lung Cancer on Computed Tomography Images Using the Adaptive Hierarchical Heuristic Mathematical Model. IEEE Access, 8, 86400–86410.

[12] Utpol Kanti Das; Juel Sikder; Umme Salma; A M Shahed Anwarn "Intelligent Cancer Detection System" 2021 International Conference on Intelligent Technologies (CONIT)