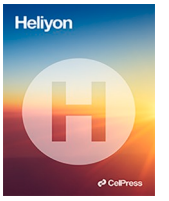
Heliyon 10 (2024) e24403

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

YOLO and residual network for colorectal cancer cell detection and counting

Inayatul Haq [a], Tehseen Mazhar [b], Rizwana Naz Asif [c], Yazeed Yasin Ghadi [d], Najib Ullah [e], Muhammad Amir Khan [f,*], Amal Al-Rasheed [g]

[a] School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, 450001, China
[b] Department of Computer Science, Virtual University of Pakistan, Lahore, 55150, Pakistan
[c] School of Computer Science, National College of Business Administration and Economics, Lahore, 54000, Pakistan
[d] Department of Computer Science and Software Engineering, Al Ain University, Abu Dhabi, 12555, United Arab Emirates
[e] Faculty of Pharmacy and Health Sciences, Department of Pharmacy, University of Balochistan, Quetta, 08770, Pakistan
[f] School of Computing Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia
[g] Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

ARTICLE INFO

Keywords:
Biomedical image processing
Computer-aided diagnosis
Medical image analysis
Machine learning
Deep learning
HT-29 cells
Colon cancer

ABSTRACT

The HT-29 cell line, derived from human colon cancer, is valuable for biological and cancer research applications. Early detection is crucial for improving the chances of survival, and researchers are introducing new techniques for accurate cancer diagnosis. This study introduces an efficient deep learning-based method for detecting and counting colorectal cancer cells (HT-29). The colorectal cancer cell line was procured from a company. Further, the cancer cells were cultured, and a transwell experiment was conducted in the lab to collect the dataset of colorectal cancer cell images via fluorescence microscopy. Of the 566 images, 80 % were allocated to the training set, and the remaining 20 % were assigned to the testing set. The HT-29 cell detection and counting in medical images is performed by integrating YOLOv2, ResNet-50, and ResNet-18 architectures. The accuracy achieved by ResNet-18 is 98.70 % and ResNet-50 is 96.66 %. The study achieves its primary objective by focusing on detecting and quantifying congested and overlapping colorectal cancer cells within the images. This innovative work constitutes a significant development in overlapping cancer cell detection and counting, paving the way for novel advancements and opening new avenues for research and clinical applications. Researchers can extend the study by exploring variations in ResNet and YOLO architectures to optimize object detection performance. Further investigation into real-time deployment strategies will enhance the practical applicability of these models.

1. Introduction

This study aims to develop a model for accurately detecting and counting colorectal cancer cells (HT-29). The irregular shape and

* Corresponding author.
E-mail addresses: Inayatulhaq1@gmail.com (I. Haq), tehseenmazhar719@gmail.com (T. Mazhar), rizwana@ncbae.edu.pk (R.N. Asif), Yazeed.ghadi@aau.ac.ae (Y.Y. Ghadi), najibullah9@yahoo.com (N. Ullah), amirkhan@uitm.edu.my (M.A. Khan), aalrasheed@pnu.edu.sa (A. Al-Rasheed).

https://doi.org/10.1016/j.heliyon.2024.e24403
Received 5 August 2023; Received in revised form 30 December 2023; Accepted 8 January 2024
Available online 11 January 2024
2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

dense clustering of HT-29 cells in images complicate accurate identification and quantification. To address this gap in such cases, a new approach is required to improve the accuracy and efficiency of detecting and counting HT-29 cells in a congested and overlapping form. Colon cancer is a significant public health concern, and its early detection plays a crucial role in improving treatment results and patient survival rates. Early detection is associated with higher survival rates, providing more treatment options, such as surgery, chemotherapy, or targeted therapy, particularly when the cancer is localized. This, in turn, leads to a better quality of life for patients due to less aggressive treatments and reduced side effects. Moreover, early detection can lower treatment costs than interventions at advanced stages. It also empowers individuals to take an active role in healthcare by making informed decisions about treatment options and lifestyle changes [1,2].

Colorectal cancer is a malignant neoplasm originating in the colon or rectum. On the inner lining of these organs, it typically begins as benign polyps and gradually transforms into cancerous growths. This form of cancer is common and affects both men and women. Age, family history of colorectal cancer, genetic predisposition, lifestyle choices (such as diet and physical activity), and inflammatory bowel disease all influence the development of colon cancer [3]. Changes in blood in the stool, bowel habits, abdominal discomfort, unexplained weight loss, and fatigue are common symptoms [4]. Regular screenings, such as colonoscopies, greatly increase the likelihood of successful treatment and recovery if performed promptly [5,6]. The other names for colorectal cancer (CCR) are colon or bowel cancer and colon or rectum cancer. These organs are integral components of the digestive system and are responsible for the processing and elimination of waste. Typically, the growth of benign polyps, or clusters of cells, within the colon or rectal lining marks the beginning of colorectal cancer. Some of these polyps can develop into cancerous tumors. CCR is a prevalent form of cancer and one of the most commonly diagnosed types worldwide [7,8]. HT-29 cells share similarities with cancer cells in colorectal tumors, making them useful for studying various aspects of CCR. Commonly, HT-29 cells are used as a model system to investigate cell proliferation, differentiation, apoptosis (programmed cell death), and cell signaling pathways, drug sensitivity, and gene expression in colorectal cancer. These cells are frequently used to evaluate the effects of potential anti-cancer agents, screen drugs for their therapeutic efficacy, and elucidate the molecular mechanisms underlying the development and progression of colorectal cancer [9,10]. Fig. 1(a) depicts the labeled diagram of the human colon, and Fig. 1 (b) depicts the colorectal cancer tumors (cells) [11,12].

HT-29 cells for both in vitro and in vivo experiments demonstrate the versatility and significance of these cells in cancer research. HT-29 cells are primarily used for in vitro studies in cancer research. Some researchers in studies [13–17] employed HT-29 cells to assess the anti-cancer effects of various substances, such as Syzygium cumini ethanolic extract or Lactobacillus acidophilus CICC 6074. HT-29 cells are used in Ref. [13] for in vitro experiments to assess the cytotoxic effects of AgNPs on colon cancer cells. The controlled in vitro setting allows researchers to observe cellular responses and molecular changes, providing insights into the potential therapeutic applications of the synthesized nanoparticles. HT-29 cells were also utilized for both in vitro experiments, where the anti-proliferative effects of YL-11 EPS are assessed at the cellular level, and in vivo experiments, where the efficacy of YL-11 EPS is evaluated in a tumor xenograft model [14]. This approach underscores the significance of HT-29 cells as a valuable tool for studying the potential anti-cancer properties of the tested exopolysaccharide in both controlled laboratory conditions and living organisms [15]. Also, HT-29 cells are employed for in vitro cancer research, specifically evaluating the anti-cancer effects of Syzygium cumini ethanolic extract (SCE). Utilizing standard in vitro techniques such as MTT assays, DNA damage assessment, and scratch tests, the research demonstrates that SCE significantly suppresses HT-29 cell growth, induces apoptosis, and reduces metastatic potential. Molecular analysis, including examining apoptosis-related genes, supports the observed effects.

The SCE could be a potential chemotherapeutic agent, proposing its suitability for in vivo trials based on promising in vitro results. This research underscores the importance of HT-29 cells in advancing our understanding of potential anti-cancer agents [16]. Another research underscores the application of HT-29 cells in cancer research, employing them for both in vitro and in vivo studies to assess the anti-cancer effects of Lactobacillus acidophilus CICC 6074. In vitro experiments demonstrate that L. acidophilus CICC 6074 inhibits HT-29 cell activity and induces apoptosis in a dose- and time-dependent manner. Mechanistic insights reveal mitochondrial membrane

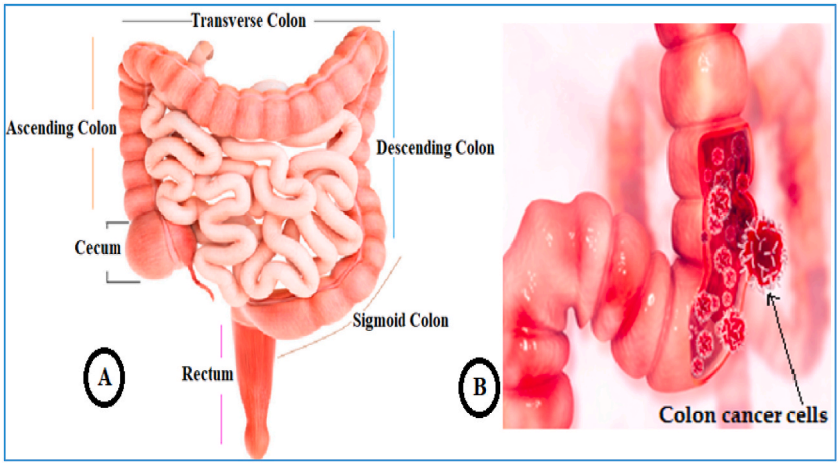


Fig. 1. Labeled diagram of the human colon: (a) colon, rectum, and (b) colorectal cancer tumors (cells) [11,12].

potential loss, release of Cytochrome *c* (Cyt *c*), and modulation of key apoptosis-related proteins. The in vivo study, utilizing an HT-29 cell-induced colon cancer mouse model, further validates these findings, showing notable apoptosis in colon cancer cells. The comprehensive approach, spanning from in vitro to in vivo investigations, establishes HT-29 cells as valuable tools for understanding the anti-cancer mechanisms of L. acidophilus CICC 6074, providing potential insights for probiotic functional food development in the market [17].

Researchers use crowd-counting to estimate the count of people in a crowd or densely populated area, typically through images or videos. It can be implemented in crowd management, public safety, event planning, and urban surveillance. Typically, crowd-counting algorithms employ computer vision methods, such as feature extraction, object detection, and machine learning (ML), to analyze an image or video frame and determine the number of individuals within it. CNNs are frequently used for crowd counting because they can learn complex visual patterns and characteristics from images [18,19]. This study is also based on crowd counting of HT-29 cells, but there is a gap in previous studies because they have only implemented deep learning (DL) techniques for crowd-counting people, and we are counting cancer cells. So, our approach is object-based for detecting and counting colorectal cancer cells (HT-29).

However, detecting cancer cells within images when densely congested and overlapping poses significant challenges for ML techniques. The complexity arises from the difficulty distinguishing individual instances, leading to intricacies in instance segmentation and challenges in determining precise object localization. Overlapping cells introduces ambiguity and interference in feature extraction, making it challenging for models to differentiate between cells and increasing the likelihood of false positives. The presence of complex backgrounds and the loss of distinct cell morphology information further compound these challenges. Additionally, the reduced signal-to-noise ratio in images with overlapping cells complicates distinguishing relevant signals, i.e., cancer cells, from noise or irrelevant features. The computational demands on ML models also escalate due to the complex nature of overlapping instances [20]. To address these challenges, advanced ML models and improved feature extraction techniques are essential, and the exploration of DL architectures like Mask R–CNN, designed for object detection and instance segmentation, becomes crucial in overcoming the complexities introduced by the crowded and overlapping arrangements of cancer cells in images [21,22].

Deep Learning techniques are used in this study because DL can detect and count the objects accurately and find the deep hidden patterns. Using DL techniques, pattern recognition can be applied to medical images to detect and extract vital patterns or features. It entails gaining useful insights from raw data and employing them to spot repeating patterns or structures. DL models like CNNs are trained on large collections of annotated medical images to learn representative patterns and features associated with various anatomical structures or pathologies. Patterns of interest, such as organs, tumors, lesions, or abnormalities, can be automatically detected and localized by these models [23,24]. Because of the tremendous applications of DL, we are using it for HT-29 cell (in medical images) detection and counting. While the YOLO algorithm stands out as a robust and efficient object detection method, its precision in detecting cancer cells depends on several factors. These factors encompass the quality of the dataset, the configuration of the model, the intricacies of the training process, and the inherent variability present in different types of cancer cells. Achieving accurate and reliable results necessitates thoroughly considering these aspects and meticulous algorithm fine-tuning to cater to the specifics of cancer cell detection tasks [25,26]. Residual Networks (ResNet) excel in image-related tasks, particularly image classification. Their versatility extends to object counting, offering practical strategies such as modifying architecture for object detection, transforming counting into density map regression, and leveraging transfer learning for fine-tuning counting-specific datasets. Features like Feature Pyramid Networks (FPN) enhance scale-aware counting. ResNet's adaptability allows regression heads to predict object counts. The applicability depends on task characteristics, object variability, and dataset nature. Specialized models like count-ception or networks for crowd counting may offer optimized performance in distinct counting scenarios [27,28].

When choosing a YOLO version for object detection, it's essential to consider factors such as the target hardware, the need for real-time processing, and the desired trade-off between accuracy and speed. The YOLO (You Only Look Once) object detectors' family has rapidly evolved. These YOLO variants share a common foundation centered around achieving real-time object detection and high-classification performance [29]. Faster-YOLO stands out due to its real-time capabilities, efficient feature extraction, adaptability to different datasets, reduced parameter tuning requirements, improved detection accuracy, and faster training speed compared to specific YOLO versions [30]. While YOLOv4 and YOLOv4-tiny are known for their efficiency, their choice depends on the project's specific requirements. YOLOv4 provides better accuracy but may be slower, while YOLOv4-tiny sacrifices some accuracy for faster inference [31,32]. The recent version of the YOLO family of algorithms is YOLOv8. These latest versions proved best in terms of accuracy and efficiency in detecting cancer [33,34] and other objects.

Table 1 contains the abbreviations used in this article.

2. Previous work

In a comprehensive study conducted from 2015 to 2016, 255 Danish national screening program participants were subjected to a fully paired analysis. All identified polyps underwent CCE, conventional colonoscopy, and histopathology examinations. The algorithm for autonomous detection demonstrated exceptional levels of precision (96.4 %), sensitivity (97.1 %), and specificity (93.0 %) [35]. Integrating DL and SLIM for screening applications was the subject of a study. The researchers trained a DL model with 1660 SLIM images of colon glands and validated its performance with 144 images of the glands. The results demonstrated an exceptionally high precision, with 98 % accuracy on the validation and 99 % on the test dataset. In addition, the model distinguished benign from malignant glands with a commendable 97 % accuracy, as measured by the area under the ROCC [36]. A classification scheme was introduced using histopathological images to discriminate between five categories of colon and lung tissues [37]. Similarly, a machine vision-based method for detecting TB in lung CT scan images was developed. The method involved the selection of ROI from TB-infected and normal lung images, followed by pre-processing and extracting statistical texture features. Supervised learning

Table 1
Abbreviations used in this article.

Acronym	Full form	Acronym	Full form
CNNs	Convolutional neural networks	DANet	Density Attention Network
CCE	Capsule colonoscopy	ASNet	Attention Scaling Network
SLIM	Spatial Light Interference Microscopy	APLoss	Adaptive Pyramid Loss
TB	Tuberculosis	OT	Optimal Transport
ROI	Regions of Interest	EMT	Epithelial-mesenchymal transition
MLP	Multi-Layer Perceptron	FCNs	Fully Convolutional Networks
ANNs	Artificial Neural Networks	SGD	Stochastic Gradient Descent
SSIM	Structural Similarity Index	TCGA	Cancer Genome Atlas
PSNR	Peak Signal-to-Noise Ratio	RFS	Relapse-free survival
CRC	Colorectal cancer	OS	Overall survival
DL	Deep Learning	ML	Machine Learning
RS	Raman spectroscopy	POST	Point-of-care testing
ROCC	receiver operating characteristic curve	CC	Colon Cancer
MAE	Mean Absolute Error	FCN	Fully convolutional networks

classifiers were employed, among which the MLP-based classifier, also known as the ANN, achieved 99 % accuracy in less than 1 s [38].

A technique was presented for automatically coloring and differentiating various regions in MR images of the human brain. The method utilizes index-based colorization and connected component techniques to distinguish between normal and effected regions by assigning colors to distinct areas. The proposed algorithm combines segmentation, ROI selection, and colorization algorithms to improve tissue identification and image contrast. Compared to industry-standard color transfer methods, the algorithm significantly improves processing efficiency and reduces time requirements [39]. A novel algorithm (CNN) was developed to compare multiple images to identify nodules and facilitate early cancer diagnosis. These results demonstrate the algorithm's effectiveness in minimizing false positives through rigorous and extensive training [40].

A single-tissue region was manually outlined on 86 CRC tissue slides and used to train a CNN using transfer learning. On an independent dataset, the CNN achieved a remarkable accuracy of over 94 % in classifying nine distinct classes of CRC tissue [41]. The ResNet architecture was implemented for classifying colorectal cancer in medical images. ResNet-18 and ResNet-50 models were trained on a dataset of images of colon glands to distinguish benign from cancer. On each test dataset, ResNet-50 outperforms ResNet-18 regarding accuracy, sensitivity, and specificity [42]. The analysis of colon cancer histopathological images utilized CNNs and segmentation models. On the AiCOLO dataset, ResNet achieved an accuracy rate of 96.98 %. In comparison, UNet and SegNet pixel-wise segmentation models reached a maximum accuracy of 81.22 % due to limited annotated datasets, necessitating the application of transfer learning techniques. Through intensive model testing exercises that involved training technique identification for precise colony tumor delineation purposes, optimal approaches were established in this study [43].

Colon cancer detection was explored through CT colonography, integrating image processing and DL techniques. The primary objective is to segregate the colon from noise in CT images, focusing particularly on identifying polyps. The proposed method involves image pre-processing, where each image undergoes segmentation into colon segments and noise categories using a pre-trained CNN. Subsequently, the system diagnoses polyp-like structures. The study underscores the substantial accuracy enhancement achieved through DL in automated colon cancer diagnosis, providing a foundational framework for future research [44]. Transfer learning was employed utilizing pre-trained CNNs such as Alexnet, VGG-16, and Inception-V3 to automatically extract high-level features from colon biopsy images. These features are then used to train a Bayesian-optimized SVM classifier. The proposed framework undergoes evaluation on four datasets, including two from Indian hospitals and two public colon image datasets. Comparative analyses underscore the superiority of the Inception-V3 network, achieving an accuracy range from 96.5 % to 99 %. This study highlights the potential of transfer learning-based approaches for efficient and accurate colon cancer detection, offering a promising framework for automated diagnosis and prognosis [45]. Table 2 contains the summary and structured overview of the diverse objectives, datasets, techniques, and results of the previous studies.

2.1. Gap in previous work

Our study introduces a DL-based method for detecting and counting HT-29 cells in colorectal cancer, filling a specific gap in the literature as follows:

F0B7 Cell-Level analysis: While the above studies focus on tissue or (tumor) level analysis, our study specifically targets detecting and quantifying individual colorectal cancer cells (HT-29). This is a crucial aspect of cancer research, especially considering the heterogeneity of cancer cell populations within tissues.

F0B7 Overlapping cell detection: Our study explicitly addresses the challenging task of detecting and counting overlapping and congested cancer cells. This is a notable gap in previous studies where the focus might have been on well-separated or easily distinguishable cells or tumors.

F0B7 Dataset and methodology: Our study introduces a dataset derived from the HT-29 cell line, contributing to the diversity of datasets used in this research domain. Using YOLOv2 for cell detection and ResNet for counting demonstrates a comprehensive and innovative methodology.

The significance of counting and detecting colorectal cancer cells is known from the following points.

F0B7 Counting and detecting individual cancer cells, especially in cases of overlapping and congested cells, contribute to a better understanding of tumor heterogeneity. This is critical for adapting effective and personalized treatment strategies.
F0B7 Identifying and quantifying cancer cells at the individual level can enhance early detection capabilities, providing clinicians with valuable information for prognosis and treatment planning.
F0B7 Cell-level analysis is essential in drug development and testing, allowing researchers to evaluate the effectiveness of potential treatments on specific cancer cell types.
F0B7 Automated cell counting and detection can support pathologists and clinicians, streamlining the diagnostic process and potentially improving diagnostic accuracy.

3. Research methodology and materials

The section encompassed model architecture, dataset preparation from lab experiments, image collection, rigorous pre-processing, and augmentations, followed by model training and testing focusing on selecting pertinent performance and evaluation metrics.

3.1. Proposed integrated deep learning model

The DL detection and counting model proposed for this study is depicted in Fig. 2. This block diagram represents the flow of data through the integrated YOLOv2 and ResNet-50 model for HT-29 cell detection. The ResNet-50 model's intermediate layers are used for feature extraction, and the YOLOv2 layers, with anchor boxes and a specific activation layer, contribute to object detection. The final fully connected layer with softmax activation produces the model's output, which includes bounding boxes and labels for HT-29 cell detection. The entire architecture is designed to seamlessly combine the strengths of both models for accurate and reliable cell detection.

3.1.1. Cancer samples and dataset preparation

3.1.1.1. Cell culture, invasion and migration. This study employed the HT-29 colorectal adenocarcinoma cell line as a surrogate for authentic colorectal cancer cells. The cell lines utilized in the study were sourced from Procell Life Science & Technology Co., Ltd., located in Wuhan, China. The cells were cultured at a temperature of 37° Celsius under an atmosphere of 5 % carbon dioxide in McCoy's 5A medium supplemented with 10 % fetal bovine serum (FBS). To prepare data the same procedure was followed by previous studies [12,46].

A transwell system, specifically a transwell 24-well chamber made by Millipore, was used to evaluate the invading and migrating capacities of the cells, also discussed in [12,47]. After being detached with trypsin, the cultured cells were centrifuged at 800 g for 3 min in order to collect them. Then, 200 μL of serum-free McCoy's 5A medium were added to the upper compartment, and 10,000 cells were seeded into a transwell. A medium containing 10 % fetal bovine serum (FBS) was added to the lower chamber. The transwell chambers were pretreated with matrigel (1:4, BD, USA) at a concentration of 1:4 before the invasion assay was carried out. The cells living in the top layer of the filter were painstakingly removed after either 24 h (for the migration assay) or 48 h (for the invasion assay). The number of cells that were still adhered to the lower surface of the filter membrane after crystal violet staining was then monitored using the Olympus microscope. Also the same criteria was followed in previous studies [12,48]. Finally, 566 images were collected and divided into a training set of 80 % and a testing set of 20 %. Figs. 5 and 7 represent one of the collected images presenting colorectal cancer cells as blue shapes. Fig. 3 shows the complete steps of the dataset (HT-29 cells images) collection [12].

3.1.2. Pre-processing and augmentation of HT-29 data

To better apply the trained models to various histopathology images and settings, we augment the data to increase the network's robustness to data variability. The augmentation techniques are applied to the original images without adding new samples to the dataset. To add variety, we spatially and chromatically alter each image. Images and their associated binary masks undergo random spatial transformations like rotation in vertical and horizontal axes and cropping at various zoom levels. The images are also rotated 90 and 180° to give the impression that the pathologist is viewing them from above. The last step of the augmentation process involves randomly converting each slide to grayscale and adjusting the brightness, contrast, and saturation to diversify the dataset further [12]. These augmentation strategies collectively enhance the adaptability and robustness of DL models, particularly essential in medical image analysis, where variations in scale, orientation, and appearance are frequently encountered.

3.1.2.1. HT-29 cells labelling. We used the Image Labeler App includes a versatile ROI Labeler tool in MATLAB 2021. This tool empowers users to generate various types of Region of Interest (ROI) labels, including rectangles, polylines, pixels, and polygons, for individual images or sequences of pictures. Fig. 4 shows the ROI and scene labeling procedure [12]. To conduct data labeling, we followed a systematic approach using the image labeler:

F0B7 In the initial phase, we loaded data that lacked any pre-existing labels. This allowed us to start with raw, unlabeled information, laying the labeling process's foundation.

Table 2
Summary of the previous studies.

Study	Objective	Dataset	Techniques	Results
[35]	Improve CCE for colorectal cancer screening	Danish National Screening Program (255 participants)	A matching algorithm, CNN	Matching algorithm provides one-to-one match; CNN achieves 96.4 % accuracy
[36]	Automate diagnosis of colorectal tissue using optical coherence tomography	18 ex vivo human specimens	Pattern-recognition optical coherence tomography	ROC of 0.998 achieved
[37]	Automate lung and colon cancer diagnosis using AI and DL	Histopathological images of lung and colon tissues	DL and DIP techniques	96.33 % accuracy in identifying cancer tissues
[38]	Machine vision-based approach for Tuberculosis detection	TB infected and normal CT scan images	Supervised learning classifiers, ANN	ANN achieves 99 % accuracy
[39]	Automated identification of brain tumors in MRI images	Grayscale human brain MR images	Connected component and index-based colorization methods	Improved PSNR and SSIM values
[40]	Develop an algorithm for lung nodule detection in CT scan data	CT scan images	CNN algorithm	84.8 % accuracy, 90.47 % precision, 90.64 % specificity
[41]	Use CNNs to extract patterns from colorectal cancer tissue slides	86 CRC tissue slides, 862 HE slides	Deep CNNs, transfer learning	Deep stroma score is an independent prognostic factor
[42]	DL for colorectal cancer detection with ResNet architecture	Colon glands images	ResNet-18 and ResNet-50	ResNet-50 outperforms ResNet-18
[43]	DL architectures for colon cancer classification and segmentation	AiCOLO colon cancer dataset, CRC-5000, NCT-CRC-HE-100K, and others	Various CNN models, transfer learning	Up to 96.98 % accuracy for patch-level classification with ResNet
[44]	Colon cancer detection through CT colonography	Not specified	Pre-trained CNN for image pre-processing	Higher accuracy (87 % for segmentation, 88 % for polyp detection) compared to classical algorithms
[45]	Automated diagnosis of colon cancer from biopsy examination	Indian hospital datasets (4X, 10X, 20X, 40X magnifications), public colon image datasets	Transfer learning architectures; Bayesian optimized SVM	Inception-V3 network achieves 96.5 %–99 % accuracy

F0B7 Subsequently, we manually labeled an image frame from a collection of images. The interactive nature of the ROI Labeler tool provided a user-friendly environment to mark ROIs by hand, assigning labels as needed.
F0B7 We employed automation techniques within the Image Labeler to expedite the labeling process across multiple frames. This automation capability significantly increased efficiency by automatically assigning labels to images throughout the sequence.
F0B7 Upon completion of the labeling process, we utilized the export feature to extract the labeled data, creating a named ground truth dataset. This dataset served as a valuable resource for training and evaluating models.
F0B7 We addressed the significance of ROI labels, representing rectangular interest regions containing relevant information. These labels included a distinct name, such as "HT-29 cells," and information about the production location.

Fig. 5 shows the labeled image for colorectal cancer cells (HT-29 cells) via the green boxes.

3.1.2.2. Classification of histopathological images. All images were resized to 224 x 224 x 3 patches to comply with each pre-trained model's input size specifications. The learning process was conducted with a batch size 32 to guarantee uniform feature representation. A total of fifty training epochs were applied to each model. The optimization technique used Stochastic Gradient Descent (SGD) with momentum. The initial learning rate of 0.0001 per second was established. The learning rate was reduced by 0.10 after every ten epochs to accommodate hierarchical depths in the data. An additional layer with full connection and softmax activation was implemented to achieve consistent and accurate patch detection across diverse network setups. The size of this stratum is directly proportional to the number of objective categories. In equation (1), the probabilistic output of the softmax function is utilized to assign each patch to the category with the highest score. "n" is the total number of classes in the given equation, whereas "I" denotes the index of each element in the output vector. As a result, "yi" denotes the output of the last convolutional layer [12].

$$P(\mathrm{y}i) = \frac{e^{yi}}{\sum_{j}^{n} e^{yj}} \tag{1}$$

3.2. Architecture of ResNet-50

ResNet −50 is a DL network model specifically designed for image recognition applications. The model can handle input images with dimensions of 224 × 224 × 3, representing color images. It starts with a convolutional layer featuring 64 filters, each sized 7x7 with a stride of 2. After each convolutional layer, batch normalization of activations is applied to aid the training process. Dimensionality reduction is achieved through a max pooling layer with a 3x3 filter size and a 2-pixel stride. ResNet-50 employs residual blocks to address the vanishing gradients problem. These blocks establish rapid connections within, linking multiple convolutional

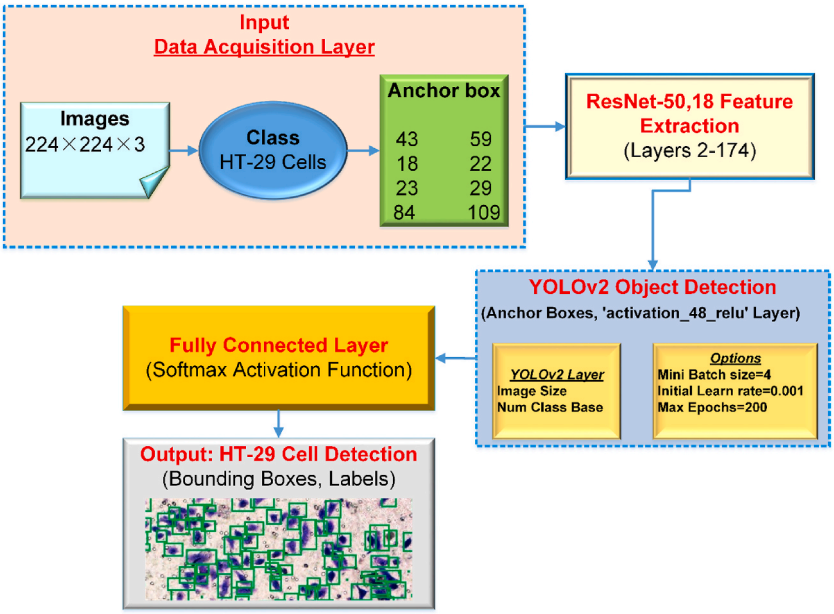


Fig. 2. Proposed integrated Deep Learning model.

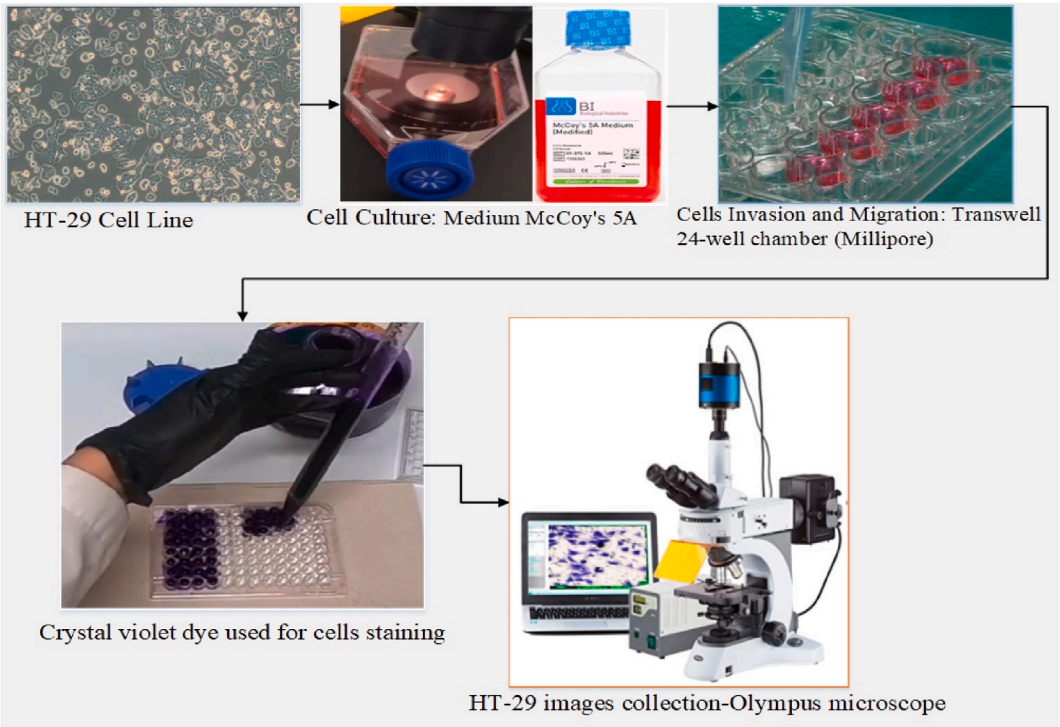


Fig. 3. The experimental steps for dataset collection [12].

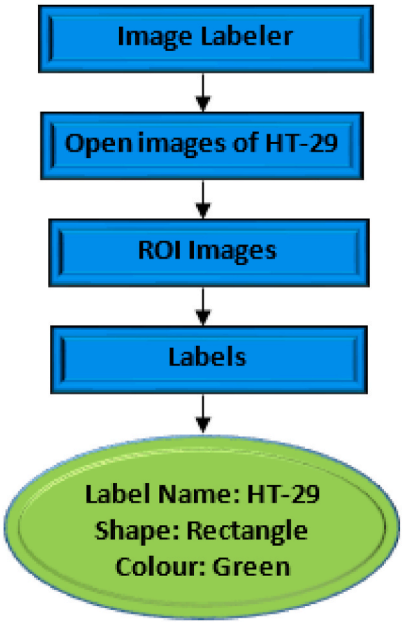


Fig. 4. Data (HT-29 cells) labelling.

layers. Subsequent architectural blocks consist of convolutional layers with varying numbers and sizes of filters, with the initial block consistently featuring three convolutional layers of different sizes. The output of each residual block, combined with the initial input through a shortcut link, enhances gradient flow, facilitating better training. Following the final residual block, global average pooling is applied to standardize the feature map dimensions to a predetermined size. To handle multiple output classes in the ImageNet dataset, a connection is established between the output of the global average pooling layer and an additional layer with 1000 fully connected units. Class probabilities are obtained by passing the output of the fully connected layer through a softmax activation function. The resulting probability distribution reflects the network's assessments of various classifications for the input image. ResNet-50's impressive architectural design incorporates the "residual learning" concept, contributing to its outstanding performance across diverse image recognition tasks [12,49]. Fig. 6 depicts the architecture of the ResNet-50 model [49].

3.2.1. ResNet for counting purposes

ResNet models can be used for counting objects in images. While ResNet architectures were primarily designed for image classification tasks, they can also be adapted for object counting by making a few modifications.

I. One common approach is to employ a two-step process: object detection followed by counting. In the detection step, an object detection model, such as Faster R–CNN or YOLO (YOLO is used here), can identify and localize objects within an image. These models are trained to generate bounding box predictions and object class probabilities. Once the objects have been detected, the counting step involves utilizing ResNet or its variants to classify each detected object as a separate instance of the desired object class. The ResNet model can be fine-tuned on the specific object class of interest using the detected object regions as inputs. By processing each detected object individually through the ResNet model, counting the number of instances of the desired object class present in the image becomes possible [50].
II. Another approach is to use regression-based methods, where the ResNet model is modified to output a count directly. Instead of classifying individual objects, the network is trained to predict the count of objects within an image as a numerical value. This can be achieved by modifying the last layer of the ResNet architecture to produce a scalar count prediction.

Both approaches require appropriate training data, where the images are labeled with object bounding boxes and corresponding object counts. The models can then be trained using these annotated datasets to learn the association between image content, object detection, and counting. It's worth noting that while ResNet models can be effective for object counting, specialized architectures such as FCNs or density-based counting methods like DensityNet are also commonly used for counting tasks. These models are designed to estimate object density maps, which can be integrated to obtain accurate object counts. These methods are also implemented in Refs. [51–55].

3.3. Model training and testing

The model proposed for this study is presented in Fig. 2. The dataset, comprising images containing HT-29 cells, is pre-processed to

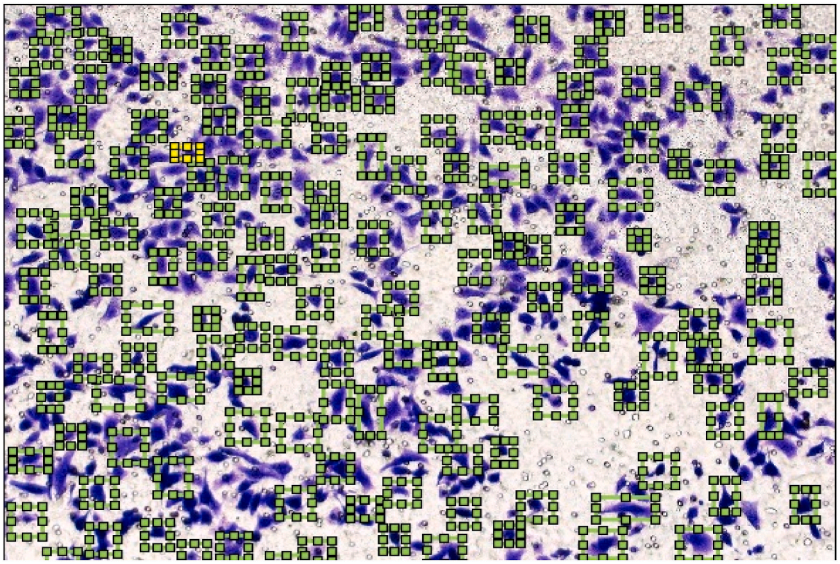


Fig. 5. labelling the HT-29 cells in images.

ensure uniformity. All images are resized to 224 x 224 x 3 patches to align with the input size specifications of the pre-trained ResNet-50 model, which is an integral component of the integrated architecture. The model architecture is a fusion of YOLOv2 and ResNet-50, 18. The ResNet-50 model is modified to extract features relevant to HT-29 cell detection. The intermediate layers (2–174) of ResNet-50 are retained, and a new layer is introduced for the final classification task. The YOLOv2 layers are integrated to facilitate object detection, utilizing anchor boxes for bounding box predictions. The feature extraction layer is set to 'activation_48_relu' in the ResNet-50 model.

The model is trained using a combination of the YOLOv2 and ResNet-50 layers. The training is conducted over 200 epochs with a batch size of 4. Stochastic Gradient Descent (SGD) with momentum is employed as the optimization technique. The initial learning rate is set to 0.001, and it is reduced by 0.10 after every ten epochs to adapt to the hierarchical depths present in the data. A final fully connected layer with softmax activation is added to ensure consistent and accurate HT-29 cell patch detection across various network setups.

Upon model training, the detection and counting of HT-29 cells are performed on both individual test images and a larger dataset. Bounding boxes and labels are predicted and visualized for a single test image. Additionally, the model is evaluated on a dataset containing multiple images, and the detection results are displayed. The testing process involves testing the trained model on a separate dataset. The performance of the integrated model is evaluated using precision, recall, and average precision metrics. The precision-recall curve is plotted based on the evaluation results, providing insights into the model's ability to accurately detect and count HT-29 cells.

3.4. Implementation tools

The proposed architectures were trained and tested on a system with an Intel(R) Xeon(R) Bronze 3204 1.9 GHz processor and 62 GB RAM. The operating system used was MS Windows 10. The DL framework in Matlab 2021 was employed to conduct all experiments. Data analysis was performed using MS Excel, and diagrams were created using the Canva tool.

3.5. Performance evaluation

We utilized the established metrics to evaluate and compare our models with state-of-the-art techniques:

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{TP + TN + FP + FN} \tag{2}$$

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{TN + FP} \times 100 \tag{3}$$

$$\mathrm{Sensitivity} = \mathrm{Recall} = \frac{\mathrm{TP}}{TP + FN} \times 100 \tag{4}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{TP + FP} \times 100 \tag{5}$$

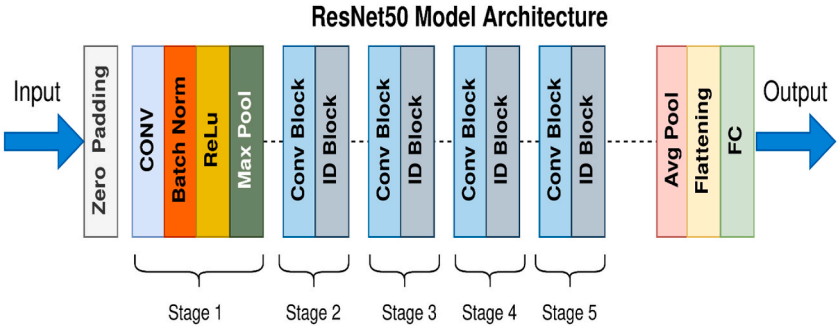


Fig. 6. The architecture of the ResNet-50 [12,49].

$$\text{False Positive Rate} = \frac{\text{FP}}{FP + TN} \times 100 \tag{6}$$

$$\text{False Positive Ratio} = 1 - \left(\frac{\text{Specificity}}{100}\right) \tag{7}$$

$$\text{False Negative Ratio} = 1 - \left(\frac{\text{Sensitivity}}{100}\right) \tag{8}$$

In the above equations (2)–(8) [12], TP means true positive, TN means true negative, FP is false positive, and FN represents false negative. Sensitivity assesses a model's capacity to correctly predict instances belonging to each category (true positives). At the same time, specificity gauges the model's proficiency in accurately predicting instances not belonging to each category (true negatives). These metrics apply to any categorical model.

4. Results and discussion

We comprehensively analyzed 566 images depicting colorectal cancer conditions and detection. The images were annotated using an image labeler to identify HT-29 cells. Subsequently, these images were exploited for training and testing using the DL architectures of ResNet-50 and ResNet-18. In our analysis, we focused on counting the number of cancer cells in each image to assess the accuracy of the proposed model. The evaluation involved measuring network performance, accuracy, and computational cost. To ensure robustness, each model underwent training and testing using ten independent non-overlapping random training and test data splits.

Furthermore, we conducted various tasks to explore and demonstrate the impact of data pre-processing on the overall results and performance of the models. The counting of HT-29 cells in the integrated YOLOv2 and ResNet-50 model is executed through a multi-step process. The trained object detector is initially applied to a single test image utilizing the detect function. This function returns crucial information, including bounding boxes, scores, and labels for identified objects in the image. Subsequently, the obtained results are visually represented on the test image using the "insertObjectAnnotation function", allowing for a clear understanding of the detection outcomes. The actual counting of HT-29 cells is achieved by summing the occurrences of specific labels as HT-29. These counts, represented by variables c1, signify the respective quantities of detected instances of HT-29 cells in the given test image. The same counting methodology is applied iteratively in a loop when processing multiple images from a dataset, providing a comprehensive analysis of the model's detection and counting capabilities across various samples. Table 3 shows the counting of HT-29 cells in images, and the total count and average values are shown in the last two rows.

The performance of the model was evaluated using the test set for HT-29 cells detection and counting. Fig. 7 shows the detection and count of HT-29 cells; Fig. 7 (a) count of HT-29 cells by ResNet-50 is 290, and Fig. 7 (b) count of HT-29 cells by ResNet-18 is 300. The Following sum function is used to count HT-29 cells in all images:

```
F0B7 ≫ sum(labels = = 'cancer')
F0B7 ans = 290
```

4.1. Performance evaluation of ResNet-50

The ResNet-50 model underwent training and testing phases, and various metrics were computed to assess the accuracy of HT-29 cell detection and counting. In the training phase of the ResNet-50 model, which utilized 70 % of the dataset, the model demonstrated a high level of precision, with a value of 0.97, indicating a minimal false-positive rate. The F1 score, a harmonic mean of precision and recall, was calculated to be 0.96, suggesting a balanced performance between precision and recall. The recall, representing the ability to capture all instances of HT-29 cells, reached 0.97. The model's overall accuracy during training was 96.66 %, indicating its

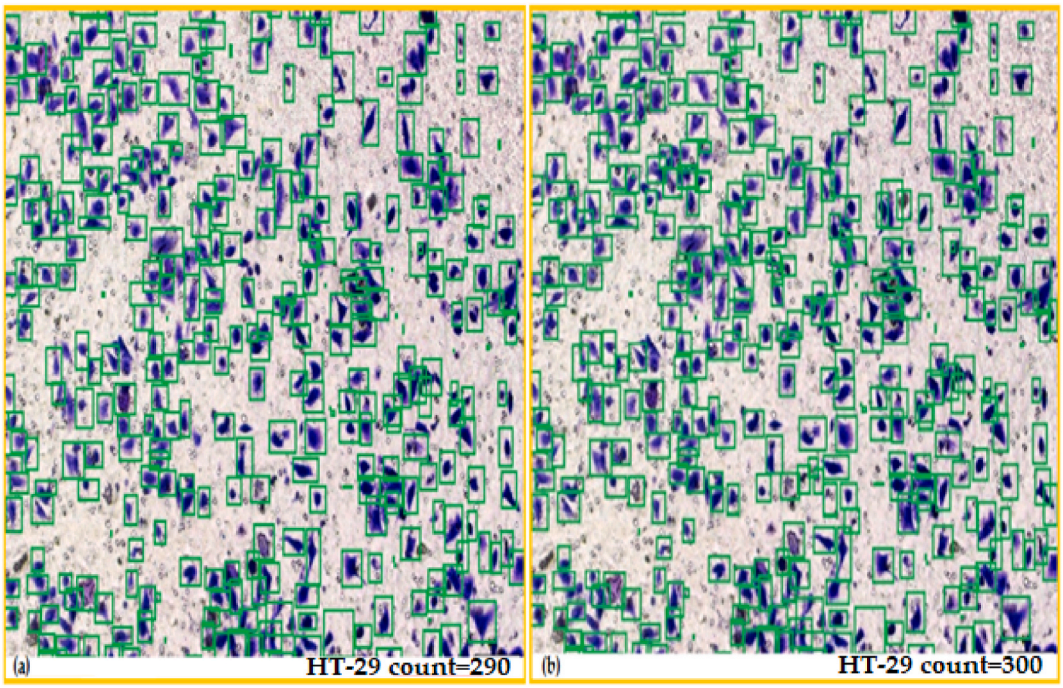


Fig. 7. Detection and total count of HT-29 cells in images: (a) ResNet-50 detection and count, (b) ResNet-18 detection and count.

proficiency in correctly classifying HT-29 cells.

During the test phase, where the model was evaluated on the remaining 30 % of the dataset, the ResNet-50 model maintained similar high-performance metrics. The precision remained at 0.97, underlining the model's consistency in minimizing false positives. The F1 score and recall were 0.96 and 0.96, respectively, indicating the model's effectiveness in balancing precision and recall on previously unseen data. The overall accuracy during testing stood at 96.33 %, reinforcing the model's generalization capability. These metrics collectively demonstrate the robustness and reliability of the ResNet-50 model in accurately identifying HT-29 cells, both during the training process and when faced with new, unseen data during the testing phase. The performance evaluation of the ResNet-50 approach is summarized in Table 5.

Fig. 8 depicts the ResNet-50 model in both its training and testing stages. The figure visually represents the model's performance, displaying the precision, F1 score, and recall values for detecting and counting HT-29 cells. The graph allows for a comprehensive understanding of how these metrics evolve throughout the training and testing processes, providing insights into the model's performance at each stage.

Fig. 9 illustrates the model's accuracy during the training and testing phases. Specifically, the figure illustrates the detection and counting precision of HT-29 cells. During the training phase, the model's accuracy reached 96.33 %. In the phase of testing, 95.3 % accuracy was recorded. This graph visually represents the model's accuracy performance throughout the training and testing processes.

Fig. 10 depicts the training progress of the ResNet-50 model, where Fig. 10 (a) shows the curve of RMSE and Fig. 10(b) shows the Loss curve. This graph represents the training process conducted on the images. Following the training phase, the model can perform object detection and counting on the test images. This graph provides insights into the accuracy and loss values achieved during

Table 3
HT-29 cells count in some of the images.

Image Number	HT-29 cells (colorectal cancer cells)
1	290
2	300
3	100
4	70
5	210
6	140
7	180
Total	1290
Average	184.285

The quantification method of colorectal cancer cells (HT-29) in images is presented in Table 4.

training. It visually depicts the model's efficacy and object detection/counting characteristics in images.

4.2. Performance evaluation of ResNet-18

In the training phase of the ResNet-18 model, encompassing 70 % of the dataset, the model exhibited exceptional precision with a value of 0.997. This high precision underscores the model's proficiency in minimizing false positives while identifying HT-29 cells. The F1 score, an amalgamation of precision and recall, reached 0.976, indicating a balanced trade-off between precision and recall. The recall, representing the model's ability to capture instances of HT-29 cells, achieved a commendable value of 0.988. The overall accuracy during training was recorded at an impressive 98.70 %, encouraging the model's accuracy in classifying HT-29 cells with high confidence.

During the subsequent testing phase, where the model was evaluated on the remaining 30 % of the dataset, the ResNet-18 model maintained its exceptional performance. The precision remained notably high at 0.989, suggesting the model's consistency in minimizing false positives on previously unseen data. The F1 score and recall were 0.968 and 0.987, respectively, demonstrating the model's effectiveness in balancing precision and recall. The overall accuracy during testing stood at 98.13 %, further indicating the model's robust generalization capability and ability to identify HT-29 cells in diverse datasets accurately.

These performance metrics collectively highlight the ResNet-18 model as a reliable and accurate tool for identifying HT-29 cells, showing its proficiency in both the training and testing phases. The performance evaluation of the ResNet-18 approach is summarized in Table 6 and depicted in Fig. 11.

The overall Accuracy of the ResNet-18 Model during the testing and testing phases of HT-29 is depicted in Fig. 12.

Fig. 13 depicts the training progress of the ResNet-18 model where Fig. 13 (a) shows the curve of RMSE and Fig. 13(b) shows the Loss curve.

4.3. Comparative analyses of employed models

The comparison of the ResNet-50 and ResNet-18 models' accuracy is depicted in Fig. 14. The comparative analysis between the ResNet-50 and ResNet-18 models for detecting and counting HT-29 cells reveals compelling insights into their respective performances. Notably, both models demonstrated impressive precision values during the training phase, with ResNet-50 achieving 0.97 and ResNet-18 surpassing it with a precision of 0.997. This indicates the ResNet-18 model's enhanced capability in minimizing false positives, showcasing its precision in identifying HT-29 cells. The F1 scores, harmonizing precision, and recall were commendably high for both models, with ResNet-18 slightly outperforming ResNet-50 (0.976 vs. 0.96). This suggests that the ResNet-18 model strikes a slightly better balance between precision and recall in the scenario of HT-29 cell detection.

In terms of recall, ResNet-18 consistently achieved high values, particularly evident during the training phase, where it reached 0.988, surpassing ResNet-50's recall of 0.97. This signifies the ResNet-18 model's superior ability to capture instances of HT-29 cells, making it more proficient at comprehensively identifying the target class. Moreover, both models' overall accuracy during training and testing was notably high, with ResNet-18 maintaining a marginal lead (98.70 % vs. 98.13 %), emphasizing its robust generalization capability. While both ResNet-50 and ResNet-18 exhibit impressive performance in detecting and counting HT-29 cells, the nuanced differences suggest that the ResNet-18 model holds a slight edge regarding precision, recall, and overall accuracy. These findings underscore the importance of model architecture in achieving optimal results for specific detection tasks, with ResNet-18 proving well-suited for the singular class detection of HT-29 cells in images.

The unexpected observation where ResNet-18 outperforms ResNet-50 in terms of accuracy in this study can be attributed to various influencing factors.

F0B7 One crucial aspect to consider is the unique characteristics of the dataset used for this study. The dataset's specific distribution, size, and complexity may favor the learning patterns captured by ResNet-18 over ResNet-50. The shallower architecture of ResNet-18 might be especially advantageous in cases where the dataset is relatively small or lacks the diversity required for ResNet-50 to generalize effectively.

F0B7 Another factor contributing to the observed performance could be the sensitivity of each model to the distribution of data or the inherent difficulty of the HT-29 cell detection task. As the dataset contains challenging instances or exhibits class imbalance, the reduced capacity of ResNet-18 might help prevent overfitting and perform more robustly in such scenarios.

F0B7 The influence of training hyperparameters should not be overlooked. The choice of learning rate, batch size, and data augmentation techniques can significantly impact model convergence and generalization. It is plausible that the specific combination of hyperparameters used in this study favors ResNet-18 over ResNet-50 for the HT-29 cell detection task.

F0B7 Also, the sensitivity of neural network architectures to random initialization could be a contributing factor. The random nature of weight initialization means that different training process runs may lead to varied results. If ResNet-18 happened to start with a more favorable set of initial weights for the HT-29 cell detection problem, it could have influenced its superior performance.

The cases below are where the accuracy of a ResNet-18 model is better than that of a ResNet-50 model, like here in the case of our dataset HT-29 cells. Although ResNet-50 is a deeper and more complex architecture than ResNet-18, it only sometimes guarantees higher accuracy.

Here, we are assuming a few scenarios where ResNet-18 might outperform ResNet-50:

Table 4
Final count of colorectal cancer cells (HT-29) in all images.

Class in images	Count ofimages	Counting procedure
HT-29 cells (Colorectal cancer)	566	566 x 184.285 = 104305.31

Table 5
Performance evaluation of the ResNet-50.

Model	Phase	Classes	Precision	F1 Score	Recall	Overall Accuracy
ResNet-50	Training 70 %	HT-29 cells	0.97	0.96	0.97	96.66 %
	Testing 30 %	HT-29 cells	0.97	0.96	0.96	96.33 %

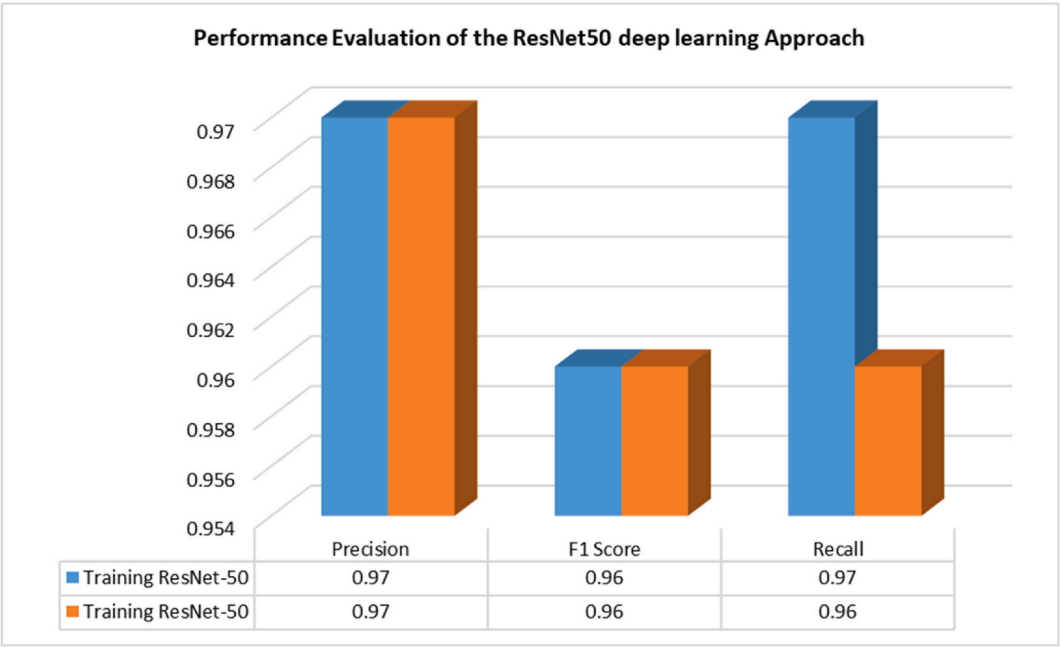


Fig. 8. Performance and evaluation of ResNet-50 during training and testing phases of HT-29.

F0B7 When the available training dataset is relatively small, ResNet-18, with its fewer parameters and simpler architecture, may generalize better and avoid overfitting compared to ResNet-50. ResNet-18 is a variant of ResNet designed to generate more robust representations without memorizing the training data [56–59].

F0B7 If the image classification task involves relatively simple patterns or features that can be captured effectively by shallower architectures, ResNet-18 might perform better. ResNet-50's deeper layers might not contribute significantly to the task if the additional depth is not required to capture the complexity of the image patterns [56–59].

F0B7 ResNet-50 requires more computational resources and memory due to its increased depth and parameter count. In scenarios where computational resources are limited, ResNet-18 may be preferred as it can provide a reasonable trade-off between accuracy and computational efficiency [56–59].

F0B7 In some cases, ResNet-50's additional capacity might lead to overfitting if the dataset is not large enough or diverse. ResNet-18's simpler architecture could be advantageous in such situations by offering better generalization and avoiding overfitting issues [56–59].

However, these scenarios are incomplete, and the performance comparison between ResNet-18 and ResNet-50 can vary depending on the specific dataset, task, and training conditions. Experimenting and evaluating multiple architectures is generally recommended to determine the best model for a particular problem. These cases are also reported in Refs. [56–59] studies. Conducting an ablation study or sensitivity analysis would reveal crucial factors that influence the performance of models in HT-29 cell detection, providing further insights through detailed investigation.

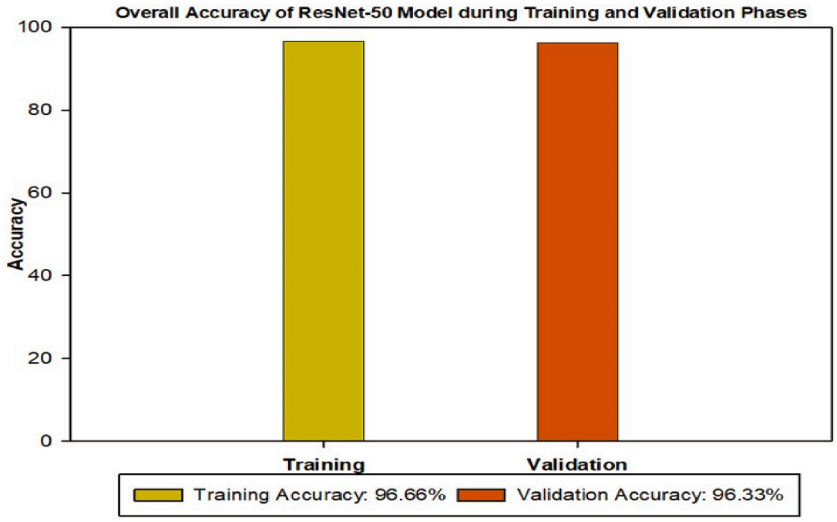


Fig. 9. Accuracy of ResNet-50 during the training and testing phases of HT-29.

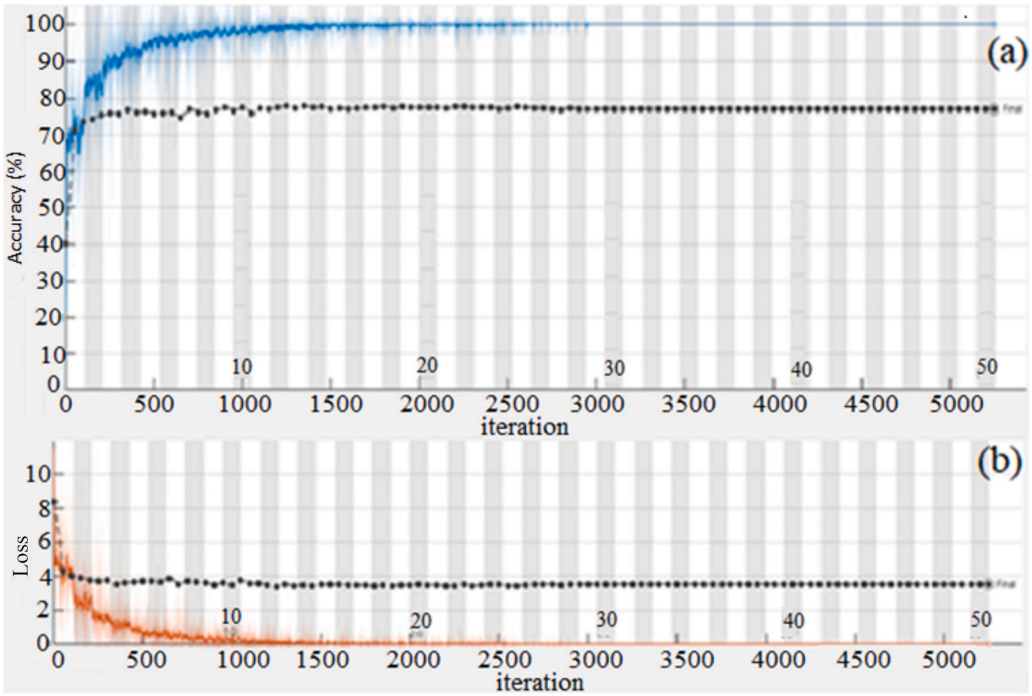


Fig. 10. ResNet-50 training progress: (a) shows the RMSE, and (b) shows the Loss.

Table 6
Performance evaluation of the ResNet-18.

Model	Phase	Class	Precision	F1 Score	Recall	Overall Accuracy
ResNet-18	Training 70 %	HT-29 cells	0.997	0.976	0.988	98.70 %
	Testing 30 %	HT-29 cells	0.989	0.968	0.987	98.13 %

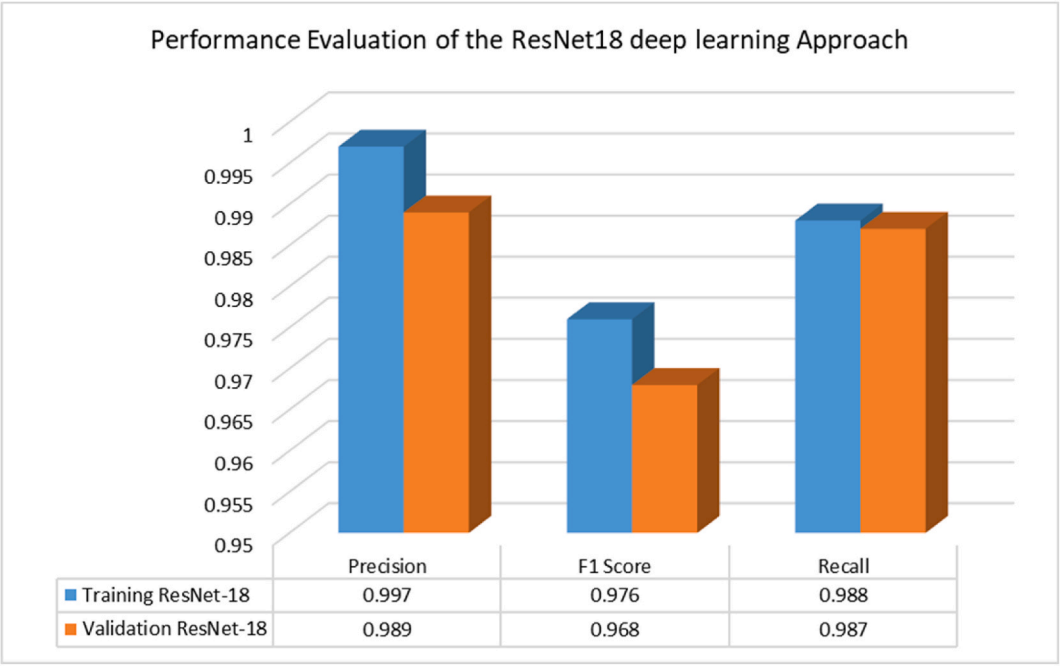


Fig. 11. ResNet-18 performance evaluation of training and testing phases of HT-29.

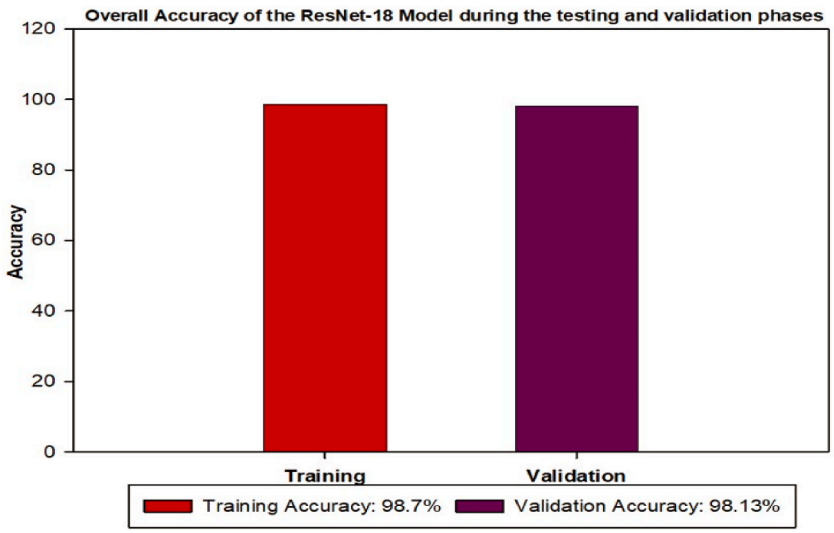


Fig. 12. Overall Accuracy of ResNet-18 model during testing and testing phases of HT-29.

4.4. Comparison with previous studies

In our focus on detecting and counting colorectal cancer cells (HT-29), the higher accuracy achieved by ResNet-18 compared to ResNet-50 can be attributed to several factors. The dataset, comprising images of congested and overlapping colorectal cancer cells obtained through fluorescence microscopy, plays a crucial role. The nature of the dataset, emphasizing crowded cell arrangements, aligns with ResNet-18's architecture, which may be more effective in capturing complex details and relationships within densely packed cell structures. Additionally, the choice of YOLOv2 for cell detection could complement ResNet-18's ability to count cells accurately. The lighter architecture of ResNet-18 might be advantageous for this specific task, as it might prevent overfitting on the given dataset. The performance differences between ResNet-18 and ResNet-50 in our study are compared to 15 other studies in Table 7.

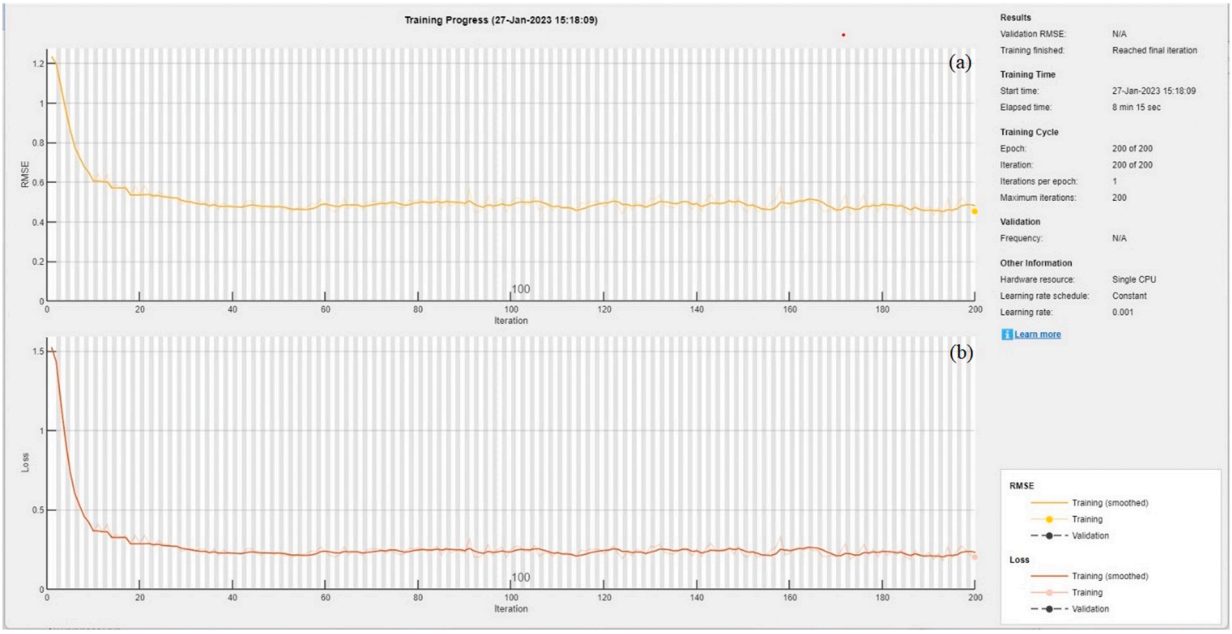


Fig. 13. ResNet-18's training progress: (a) shows RMSE, and (b) shows the Loss.

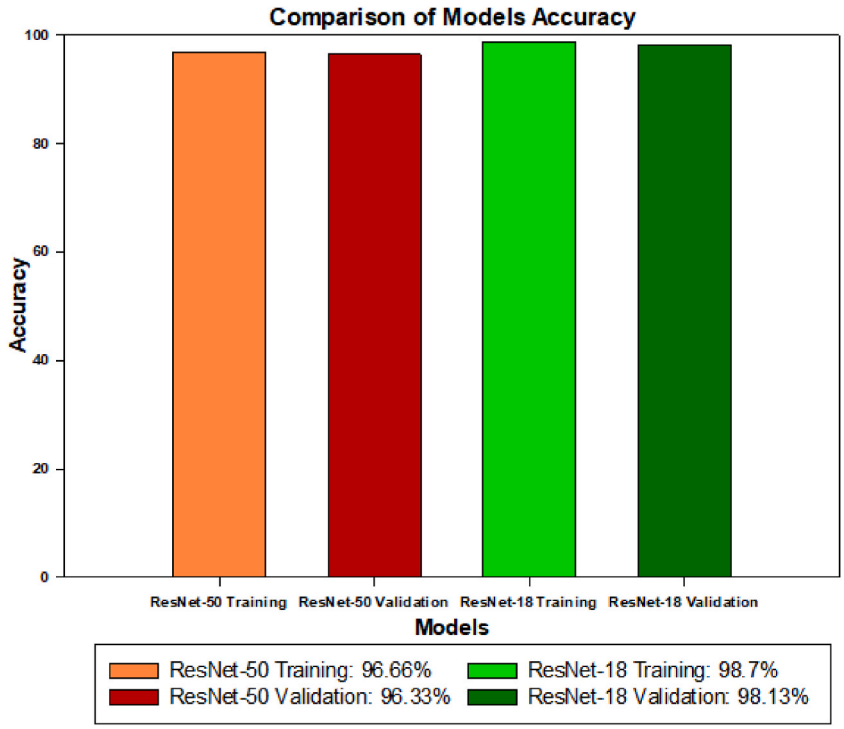


Fig. 14. The comparison of the ResNet-50 and ResNet-18 models' accuracy.

5. Conclusion

This research aimed to develop an effective model for detecting and counting colorectal cancer (HT-29) cells in images, employing an integrated approach that combines YOLOv2 with ResNet-50 and ResNet-18 architectures. The primary objective was to enhance the accuracy and reliability of cell detection, addressing a crucial aspect of medical image analysis in the context of colorectal cancer

Table 7
Comparison with previous studies based on residual networks.

Study	Goal	Methodology	Model/Algorithm	Dataset	Results	Accuracy reasons
[60]	Breast cancer diagnosis using ResNet-50 CNN	Histopathological image classification	ResNet-50 (Transfer Learning)	BreakHis dataset	99 % accuracy	Transfer learning, ResNet-50 architecture
[42]	Detection of colorectal cancer with ResNet	Image classification	ResNet-18, ResNet-50	Colon glands images	>80 % accuracy, >87 % sensitivity, >83 % specificity	ResNet-50 outperforms ResNet-18
[61]	Classification of skin cancer, brain tumor, breast cancer, and colon cancer	Image classification	Proposed CNN model vs. ResNet-50	Multiple datasets	Improved accuracy in every category	Proposed model superiority
[62]	Detection of cancer from histopathology images	ML with CNN ResNet-50	ResNet-50	Breast Cancer Wisconsin dataset, Breast Histopathology images dataset	91.7 % accuracy	Training CNN with ResNet-50 architecture
[63]	Breast cancer diagnosis using ResNet-50	Computational breast cancer diagnosis	ResNet-50 (Transfer Learning)	BreakHis dataset	99 % accuracy	Transfer learning, ResNet-50 architecture
[64]	Skin cancer classification using ResNet-50	Image classification	ResNet-50, VGG16, CNN	HAM10000 dataset	84.87 % accuracy, Precision 0.86, Recall 0.85, F1-score 0.85	The high predictive capability of ResNet-50
[65]	Breast abnormality classification	Mammographic imaging	ResNet-50 with SCDA augmentation	INbreast and MINI-MIAS datasets	95.74 % accuracy (8.55 % specificity, 92.83 % sensitivity)	SCDA augmentation method
[66]	Prediction of primary tumor sites in spinal metastases	MRI-based ResNet-50 CNN	ResNet-50 CNN	Spinal metastases patients	52.97 % top-1 accuracy (AUC-ROC: 0.77)	Potential to prioritize examinations and treatments
[67]	Melanoma cancer classification	Image classification	ResNet-50, ResNet-40, ResNet-25, ResNet-10, ResNet-7	ISIC 2018 dataset	Best model: ResNet-50 without augmentation (Validation accuracy: 0.83, F1 score: 0.46)	Best performance among ResNet models
[68]	Brain tumor classification	Multi-class classification	ResNet-50 (Transfer Learning)	Brain tumor datasets	Accuracy: 95.3 %	Transfer learning with pre-trained ResNet-50
[69]	Colon cancer identification	Image classification	ResNet-18 vs. Squeezenet	30 samples	ResNet-18 accuracy: 86.51 %	ResNet-18 outperforms Squeezenet
[70]	IoT framework for brain tumor detection	Brain tumor detection	Optimized modified ResNet-18 (OMRES)	Not specified	Accuracy: 98.67 %	RMSProp algorithm with a dropout rate of 0.5
[71]	Medical image retrieval using ResNet-18	Image retrieval	ResNet-18	Multi-modality dataset	Classification accuracy: 92 % (mAP: 0.90)	Retrieval based on pre-trained ResNet-18
[72]	Breast cancer detection from ultrasound images	Breast cancer detection	ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152	BUSI dataset	Varied metrics evaluated	ResNets superiority, GradCAM++ visualizations
[73]	Pancreas segmentation using ResNet-18	Pancreas segmentation	ResNet-18	CT scan images	Dice similarity index: 98.29, Jaccard Index: 96.63, Bfscore: 84.65	DL outperforms traditional methods
Our study	Colorectal cancer cell (HT-29) detection and counting	Cell counting using YOLOv2 and ResNet-50,18	YOLOv2, ResNet-18, ResNet-50	Colorectal cancer cell images	ResNet-18: 98.70 %, ResNet-50: 96.66 % accuracy	YOLOv2 and ResNet hybrid architecture, dataset type, the influence of training hyperparameters, and single class for accurate cell counting

research. The ResNet-50 and ResNet-18 models were then trained using a carefully curated dataset, and their performances were rigorously evaluated. The results revealed that the ResNet-18 model consistently outperformed ResNet-50 regarding precision, recall, and overall accuracy during the training and testing phases. This unexpected outcome could be attributed to various factors, including the specific characteristics of the dataset, the sensitivity of each architecture to different data distributions, and the chosen hyperparameters during training. The novelty of this research lies in its exploration of integrated models and the detailed evaluation of their performances in the specific context of HT-29 cell detection. This study contributes valuable insights to healthcare and cancer research by employing cutting-edge DL techniques. The ResNet model's good accuracy makes it a robust tool for precise HT-29 cell detection. This research explores the technical aspects of model development, emphasizing the intricate interplay among architecture, dataset characteristics, and hyperparameter choices. The findings highlight the significance of empirical testing and context-specific considerations in crafting resilient models for medical image analysis, ultimately advancing cancer detection and healthcare technologies. In future work, we aim to further investigate these models using different datasets, more classes, and hyperparameters.

Funding

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Ethics approval and consent to participate

This study does not involve patient data. The HT-29 cell line was bought from a company (Procell Life Science & Technology Co., Ltd. Wuhan, China). Therefore, it does not require ethical approval and number.

Consent for publication

The manuscript has consensus from all authors for publication.

Availability of data and materials

We regret to inform that the dataset underlying this research cannot be made publicly available due to limitations imposed by our lab (institution). However, inquiries regarding specific aspects of the data should be directed to Inayatul Haq (inayatulhaq1@gmail.com) upon reasonable request.

CRediT authorship contribution statement

Inayatul Haq: Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Tehseen Mazhar:** Validation, Software, Methodology, Investigation, Formal analysis. **Rizwana Naz Asif:** Validation, Software, Resources, Methodology. **Yazeed Yasin Ghadi:** Visualization, Validation, Resources, Investigation, Funding acquisition. **Najib Ullah:** Visualization, Validation, Resources, Investigation, Conceptualization. **Muhammad Amir Khan:** Visualization, Validation, Resources, Investigation, Formal analysis. **Amal Al-Rasheed:** Visualization, Validation, Investigation, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We express sincere gratitude to Princess Nourah bint Abdulrahman University (PNUR) for their pivotal support through Researchers Supporting Project (PNURSP2023R235). Their financial assistance, including covering Article Processing Charges (APC), significantly contributed to the success of this study, enhancing the overall quality of our research.

References

[1] D. Crosby, et al., Early detection of cancer, Science 375 (6586) (2022) eaay9040.
[2] E. Medico, et al., The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets, Nat. Commun. 6 (1) (2015) 7002.
[3] J.Y. Buikhuisen, A. Torang, J.P. Medema, Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges, Oncogenesis 9 (7) (2020) 66.
[4] M. Borre, et al., Diet and bowel symptoms among colon cancer survivors, Acta Oncol. 61 (10) (2022) 1192–1199.
[5] M. Ahmed, Colon cancer: a clinician's perspective in 2019, Gastroenterol. Res. 13 (1) (2020) 1–10.
[6] I. Pacal, et al., A comprehensive review of deep learning in colon cancer, Comput. Biol. Med. 126 (2020) 104003.
[7] Society, AC colorectal cancer [cited 2023. 31 May,]; Available from: https://www.cancer.org/cancer/types/colon-rectal-cancer.html, 2023.
[8] Y. Xi, P. Xu, Global colorectal cancer burden in 2020 and projections to 2040, Translational oncology 14 (10) (2021) 101174.

[9] Center, M.S.K.C. *HT*-29: human colorectal adenocarcinoma cell line (ATCC HTB-38) [cited 2023. 31 May,]; Available from: https://www.mskcc.org/research-advantage/support/technology/tangible-material/human-colorectal-adenocarcinoma-cell-line-ht-29, 2023.
[10] N. Bie, et al., A polysaccharide from Grifola frondosa fruit body induces HT-29 cells apoptosis by PI3K/AKT-MAPKs and NF-κB-pathway, Int. J. Biol. Macromol. 147 (2020) 79–88.
[11] S. Basu, Colon cancer, 03 Mar 2021 [cited 2023 29 November,]; Available from: https://www.netmeds.com/health-library/post/colon-cancer-causes-symptoms-and-treatment, 2021.
[12] M.T. Haq I, R. Naz Asif, Y. Yasin Ghadi, R. Saleem, F. Mallek, H. Hamam, A deep learning approach for the detection and counting of colon cancer cells (HT-29 cells) bunches and impurities, PeerJ Computer Science 9 (2023).
[13] P.C. Shathviha, et al., β-sitosterol mediated silver nanoparticles induce cytotoxicity in human colon cancer HT-29 cells, Avicenna J. Med. Biotechnol. (AJMB) 13 (1) (2021) 42.
[14] F. Li, et al., Antitumor mechanisms of an exopolysaccharide from Lactobacillus fermentum on HT-29 cells and HT-29 tumor-bearing mice, Int. J. Biol. Macromol. 209 (2022) 552–562.
[15] Y. Chen, et al., Luteolin induces pyroptosis in HT-29 cells by activating the Caspase1/Gasdermin D signalling pathway, Front. Pharmacol. 13 (2022) 952587.
[16] A. Khodavirdipour, R. Zarean, R. Safaralizadeh, Evaluation of the anti-cancer effect of Syzygium cumini ethanolic extract on HT-29 colorectal cell line, J. Gastrointest. Cancer 52 (2021) 575–581.
[17] Y. Guo, et al., Lactobacillus acidophilus CICC 6074 inhibits growth and induces apoptosis in colorectal cancer cells in vitro and in HT-29 cells induced-mouse model, J. Funct.Foods 75 (2020) 104290.
[18] N. Ilyas, A. Shahzad, K. Kim, Convolutional-neural network-based image crowd counting: review, categorization, analysis, and performance evaluation, Sensors 20 (1) (2019) 43.
[19] V. Ranjan, H. Le, M. Hoai, Iterative crowd counting, in: Proceedings of the European Conference on Computer Vision, ECCV), 2018.
[20] R. Alkadi, et al., A deep learning-based approach for the detection and localization of prostate cancer in T2 magnetic resonance images, J. Digit. Imag. 32 (2019) 793–807.
[21] R. Buenrostro-Mariscal, et al., A review of deep learning applications for the next generation of cognitive networks, Appl. Sci. 12 (12) (2022) 6262.
[22] J. Guerrero-Ibañez, J. Contreras-Castillo, S. Zeadally, Deep learning support for intelligent transportation systems, Transactions on Emerging Telecommunications Technologies 32 (3) (2021) e4169.
[23] F. Montanari, R. German, A. Djanatliev, Pattern recognition for driving scenario detection in real driving data, in: 2020 IEEE Intelligent Vehicles Symposium (IV), IEEE, 2020.
[24] Y. Li, et al., Medical image fusion method by deep learning, International Journal of Cognitive Computing in Engineering 2 (2021) 21–29.
[25] M.E. Salman, et al., Automated prostate cancer grading and diagnosis system using deep learning-based Yolo object detection algorithm, Expert Syst. Appl. 201 (2022) 117148.
[26] N. Aishwarya, et al., Skin cancer diagnosis with yolo deep neural network, Procedia Comput. Sci. 220 (2023) 651–658.
[27] H. Tahir, M.S. Khan, M.O. Tariq, Performance analysis and comparison of faster R-CNN, mask R-CNN and ResNet50 for the detection and counting of vehicles, in: 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE, 2021.
[28] S.-D. Yang, et al., Class-agnostic few-shot object counting, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021.
[29] M. Hussain, YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection, Machines 11 (7) (2023) 677.
[30] Y. Yin, H. Li, W. Fu, Faster-Yolo, An accurate and faster object detection method, Digit. Signal Process. 102 (2020) 102756.
[31] R. Gai, N. Chen, H. Yuan, A detection algorithm for cherry fruits based on the improved YOLO-v4 model, Neural Comput. Appl. 35 (19) (2023) 13895–13906.
[32] S. Alam, et al., A Comparative Study of Object Detection Models for Real Time Application in Surveillance Systems, Brac University, 2022.
[33] B. Selcuk, T. Serif, Brain tumor detection and localization with YOLOv8, in: 2023 8th International Conference on Computer Science and Engineering (UBMK), IEEE, 2023.
[34] Z. Ji, et al., ELCT-YOLO: an efficient one-stage model for automatic lung tumor detection based on CT images, Mathematics 11 (10) (2023) 2344.
[35] V. Blanes-Vidal, G. Baatrup, E.S. Nadimi, Addressing priority challenges in the detection and assessment of colorectal polyps from capsule endoscopy and colonoscopy in colorectal cancer screening using machine learning, Acta Oncol. 58 (sup1) (2019) S29–S36.
[36] Y. Zeng, et al., Real-time colorectal cancer diagnosis using PR-OCT with deep learning, in: Optical Coherence Tomography, Optica Publishing Group, 2020.
[37] M. Masud, et al., A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework, Sensors 21 (3) (2021) 748.
[38] I. Haq, et al., Machine vision approach for diagnosing tuberculosis (TB) based on computerized tomography (CT) scan images, Symmetry 14 (10) (2022) 1997.
[39] I. Haq, et al., A novel brain tumor detection and coloring technique from 2D MRI images, Appl. Sci. 12 (11) (2022) 5744.
[40] I. Haq, et al., Lung nodules localization and report analysis from computerized tomography (CT) scan using a novel machine learning approach, Appl. Sci. 12 (24) (2022) 12614.
[41] J.N. Kather, et al., Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study, PLoS Med. 16 (1) (2019) e1002730.
[42] D. Sarwinda, et al., Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer, Procedia Comput. Sci. 179 (2021) 423–431.
[43] A.B. Hamida, et al., Deep learning for colon cancer histopathological images analysis, Comput. Biol. Med. 136 (2021) 104730.
[44] A.M. Godkhindi, R.M. Gowda, Automated detection of polyps in CT colonography images using deep learning algorithms in colon cancer diagnosis, in: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), IEEE, 2017.
[45] T. Babu, et al., Colon cancer prediction on histological images using deep learning features and Bayesian optimized SVM, J. Intell. Fuzzy Syst. 41 (5) (2021) 5275–5286.
[46] Z. Du, et al., Label-free detection and enumeration of rare circulating tumor cells by bright-field image cytometry and multi-frame image correlation analysis, Lab Chip 22 (18) (2022) 3390–3401.
[47] X. Jin, et al., The overexpression of IQGAP1 and β-catenin is associated with tumor progression in hepatocellular carcinoma in vitro and in vivo, PLoS One 10 (8) (2015) e0133770.
[48] B. Chen, et al., STC2 promotes the epithelial-mesenchymal transition of colorectal cancer cells through AKT-ERK signaling pathways, Oncotarget 7 (44) (2016) 71400.
[49] S. Mukherjee, Explaining how ResNet-50 works and why it is so popular [cited 2023. 01 June,]; Available from: https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758, 2022.
[50] P. Kumar, B. Swaminathan, U. Karthikeyan, A novel deep learning model based on Yolo-V2 and Resnet for pedestrian detection, Annals of the Romanian Society for Cell Biology (2021) 2258–2268, 2258–2268.
[51] Y. Zhang, et al., Single-image crowd counting via multi-column convolutional neural network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[52] V.A. Sindagi, V.M. Patel, Cnn-based cascaded multi-task learning of high-level prior and density estimation for crowd counting, in: 2017 *14th* IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), IEEE, 2017.
[53] H. Idrees, I. Saleemi, M. Shah, Multi-source, Multi-Scale Counting in Dense Crowd Images, Google Patents, 2018.
[54] D. Onoro-Rubio, R.J. López-Sastre, Towards perspective-free object counting with deep learning, in: Computer Vision–ECCV 2016: 14th European Conference, Springer, Amsterdam, The Netherlands, 2016. October 11–14, 2016, Proceedings, Part VII 14.
[55] D. Babu Sam, S. Surya, R. Venkatesh Babu, Switching convolutional neural network for crowd counting, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[56] K. He, et al., Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[57] G. Huang, et al., Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[58] M. Tan, Q. Le Efficientnet, Rethinking model scaling for convolutional neural networks, in: International Conference on Machine Learning, PMLR, 2019.
[59] S. Zagoruyko, N. Komodakis, s, Wide residual network (2016) arXiv preprint arXiv:1605.07146.
[60] Q.A. Al-Haija, A. Adebanjo, Breast cancer diagnosis in histopathological images using ResNet-50 convolutional neural network, in: 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), IEEE, 2020.
[61] P.N.R. Bodavarapu, et al., Optimized deep neural model for cancer detection and classification over ResNet, in: Smart Technologies in Data Science and Communication: Proceedings of SMART-DSC 2021, Springer, 2021.
[62] S.A. Shadab, et al., Detection of cancer from histopathology medical image data using ML with CNN ResNet-50 architecture, in: Computational Intelligence in Healthcare Applications, Elsevier, 2022, pp. 237–254.
[63] Q.A. Al-Haija, G.F. Manasra, Development of breast cancer detection model using transfer learning of residual neural network (resnet-50), American Journal of Science & Engineering 1 (3) (2020) 30–39.
[64] A. Mehra, et al., Skin cancer classification through transfer learning using ResNet-50, Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2020 2 (2021) 55–62. Springer.
[65] X. Yu, et al., ResNet-SCDA-50 for breast abnormality classification, IEEE ACM Trans. Comput. Biol. Bioinf 18 (1) (2020) 94–102.
[66] K. Liu, et al., Prediction of primary tumor sites in spinal metastases using a ResNet-50 convolutional neural network based on MRI, Cancers 15 (11) (2023) 2974.
[67] A. Budhiman, S. Suyanto, A. Arifianto, Melanoma cancer classification using resnet with data augmentation, in: 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, 2019.
[68] M.B. Sahaai, et al., ResNet-50 based deep neural network using transfer learning for brain tumor classification, in: AIP Conference Proceedings., AIP Publishing, 2022.
[69] K. Vasu, Effective classification of colon cancer using resnet-18 in comparison with squeezenet, J. Pharm. Negat. Results (2022) 1413–1421.
[70] S.A. El-Feshawy, et al., IoT framework for brain tumor detection based on optimized modified ResNet 18 (OMRES), J. Supercomput. 79 (1) (2023) 1081–1110.
[71] S. Ayyachamy, et al., Medical image retrieval using Resnet-18, in: Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications, SPIE, 2019.
[72] A. Das, S. Rana, Exploring residual networks for breast cancer detection from ultrasound images, in: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, 2021.
[73] S. Kakarwal, P. Paithane, Automatic pancreas segmentation using ResNet-18 deep learning approach, System research and information technologies (2) (2022) 104–116.