Detection of live breast cancer cells in bright-field microscopy images containing white blood cells by image analysis and deep learning

Golnaz Moallem,[a] Adity A. Pore,[b] Anirudh Gangadhar,[b] Hamed Sari-Sarraf,[a],* and Siva A. Vanapalli[b],*

[a]Texas Tech University, Department of Electrical and Computer Engineering, Lubbock, Texas, United States
[b]Texas Tech University, Department of Chemical Engineering, Lubbock, Texas, United States

Abstract

Significance: Circulating tumor cells (CTCs) are important biomarkers for cancer management. Isolated CTCs from blood are stained to detect and enumerate CTCs. However, the staining process is laborious and moreover makes CTCs unsuitable for drug testing and molecular characterization.

Aim: The goal is to develop and test deep learning (DL) approaches to detect unstained breast cancer cells in bright-field microscopy images that contain white blood cells (WBCs).

Approach: We tested two convolutional neural network (CNN) approaches. The first approach allows investigation of the prominent features extracted by CNN to discriminate *in vitro* cancer cells from WBCs. The second approach is based on faster region-based convolutional neural network (Faster R-CNN).

Results: Both approaches detected cancer cells with higher than 95% sensitivity and 99% specificity with the Faster R-CNN being more efficient and suitable for deployment presenting an improvement of 4% in sensitivity. The distinctive feature that CNN uses for discrimination is cell size, however, in the absence of size difference, the CNN was found to be capable of learning other features. The Faster R-CNN was found to be robust with respect to intensity and contrast image transformations.

Conclusions: CNN-based DL approaches could be potentially applied to detect patient-derived CTCs from images of blood samples.

1 Introduction

Breast cancer is one of the leading causes of death with one in eight women expected to be diagnosed with invasive breast cancer in the United States.[1] An estimated 284,200 new cases are expected to be diagnosed in 2021 alone with about 44,130 deaths due to breast cancer.[2] A majority of deaths in breast cancer patients are due to metastasis, where circulating tumor cells (CTCs) shed from the primary tumor, spread to distant organs, and form secondary tumors.[3,4] Isolating, enumerating, and characterizing these CTCs are therefore valuable means to monitor disease progression and manage treatment therapies for cancer patients.[5–9]

The CTCs are present in exceedingly low counts in peripheral blood (1 to 100 CTC s/ml),[5] which has resulted in significant efforts to develop methods for isolating and characterizing

---

*Address all correspondence to Hamed Sari-Sarraf, Hamed.Sari-Sarraf@ttu.edu; Siva A. Vanapalli, Siva.Vanapalli@ttu.edu

these rare cells. Current CTC isolation techniques may be classified into two broad categories: affinity-based (labeled) and label-free techniques. Technologies based on the former approach rely on the interaction of cell-surface receptors on CTCs and specific antibodies, leading to the isolation of only the CTCs that bind to these antibodies.[10–13] In contrast, label-free approaches isolate CTCs based on physical characteristics, e.g., size,[14–16] deformability,[17] dielectric properties,[18,19] and density.[20]

Despite the development of numerous separation methods for isolating CTCs from whole blood, the enriched CTC samples are almost always accompanied by a large number of white blood cells (WBCs).[21–26] The presence of these background nucleated cells makes it necessary to use molecular markers and immunostaining to distinguish CTCs from WBCs. However, the immunostaining process kills CTCs; thus, making it impossible to use them in downstream assays such as expansion studies or single-cell transcriptomics, all of which require the CTCs to be alive and functional.[27–29] Thus, there is a need for staining-free approaches that can enumerate live CTCs in a background of WBCs.

Current approaches for detecting stained CTCs range from manual scoring to techniques based on image processing and machine learning. There is a growing interest in using machine learning to detect and classify cells in patient blood samples as it eliminates drawbacks of manual scoring and threshold-based object detection.[30–32] For example, recently, Zeune et al.[33] developed a deep learning (DL) network to automatically identify and classify fluorescently stained CTCs obtained from the blood of metastatic cancer patients with a reported accuracy of over 96%. DL networks are attractive as they do not require any feature engineering but are capable of automatically discovering the best set of features that discriminate the cells of interest and can directly identify them from raw input images.[34]

Given the success of DL models to identify CTCs in stained images, new efforts are emerging to implement similar approaches in unstained samples. However, since CTCs are quite heterogeneous and are present in low counts in patient samples, generating ground truth data is challenging. Thus, as a first step, DL-based frameworks are being pursued where *in vitro* cancer cells are mixed with blood cells, and the efficacy of DL models is being investigated to detect live cancer cells in a background of blood cells[35–38] As shown in Table 1, limited studies have been conducted on DL-based frameworks for cancer cell identification from a background of blood cells, with images acquired using quantitative phase, dark-field, and bright-field microscopy. A notable study of this kind is the study by Wang et al.,[35] who demonstrated detection of live CTCs in the blood of renal cancer patients as well as *in vitro* human colorectal cancer cells (HCT-116) using a deep convolutional neural network (CNN) and reported an accuracy of 88.6%.

In this study, we utilize CNN approaches to demonstrate an automatic, sophisticated detection model for Michigan cancer foundation-7 (MCF-7) breast cancer cells in bright-field images of mixed population samples containing MCF-7 cells and WBCs. We develop two CNN models: (i) the first approach referred to as the "decoupled cell detection" allows us to investigate the prominent features extracted by CNN to discriminate cancer cells from WBCs. To the best of our knowledge, the features CNN uses to discriminate cells have not been explained by previous studies. (ii) The second approach is based on faster region-based convolutional neural network[39] referred to as the Faster R-CNN, which is more efficient and therefore suitable for deployment. Additionally, we introduce a novel automatic technique to generate the training set for the Faster R-CNN approach. This algorithm employs the bright-field image and the associated fluorescent images as ground truth. Additionally, we investigate the effect of image transformation on the classification accuracy to assess how well the DL model performs when variability is introduced in image quality. Finally, we discuss avenues for moving beyond the model system studied here and the challenges that need to be addressed to apply DL techniques to live CTC detection in patient blood samples.

2 Sample Preparation, Image Acquisition, and Data Sets

In this section, we discuss the methods for culturing cancer cells, isolating WBCs, and how we generated the image data sets of pure cell populations and mixed cell populations. We also clearly outline the image data sets used in developing our machine learning methods.

Table 1 Summary of studies that applied DL methods for the identification of unstained *in vitro* cancer cells in a background of blood cells. The performance measures shown for our work correspond to the Faster R-CNN model.

Study	Cancer cells	Background cells	Test images	Ground truth images	Classification/ detection model used	Accuracy/sensitivity/specificity/precision
Chen et al.[36]	Colon (SW-480)	T-cells (WBC subset)	Quantitative phase images	Quantitative phase images	DNN classifier trained on AUC	95.5%
						Not reported
						Not reported
						Not reported
Ciurte et al.[37]	Breast (Hs 578T, MCF-7), Colon (DLD-1)	Blood cells (RBCs + WBCs)	Dark-field images	Fluorescent images	Adaboost classifier	Not reported
						92.87%
						99.98%
						Not reported
Wang et al.[35]	Colorectal (HCT-116), metastatic renal cell carcinoma CTCs	WBCs	Bright-field images	Fluorescent images	ResNet50 CNN classifier	88.6% for patient blood sample images and 97% for cultured cell images
						Not reported
						Not reported
						Not reported
This work	Breast (MCF-7)	WBCs	Bright-field images	Fluorescent images	Custom CNN classifier, Faster R-CNN detection model	99.45%
						99.1%
						Not applicable
						99.8%

2.1 Cell Culture

The breast cancer cell line MCF-7 was obtained from the American Type Cell Collection (Manassas, Virginia). MCF-7 cells were cultured in Dulbecco's Modified Eagle Medium (Gibco, Gaithersburg, Maryland) containing 10% fetal bovine serum (Gibco, Gaithersburg, Maryland), 1% penicillin/streptomycin (Gibco, Gaithersburg, Maryland) and 1% sodium pyruvate (Gibco, Gaithersburg, Maryland). The cell culture media were prepared according to the manufacturer's protocol. The cells were incubated at 37 in a 5% CO_2 environment.

2.2 MCF-7 Cell Staining

The cells were labeled with CellTracker™ Green CMFDA (Invitrogen™, Waltham, Massachusetts). The dye was prepared according to the manufacturer's protocol to make a stock solution of 10 mM. A working solution of 1 μM was made by diluting the stock solution in serum-free media. This working solution was added to the tissue culture flask containing cells followed by an incubation step for 45 min at 37. After incubation, the cells were washed with phosphate buffer saline (Gibco, Gaithersburg, Maryland) three times to get rid of any excess dye. The cells were further trypsinized, resuspended in fresh media, and stained with the nuclear stain 4′,6-diamidino-2-phenylindole (DAPI, Molecular Probes, Eugene, Oregon). A working concentration of 100,000 cells/ml was used for the experiments.

2.3 WBC Isolation and Staining

Normal human whole blood was ordered from BioIVT (Hicksville, New York) for isolation of WBCs from whole blood. 1 ml of whole blood was lysed using ammonium–chloride–potassium (ACK) lysing buffer (Quality Biological, Gaithersburg, Maryland) for isolation of WBCs. WBCs were stained with the nuclear stain DAPI. A working concentration of 3.6×10^6 cells/ml was used for the experiments.
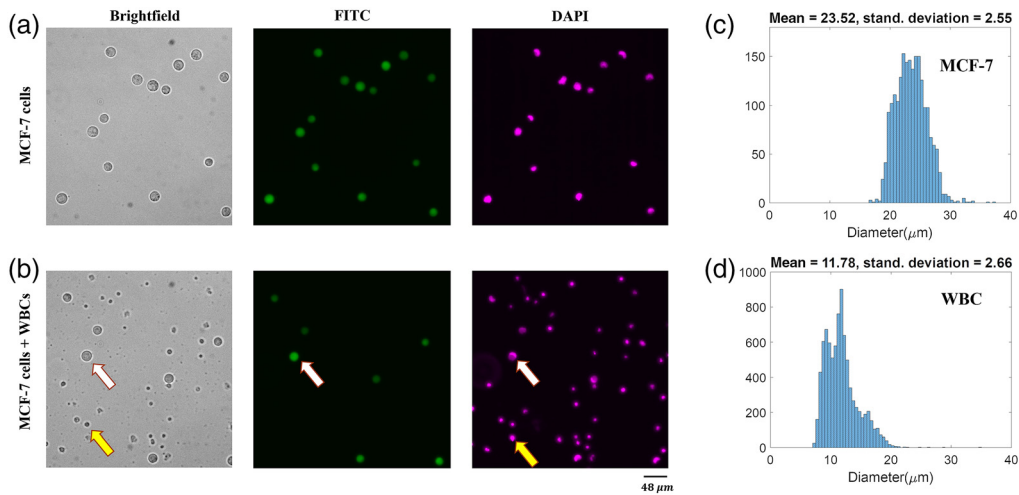
2.4 Pure Cell Population Images

Slides of MCF-7 cells were prepared by adding 10 μl of working solution of MCF-7 cells on the cover glass (Richard-Allan Scientific, Kalamazoo, Michigan) equipped with 10 mm × 24 mm spacers on the edges. This solution was then sandwiched using another cover glass and imaged using Olympus IX81 microscope (Massachusetts). The microscope is equipped with a Thorlabs automated stage (New Jersey) and a Hamamatsu digital camera (ImagEM X2 EM-CCD, New Jersey) controlled by SlideBook 6.1 (3i Intelligent Imaging Innovations Inc., Denver). Bright-field and fluorescence images were acquired using 20× objective (0.8 μm/pixel, 512 × 512 pixels) with the fluorescent filters DAPI and fluorescein isothiocyanate (FITC). Exposure times between 30 and 200 ms were used for image acquisition.

Similarly, slides of WBCs were prepared by adding 10 μl of working solution of WBCs on the sandwiched cover glass equipped with spacers. Bright-field and fluorescence images were acquired using 20× objective (0.8 μm/pixel, 512 × 512 pixels) and the fluorescent filter DAPI. Exposure times use for the WBC pure cell training set image acquisition were 30 to 100 ms.

2.5 Mixed Cell Population Images

Slides for imaging of mixed cell populations were prepared by mixing 5 μl of WBC working solution with 5 μl of MCF-7 working solution. This 10 μl solution was added to the cover glass and imaged using the similar sandwich method as described previously.[40,41] Bright-field and fluorescent (DAPI, FITC) image acquisition was done at 20× magnification (0.8 μm/pixel, 512 × 512 pixels). Images were acquired between 30- to 200-ms exposure times. Cell body and nucleus of MCF-7 cells were fluorescently labeled using CellTracker™ Green CMFDA and DAPI, respectively, whereas those of WBCs were only labeled with DAPI, enabling distinction of the two cell types under fluorescence imaging.

Fig. 1 Representative images and size distribution of MCF-7 cancer cells and WBCs. (a) Bright-field and the corresponding fluorescent images of live MCF-7 cells. (b) Bright-field and corresponding fluorescent images of a mixed-cell population of MCF-7s and WBCs. White arrows denote an MCF-7 cell, which can be co-localized in the ground truth fluorescent FITC and DAPI images. Yellow arrows denote a WBC, which can be co-localized in the DAPI image but not in the FITC image. (c) Size distribution of MCF-7 cells, $n = 2071$. (d) Size distribution of WBCs, $n = 8575$.

2.6 *Image Data Sets*

The bright-field images of MCF-7s and WBCs were accompanied by ground truth DAPI and FITC fluorescent images for exploration, training, and evaluation of DL models used in this study. We acquired 22 pure MCF-7 cell images and 21 pure WBC images totaling 1700 MCF-7 cells and 11,518 WBCs. For the mixed cell population, we acquired 592 images that contained 4722 MCF-7 cells and 12,314 WBC cells. Figures 1(a) and 1(b) show examples of pure cell and mixed cell bright-field images along with their associated fluorescent counterparts. Figure 2 shows the details and statistics of the image sets that are used in the following sections.

3 Results and Discussion

In this study, we investigated two CNN-based approaches to detect MCF-7 cancer cells from a background of WBCs. In the first approach, referred to as decoupled cell detection, all cells are localized using the maximally stable extremal regions (MSER) algorithm,[42] and then classified using a trained CNN, i.e., the cell localization and identification tasks are carried out in two distinct steps. The benefit of the decoupled approach is that it allows investigation of the distinctive features that convolutional layers extract to distinguish MCF-7 cells and WBCs.
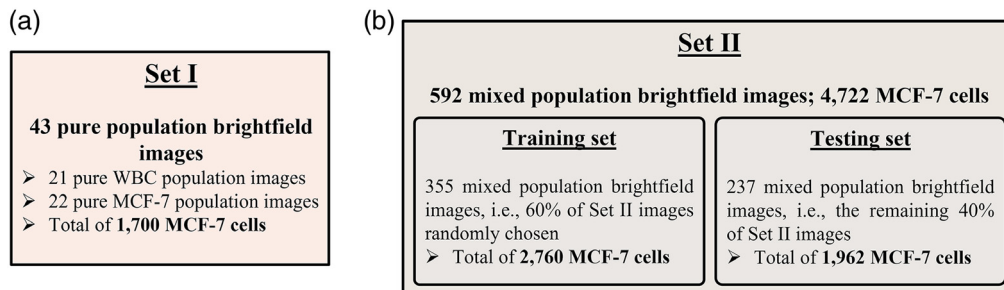
Fig. 2 Image sets used in this study. (a) Set I contains only pure cell population images and is used in developing the framework of the decoupled approach. (b) Set II includes mixed cell population images and is used in the training and testing of the optimized decoupled approach and the Faster R-CNN approach. This set is also used to compare the performance of the two approaches.

Additionally, in the decoupled approach, one can differentiate the localization and classification error rates.

In the second approach, we detect MCF-7 cells using a Faster R-CNN model. In this approach, the cell localization and identification tasks are integrated without having separate access to the output of the localization step or the input of the classification step. As a result, it is difficult with Faster R-CNN model to understand the features being used to discriminate cells. However, due to its speed and efficiency of execution, the Faster R-CNN approach is better suited for deployment.

Below, we discuss the results from our investigations with these two approaches to detect MCF-7 cells in a background of WBCs. We note that the network performance is reported in terms of accuracy, sensitivity, and precision.[43] Here, true positives (TPs) were defined as the number of detected MCF-7 cells. False negatives (FNs) were defined as the number of MCF-7 cells missed in the detection procedure, and false positives (FPs) were defined as the number of detections that did not correspond to MCF-7 cells.

3.1 Decoupled Cell Detection

In this approach, the MCF-7 cells are detected in two decoupled modules: cell localization and cell identification modules. To detect the MCF-7 cells in a given bright-field image: (1) the bright-field image is fed to the localization module to localize the cells using the MSER algorithm, and (2) tiles of localized single cells are provided to the identification module to distinguish MCF-7 cells employing a trained CNN. The architecture and training details of the CNN is discussed in the following sections. As noted earlier, one of the purposes of developing this framework was to study the distinctive features that convolutional layers extract to distinguish MCF-7 cells and WBCs. Therefore, it was essential to have access to the input of the CNN classifier in order to solely study and evaluate its performance.

To develop the decoupled cell detection approach, we first designed and tested a shallow CNN using tiles of individual MCF-7 cells and WBCs. Subsequently, we evaluated the prominent cell features that allows successful classification with this trained CNN. Next, we optimized the CNN architecture to improve the detection of MCF-7 cells. Finally, we applied the decoupled cell detection approach to mixed cell images, i.e., individual images that contained a mixture of MCF-7 cells and WBCs, i.e., Fig. 2(b). The results from this systematic investigation are discussed below.

3.1.1 Shallow CNN efficiently discriminates MCF-7 cells from WBCs

To generate the training and testing set for the shallow CNN, the MSER algorithm was applied to localize the cells in the "pure" bright-field images in set I, i.e., images containing only MCF-7 cells or WBCs. For each localized cell, a 36×36 tile was cropped with the cell at the center. The tile size was determined based on the MCF-7 estimated cell size distribution [Fig. 1(c)] and image resolution. This process ensured that both WBCs and MCF-7 cells were contained entirely in the cropped tiles. The tiles were carefully inspected, verified, and labeled manually employing the FITC and DAPI masks. Those tiles with a complete WBC located in the center of the corresponding tile were labeled "WBC." Similarly, tiles with a complete MCF-7 cell positioned in the central region of the corresponding tile were tagged "MCF7." Set I-training set was formed by creating balanced classes of "WBC" and MCF7 tiles, each class comprising 1190 tiles. Set I-testing set contained 510 tiles per class.

We initially designed and trained a shallow CNN with two convolutional layers followed by a fully connected layer to classify the tiles into WBC and MCF7 categories. The convolutional layers had 5×5 kernel sizes followed by a 2×2 pooling layer to minimize the encoding depth due to the relatively small size of the tiles. The trained CNN was found to achieve satisfactory performance with a training accuracy of 99.51% and a test accuracy of 99.54%. These performance metrics justified the use of the relatively shallow architecture of the CNN as it achieved a high accuracy level without overfitting. In addition, the class probability histograms showed the network's high confidence level of above 99% on classification decisions. We concluded that the trained CNN successfully performed the discrimination task.

3.1.2 *Identifying the discriminatory cellular features*

We investigated the cellular features that enabled the successful classification by the shallow CNN. Since in Figs. 1(c) and 1(d) we observed that the mean diameter of MCF-7 cells is approximately twice as large as the mean diameter of WBCs, we hypothesized that cell size could be a distinguishing feature. To test this hypothesis, we made the size of WBCs similar to MCF-7 cells. This was achieved by doubling the size of set I-testing WBC tiles, which doubled the size of the WBCs contained in them, and then they were cropped to make them compatible with the network input. When we tested the trained CNN, i.e., the model trained on non-modified MCF7 and WBC tiles, on the newly generated WBC tiles, we observed a deterioration of 61% in the classification accuracy, which suggested that the trained CNN indeed relied on the size feature to make the classification decision. In other words, a logical explanation for why the network decided that resized WBCs look less like WBCs and more like MCF7s was because it learned to use size as the main feature of distinction.

Next, we hypothesized that in addition to cell size, the shallow CNN could extract other geometric and photometric features such as texture of the cell to perform the classification task successfully. To test this hypothesis, we eliminated the size difference between the cell types and trained our CNN in two ways. (i) We doubled the size of set I-training WBC tiles and used them together with set I-training MCF7 tiles to train the CNN. In doing so, we achieved a 99.33% training accuracy and a 98.14% testing accuracy. (ii) Alternatively, set I-training MCF7 tiles were halved in size and combined with set I-training WBC tiles to train the CNN. The new network designed to carry out this experiment processed 15×15 (not 36×36) input images. To meet this requirement, MCF7 tiles were simply resized by a factor of 0.5 and WBC tiles were cropped. This resulted in a 98.32% training and a 96.86% test accuracy. Note that in each case, set I-testing tiles were modified the same way as set I-training tiles.

The high training and testing accuracy suggest that although size is a prominent discriminatory feature used by the designed CNN, the network does utilize other features such as texture to perform cell classification with a high, albeit reduced, level of accuracy if necessary. This result helps explain the model's ability to distinguish the two cell types, even when there is an overlap in the cell size distributions [Fig. 1(c) and 1(d)].

3.1.3 *Optimizing the CNN architecture*

The training was carried out on two different CNN architectures, one with two and the other with three convolutional layers while keeping the kernel sizes as 5×5 and pooling layer as 2×2. To boost the TP detection rate of the CNN, we tuned the MSER parameters to maximize the localization of MCF-7 cells at the expense of a proportional increase in the overall number of tiles that were output by the MSER module. The classification task was then recast as a two-class problem in which one class represented MCF-7 cells, and the second represented all non-MCF-7 objects that included WBCs and debris. The results showed that the CNN with three convolutional layers had testing sensitivity and precision of 98.8%, outperformed the CNN with two layers by more than 3%. We therefore used the CNN with three convolution layers in subsequent studies for classification in the decoupled cell detection approach.

So far, our investigations were focused on the pure cell images in set I to develop and optimize a shallow CNN and identify cellular features that enable classification. Next, we trained and evaluated the decoupled cell detection approach, as shown in Fig. 3, on images containing both MCF-7 cells and WBCs, i.e., set II images as described in Fig. 2(b). For training purposes, 60% of set II images were randomly chosen and used, which hereafter will be referred to as the "training images." The remaining 40% of set II images were used to evaluate the developed framework, i.e., the "testing images."

In the cell localization module, i.e., module 1, the MSER algorithm is applied to the input image to localize the objects. The MSERs are those connected regions that do not change in size and morphology as one incrementally thresholds the image over a wide range of thresholding values. There are three parameters to be adjusted for the MSER algorithm: (1) size range, i.e., smallest and largest extremal regions to be selected, (2) stability testing range, i.e., minimum number of thresholding iterations that results in a stable extremal region, and (3) maximum
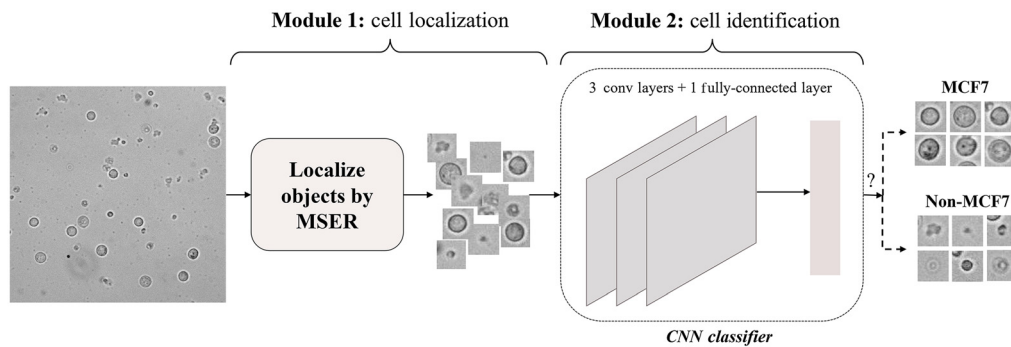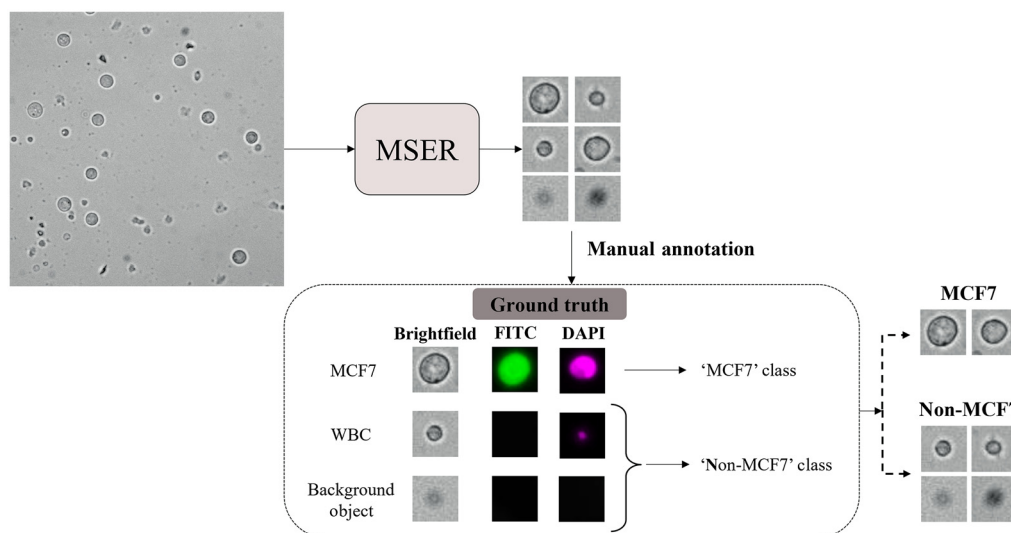
Fig. 3 The framework of the decoupled cell detection approach. As shown, the cell detection is performed in two modules: cell localization and cell identification. The localization module localizes the cells present in the input bright-field image. The CNN-based identification module then identifies the MCF-7 cells among the localized cells.

variation, i.e., the maximum size variation that is allowed for the region to be counted as stable. To detect the MSERs, the image is iteratively binarized at different thresholding levels and connected regions are estimated. The detector selects any connected region that is within the size range and exhibits less area variation than the maximum variation parameter within the specified stability testing range. In this framework, the size range was set based on the lower and upper bounds of the MCF-7 size distribution to maximize the localization of MCF-7 cells. Specifically, this parameter was set to [lower bound -5%, upper bound $+5\%$]. Outputs from the localization module are 36×36 tiles centered on the corresponding objects. The tile size was calculated based on the upper bound of the MCF-7 cell size distribution and image resolution.

In the cell identification module, i.e., Module 2, the cropped tiles are fed to the trained CNN model that, as described above, consists of three convolutional layers and a fully connected layer. The convolutional layers have a 5×5 kernel size, followed by a 2×2 pooling layer. The fully connected layer outputs a two-dimensional vector containing the MCF7 and "non-MCF7" class prediction scores.

The training and testing set label assignment workflow for developing the decoupled cell detection framework is shown in Fig. 4. In this workflow, first, MSER was applied to each input image, and tiles of localized objects were generated. Afterward, the cropped tiles were manually

Fig. 4 Training and testing set label assignment strategy for developing the framework of the decoupled cell detection approach. As demonstrated, the objects in the input image are localized employing the MSER algorithm. Then, the generated tiles of the localized objects are manually categorized based on their corresponding signatures in the FITC and DAPI filters.

categorized based on their corresponding signatures in DAPI and FITC masks. Every localization depicting a signature in both DAPI and FITC masks was labeled as MCF7 since cancer cells were live stained with both a nucleus (DAPI) and a cell body (FITC) marker, whereas tiles lacking any signature in the FITC mask were annotated as non-MCF7. The input to this workflow is the training or testing images, i.e., Fig. 2(b), and the generated training or testing set is a collection of cell tiles, each associated with a label.

We used the training images to generate the training set containing 2760 MCF7 tiles and 2760 tiles in the non-MCF7 class. The CNN was trained using the ADAM optimizer for 35 epochs. The training process took 74 s on an Nvidia TITAN RTX. The trained classification CNN achieved 98.99% training sensitivity and 99.8% training precision.

3.1.4 *Performance of the decoupled cell detection approach*

We evaluated the decoupled approach on the generated testing set, i.e., a set of 1962 MCF7 tiles and 1962 non-MCF7 tiles, using three performance metrics: sensitivity, precision, and average execution time per image as shown in Table 2. Overall, the decoupled approach performed with 95.5% accuracy, 95.3% sensitivity, and 99.8% precision, which indicates its ability to successfully achieve the MCF-7 identification task. Figure 5 shows representative examples of TPs, FPs, and FNs observed during the analysis. In the decoupled approach, MCF-7 cells localized and correctly classified were counted toward TPs [Fig. 5(a)]. The combination of (i) the MCF-7 cells

Table 2 The performance statistics of the two CNN approaches used in this study on the testing images.

| Approach | TPs | FNs | | FPs | | Sensitivity (%) | Precision (%) | Average inference time per image (s) |
		Module1	Module 2	Non-MCF7 objects	Extra detections			
Decoupled Cell Detection	1869	71	22	3	1	95.3	99.8	0.6
Faster R-CNN Cell detection	1945	17		4	0	99.1	99.8	0.3
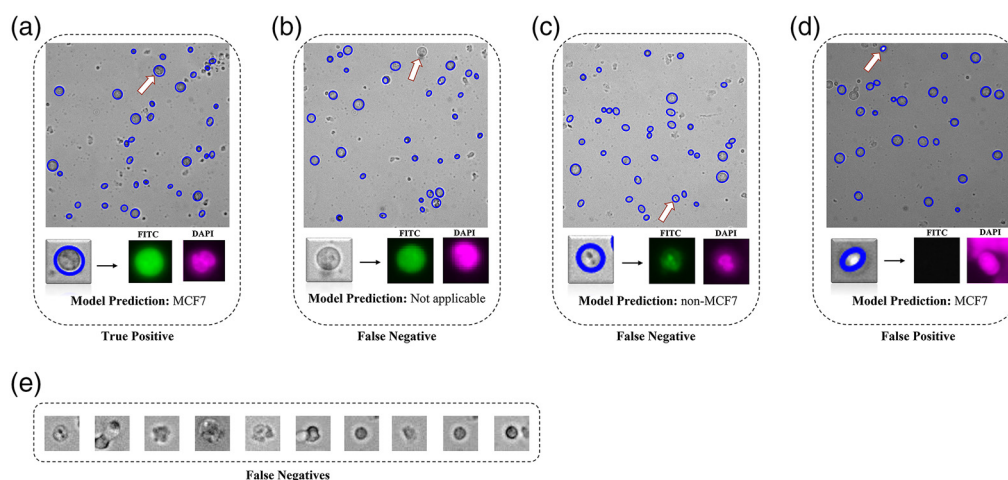
Fig. 5 Representative examples of TP, FP, and FN outcomes from the decoupled cell detection approach. (a) TPs are those MCF-7 cells that are localized and correctly identified as MCF-7 cells by the model. (b) FNs are MCF-7 cells that are not localized and therefore the model is not able to reach the identification stage (c) FNs can also be MCF-7 cells that are localized, but incorrectly classified by the model as non-MCF7. (d) FPs are the non-MCF-7 objects that are localized and classified as MCF7. (e) Additional examples of FNs.

not localized [Fig. 5(b)] and (ii) the MCF7 tiles labeled as "non-MCF7" [Fig. 5(c)] were counted as FNs. Lastly, the non-MCF7 tiles categorized into the MCF7 class were considered as FPs [Fig. 5(d)].

We further investigated the source of FNs. As shown in Table 2, the MSER localization module is unable to localize 71 MCF-7 cells, whereas the identification module misclassifies 22 MCF-7 cells. Therefore, the localization errors are responsible for more than 75% of the FNs. The majority of FNs encountered during identification are MCF-7 cells whose contrast, shape, and size differ from an average MCF-7 cell [see examples in Fig. 5(e)]. These two observations indicate that the convolutional layers of the CNN classifier were able to grasp the common features of MCF-7 cells for identification. As observed, the decoupled nature of this approach enabled us to trace back the missed detections of the CNN classifier and study them to gain insights into the discriminatory power of features that convolution layers extract in classifying MCF-7 cells versus WBCs.

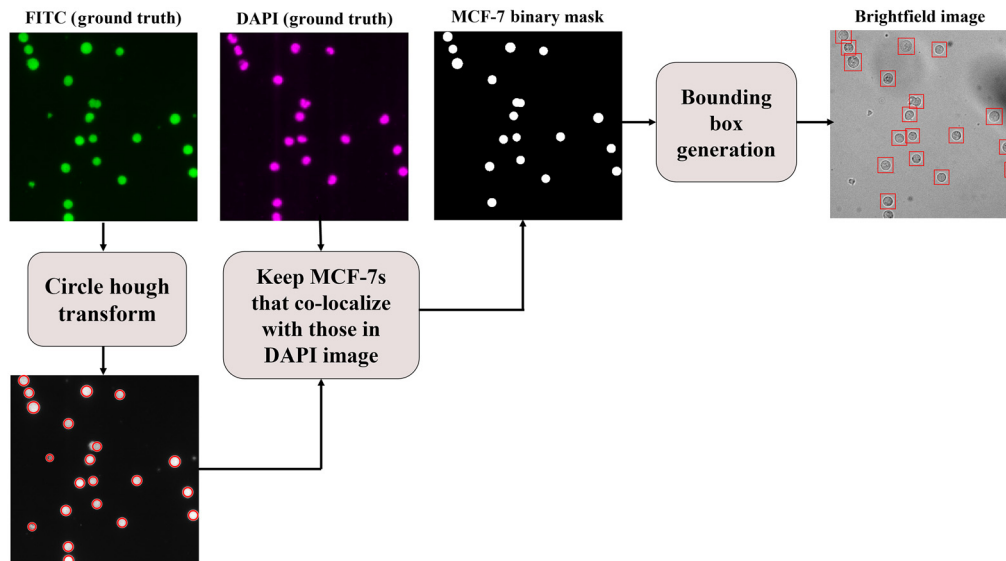### 3.2 Faster R-CNN Cell Detection

The decoupled cell detection approach performs the MCF-7 localization and identification in two steps. However, the sensitivity of this approach is being affected by the large number of FNs due to localization errors. We therefore explored a Faster R-CNN model, which integrates these two steps without having separate access to the output of the localization step or the input of the classification step. The Faster R-CNN model takes a bright-field image as input and outputs a bounding box (BB) for every MCF-7 cell present in the input image (Fig. 6). Two major modules underlie the Faster R-CNN model: (1) a localization module, i.e., a region proposal network (RPN) that discovers regions in the input image that are likely to include an MCF-7 cell; and (2) a classification module, i.e., Fast RCNN that labels the region proposals, which essentially include an MCF-7 cell and estimates a BB that properly confines the cell.

3.2.1 Automated dataset label assignment procedure

For training or evaluating a Faster R-CNN model, every image must be accompanied by a BB associated with each MCF-7 cell present in the image. The manual generation of BBs can be a tedious, subjective, and labor-intensive task. Here, we propose a technique for generating the BBs automatically (see Fig. 7). In this technique, the MCF-7 cell signatures in the FITC mask are first localized utilizing the circular Hough transform algorithm.[44] The FITC mask background is less noisy and complex than its corresponding bright-field image and, therefore, a better option
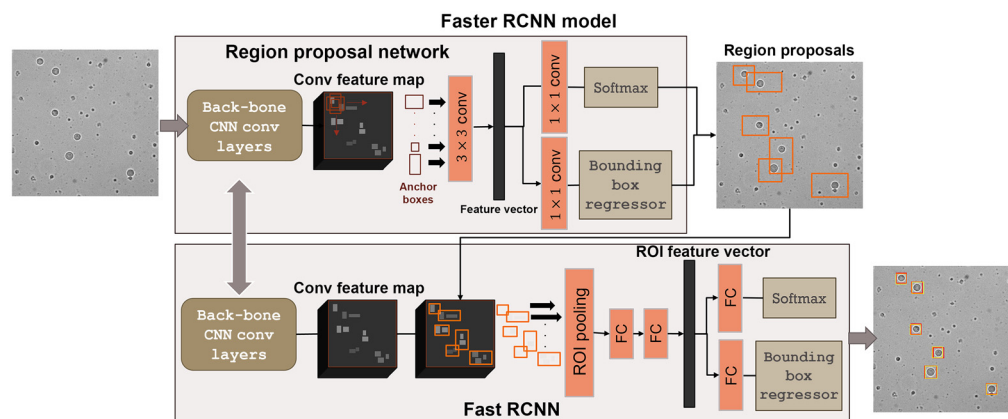
Fig. 6 The Faster R-CNN cell detection approach. The Faster R-CNN detection model consists of two main integrated modules: (i) a RPN that generates the region proposals (indicated by the orange bounding boxes in the image) that are most likely to include an MCF-7 cell, and (2) a Fast RCNN that labels the generated region proposals and outputs a BB for each detected MCF-7 cell. In the output image, the yellow boxes give the estimated bounding boxes assigned to the MCF-7 cell and the red boxes depict the corresponding ground truth.

Fig. 7 Training and testing set label assignment procedure for the Faster R-CNN cell detection approach. For every bright-field image, the MCF-7 cells are first localized in the corresponding FITC mask using the circular Hough transform. Then, each localization, if not touching the image boundaries, is screened in the DAPI mask. If the localized region has a signature in the DAPI mask, then it is added to the MCF-7 binary mask. Finally, a BB is generated for every blob in the MCF-7 binary mask.

to use for localization. For every localization, if not touching the image boundaries, the corresponding region in the DAPI mask is screened. If a signature exists in the screened region, then the corresponding localization is added to the MCF-7 binary mask. Lastly, for every blob in the MCF-7 binary mask, a 36×36 BB is generated and centered on that blob.

With this technique, 613 mixed population and pure cell images were annotated in less than 117 s, which is a significant improvement compared to manual annotation taking more than 100 s per image on average using MATLAB's Image Labeler tool. Additionally, this technique produced equal-sized BBs that are centered on the cells; an outcome that is difficult to achieve manually. Consequently, the proposed automatic dataset label assignment algorithm is a fast, efficient, and consistent technique to generate BB annotations for MCF-7 bright-field images.

For the Faster R-CNN cell detection approach, we utilized the same training images used in the decoupled cell detection approach. Utilizing dataset label assignment framework demonstrated in Fig. 7, we generated a training set comprising the 355 training images with a total of 2760 MCF-7 BBs. We employed the ResNet50[45] as the back-bone CNN for the Faster R-CNN model. The negative and positive overlap ranges were set to [0,0.5] and [0.6,1], respectively. The model was trained for 30 epochs on an Nvidia TITAN RTX in 285 min. The model's performance on the training set indicated an average intersection-over-union (IoU) of 80.3% and a prediction score of 99.6%. The trained model achieved 99.6% training sensitivity and 99.18% training precision.

For the Faster R-CNN detection approach, the TPs, FNs, and FPs were defined in the same way as the decoupled approach, however they were interpreted differently due to the dissimilarities in the mechanism of the two approaches. All MCF-7 cells assigned a BB having IoU > 50% with respect to the annotation BB were counted toward TPs [Fig. 8(a)]. An MCF-7 cell missing an estimated BB was considered a FN [Fig. 8(b)]. There were two kinds of FPs as shown in Figs. 8(c) and 8(d): (1) redundant BBs assigned to a single MCF-7 and (2) non-MCF7 objects assigned a BB.

3.2.2 *Performance comparison of the two approaches*

As discussed in the previous sections, we developed two different frameworks that detect MCF-7 cells in the mixed cell population images. The first developed approach, i.e., decoupled cell
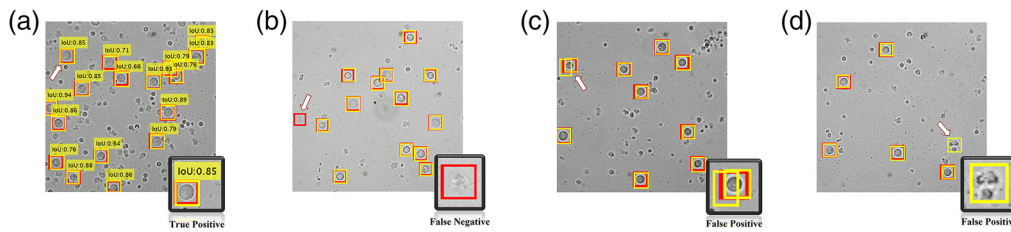
Fig. 8 Representative examples of (a) TP (b) FN and (c) and (d) FP from the Faster R-CNN approach. The arrow in each image and the bottom right inset highlights these examples. The yellow boxes give the estimated bounding boxes assigned to the MCF-7 cell and the red boxes depict the corresponding ground truth.

detection, performed with 95.3% sensitivity and 99.8% precision on testing images. To achieve a better sensitivity rate, we pursued and presented an alternative cell detection approach, i.e., Faster R-CNN detection approach. Table 2 gives the statistical measurements and corresponding performance metrics of the two cell detection approaches on the testing images. Both approaches exhibit a comparable precision, whereas the sensitivity of the Faster R-CNN approach is better than the decoupled approach, indicating that the Faster R-CNN approach is less likely to miss MCF-7 cell detections. It is worth noting that unlike the decoupled approach, we cannot trace back the missed detections (FNs) specifically to the localization or classification modules in the Faster R-CNN model.

Comparing the performance of both approaches indicate that the Faster R-CNN cell detection approach is preferred because it provides better overall performance in detecting MCF-7 cells in roughly half the time. Finally, the cell detection framework proposed by Wang et al.[35] is comparable to our decoupled cell detection approach in that it too is comprised of cell localization and cell identification modules. However, their work does not offer feature analysis, nor do they assess the performance of each module independently. Moreover, our proposed deployed detection model, i.e., Faster R-CNN model, outperforms Wang et al.'s detection framework.

3.3 *Impact of Image Intensity Transformations on the Faster R-CNN Approach*

A major factor determining the effectiveness of a trained DL model is how well it performs when exposed to images with variations in the background intensity and contrast, as well as intensity variations within cells and cell boundaries. Such variations are not uncommon in experimental images due to variability in cell sample preparation and imaging conditions (e.g., focus position and exposure times). To understand the effect of such variations on the performance of Faster R-CNN model, we performed image transformations and assessed the performance of the Faster R-CNN approach on the transformed images. In addition to sensitivity and precision, $F1$-score was used to evaluate the overall performance of the Faster R-CNN detection model in the presence of image transformations. $F1$-score is defined as the harmonic average of recall and precision, evaluating the overall changes in the performance of the detection model, when deployed on a dataset.

We performed image transformations and calculated the aggregate intensity distribution of a dataset by averaging the accumulated intensity distributions of all the images in the corresponding dataset. The intensity histogram of every image, $h(I)$, was altered in the following three ways:

1. Shift the image intensity histogram to generate a brighter version of the image. The offset (B_1) added to every pixel intensity is randomly drawn from a uniform distribution with bounds $[\frac{\text{mean}(h(x))}{4}, \frac{\text{mean}(h(x))}{2}]$. The altered image is modeled with $I_{\text{Int}}(x, y) = I(x, y) + B_1$ with corresponding $h_{\text{Int}}(I)$. This transformation increases the mean of $h(I)$ by 25% to 50%. The upper limit was set to avoid any pixel intensity saturation. The aggregate of all generated $h_{\text{Int}}(I)$ is represented by H_I.
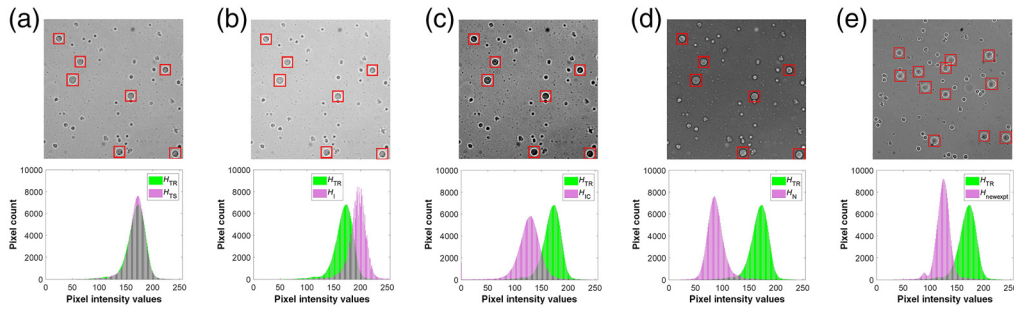
Fig. 9 Impact of image transformations on the aggregate intensity distributions. (a) An original testing image and the corresponding aggregate intensity distribution, H_{TS} (b) Intensity altered image and the corresponding aggregate intensity distribution H_I (c) Intensity and contrast altered image and the corresponding aggregate intensity distribution H_{IC} (d) Negative image and the corresponding aggregate intensity distribution H_N (e) New experiment image and the corresponding intensity distribution H_{newexpt}. Each of the altered intensity distributions are compared with the intensity distribution H_{TR} from the original training set.

2. Increase the image contrast through stretching the image intensity distribution, $h(I)$, around the mean (increase the intensity variance). For this purpose, the offset (B_2) is chosen so that $h(I)$ is centered in the [0, 255] range. This offset is added to every pixel intensity value. Afterward, a random value (A_2) is chosen from a uniform distribution with bounds [1.5, 2.5] and multiplied to every pixel value. This results in a darker version of the image with higher contrast. The altered image and its corresponding histogram are represented with $I_{IC}(x, y) = A_2 \times (I(x, y) + B_2)$ and $h_{\text{IC}}(I)$, respectively. By this transformation, the variance of $h(I)$ has an increase between 125% to 525%. The aggregate of all generated $h_{IC}(I)$ is represented by H_{IC}.

3. Invert the pixel intensity values to generate the negative of the corresponding image. The altered image and its corresponding histogram are represented with $I_N(x, y) = 255 - I(x, y)$ and $h_N(I)$, respectively. The aggregate of all generated $h_N(I)$ is represented by H_N.

In addition to the above three transformations that were performed on the original image sets discussed in the previous sections, we also generated a new experimental image set by altering the focus and exposure time during image acquisition. The aggregate intensity distribution for this new data set is denoted by H_{newexpt}.

In Fig. 9(a), we show the aggregate intensity distribution H_{TR} and H_{TS} for training and testing sets, respectively. As expected, the intensity distributions have similar mean and variance indicating that the detection framework is trained on a similar distribution as the testing set. In contrary, as shown in Figs. 9(b), 9(c), and 9(e), H_I, H_{IC}, and H_{newexpt} have different mean values compared to H_{TR}. Additionally, Fig. 9(c) shows notable differences in the variance of H_{IC} and H_{TR}. Note that Fig. 9(d) displays a mirroring effect between H_{TR} and H_N. This property suggests that, in contrast to the original training images, the cells in the negative images appear brighter than the background.

Table 3 compares the performance of the Faster R-CNN approach on the original testing set, the altered versions of the original testing set, and the new experimental dataset. As shown, the detection model manifests its best performance, i.e., achieves highest $F1$-score, on the original testing set. This observation is rather intuitively comprehensible since the original testing dataset has a similar distribution as the original training set, and so, the detection model is trained for the images in the original testing set. Taking the performance of the model on the testing set as the reference, for 25% to 50% intensity variation and 125% to 525% variance or equivalently contrast change, the F1-score was decreased to <0.1% and 0.2%, respectively. These observations indicate that the Faster R-CNN approach robustly responds to the intensity and contrast variations. Note that the degradation in the performance of the Faster R-CNN approach extends as the transformed intensity distribution further deviates from the intensity distribution of the original training set. Based on the preceding observations, it can be fairly concluded that the detection model extracts features that are relatively robust to intensity and contrast variations.

Table 3 The performance of the Faster-RCNN approach on the original testing and transformed datasets.

Dataset	Histogram	Sensitivity (%)	Precision (%)	F1-score (%)
Original testing images	H_{TS}	99.1	99.8	99.45
Intensity-altered testing images	H_I	98.9	99.9	99.40
Intensity- and contrast-altered testing images	H_{IC}	99	99.5	99.25
Negative testing images	H_N	98.8	71.1	82.70
New experimental image set	$H_{newexpt}$	99.4	97.3	98.34

For the negative testing set, the precision noticeably decreased (-28.7%), causing the F1-score to drop 16.8%, which shows that the model is considerably less robust to transformations that reverse the intensity map of the objects with respect to the background. However, when we retrained the Faster R-CNN model on training images and their negative versions and tested the retrained detection model on the negative of testing images, the performance recovered to 99.1% sensitivity, 97.3% precision, and F1-score of 98.2%. Thus, the detection model tends to show vulnerability toward contrast reversal transformations. However, if trained on both the image and its negative version, the detection model manages to learn features of the cell that are independent of the cell versus background contrast.

3.4 Application of Deep Learning Techniques to Live Patient-Derived CTC Detection

Our results show that the DL-based Faster R-CNN model can successfully classify live *in vitro* breast cancer cells in a background of WBCs. This outcome parallels the success that DL models are achieving in the fields of biomedical image analysis and clinical diagnostics.[46,47] These advances have led to successful efforts in applying DL to detect patient-derived CTCs in fluorescently stained blood samples.[33,35] Applying similar techniques for detecting live CTCs in unstained patient blood samples would open new opportunities (e.g., *ex vivo* expansion of CTCs for drug screening and molecular analysis). However, application of DL techniques for label-free detection of patient-derived CTCs presents challenges as we discuss below.

For supervised machine learning-based models, generation of valid ground truth data is essential. In the case of patient-derived CTCs, generation of ground truth images requires live-cell fluorescent markers that are specific to CTCs but do not stain blood cells. However, tagging CTCs with live-cell fluorescent labels has been a challenge in the field, although some progress is being made. For example, Wang et al. used carbonic anhydrase 1× PE-conjugated antibody along with a live cell dye, calcein acetoxymethyl ester to detect live CTCs in metastatic renal cell carcinoma patients.[35] In another study, a group of near-infrared (IR) heptamethine carbocyanine dyes were used to identify viable CTCs recovered from prostate cancer patients.[48] These dyes were shown to actively tag cancer cells in xenograft models and have since begun to be used as live markers to further improve cancer prognosis and treatment efficacy.[49] Recently, positive selection approaches involving the use of cell surface markers, such as HER2 (human epidermal growth factor receptor-2), EpCAM (epithelial cell adhesion molecule), and EGFR (epidermal growth factor receptor) have also been implemented to recover viable CTCs.[50]

Another challenge for generating ground truth images is the heterogeneity of patient-derived CTCs. During the metastatic cascade, primary tumor cells that are of epithelial origin become migratory acquiring a mesenchymal phenotype.[11,51,52] Studies with cancer patient blood show that some of the isolated CTCs have high epithelial marker expression while others have high mesenchymal marker expression or a combination of both.[53–56] This heterogeneity in marker expression requires multiple fluorescent labels to tag live CTCs and capture the diversity present in patient blood and generate ground truth data.

Finally, machine learning-based approaches often require large data sets for training. Since CTCs are present in extremely low counts in patient blood, generating large, annotated data sets

requires access to many patients' blood samples. Collecting blood samples from a large patient population can be a time-consuming, expensive, and a logistically challenging task.

A potential avenue to address these challenges in label-free live detection of CTCs is to develop robust DL-based approaches to detect blood cells in patient samples and score the remaining non-blood cells as prospective CTCs. This negative selection-based approach will at least help in quickly screening, which patient blood samples are of most interest for further scrutiny.

5 Conclusions

In this work, we proposed an automated framework for label-free detection of MCF-7 breast cancer cells in a background of WBCs in bright-field images. An effective Faster R-CNN-based detection model was developed for detecting MCF-7 cells in the acquired bright-field images. The proposed model demonstrated 99.1% sensitivity and 99.8% specificity, and an average IoU of >80%. The MCF-7 cell detection model analyzed each bright-field image in <0.3 s, more than 300× faster than a human labeler. Also, we introduced a novel fully automated technique for training set label assignment.

Additionally, we conducted multiple studies to investigate the discriminatory features that an effective CNN derives and employs to differentiate MCF-7 cells and WBCs. These studies showed that the size of the cells was the main distinctive feature that the CNN used to distinguish MCF-7 cells from WBCs. However, in the absence of the size feature, the CNN was still capable of learning other features to perform the identification task, with an acceptable, yet decreased, accuracy level.

Finally, we examined the performance of the detection model in the presence of numerous image intensity transformations. The results demonstrated that for intensity and contrast variations, the $F1$-score of the detection model was reduced by $< 0.2\%$. Therefore, the MCF-7 cell detection model was sufficiently robust with respect to intensity and contrast image transformations. These observations inform that the detection model uses intensity- and contrast-invariant features to perform the detection task.

The results in this work indicate that DL approaches could be potentially applied to detect CTCs in the blood of cancer patients. In the future, challenges related to live-cell and multi-marker fluorescent labeling of patient-derived CTCs and community-wide access to large, labeled datasets need to be addressed.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Acknowledgments

References

1. R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA Cancer J. Clin.* **67**(1), 7–30 (2017).
2. R. L. Siegel et al., "Cancer statistics, 2021," *CA Cancer J. Clin.* **71**(1), 7–33 (2021).
3. K. Pantel, R. H. Brakenhoff, and B. Brandt, "Detection, clinical relevance and specific biological properties of disseminating tumour cells," *Nat. Rev. Cancer* **8**(5), 329–340 (2008).
4. I. Crnic and G. Christofori, "Novel technologies and recent advances in metastasis research," *Int. J. Dev. Biol.* **48**(5–6), 573–581 (2004).
5. H. W. Hou et al., "Isolation and retrieval of circulating tumor cells using centrifugal forces," *Sci. Rep.* **3**, 1259 (2013).

6. M. Mego, S. A. Mani, and M. Cristofanilli, "Molecular mechanisms of metastasis in breast cancer—clinical applications," *Nat. Rev. Clin. Oncol.* **7**(12), 693–701 (2010).

7. A. Toss et al., "CTC enumeration and characterization: moving toward personalized medicine," *Ann. Transl. Med.* **2**(11), 108 (2014).

8. R. Harouaka et al., "Circulating tumor cells: advances in isolation and analysis, and challenges for clinical applications," *Pharmacol. Therapeut.* **141**(2), 209–221 (2014).

9. Z. Habli et al., "Circulating tumor cell detection technologies and clinical utility: challenges and opportunities," *Cancers* **12**(7), 1930 (2020).

10. A. M. Sieuwerts et al., "Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells," *J. Natl. Cancer Inst.* **101**(1), 61–66 (2009).

11. J. P. Thiery and C. T. Lim, "Tumor dissemination: an EMT affair," *Cancer Cell* **23**(3), 272–273 (2013).

12. S. L. Stott et al., "Isolation of circulating tumor cells using a microvortex-generating herringbone-chip," *Proc. Natl. Acad. Sci. U. S. A.* **107**(43), 18392–18397 (2010).

13. S. M. Yahyazadeh Mashhadi et al., "Shedding light on the EpCAM: an overview," *J. Cell. Physiol.* **234**(8), 12569–12580 (2019).

14. K.-A. Hyun et al., "Microfluidic flow fractionation device for label-free isolation of circulating tumor cells (CTCs) from breast cancer patients," *Biosens. Bioelectron.* **40**(1), 206–212 (2013).

15. S. J. Tan et al., "Versatile label free biochip for the detection of circulating tumor cells from peripheral blood in cancer patients," *Biosens. Bioelectron.* **26**(4), 1701–1705 (2010).

16. M. E. Warkiani et al., "Slanted spiral microfluidics for the ultra-fast, label-free isolation of circulating tumor cells," *Lab Chip* **14**(1), 128–137 (2014).

17. S. C. Hur et al., "Deformability-based cell classification and enrichment using inertial microfluidics," *Lab Chip* **11**(5), 912–920 (2011).

18. H.-S. Moon et al., "Continuous separation of breast cancer cells from blood samples using multi-orifice flow fractionation (MOFF) and dielectrophoresis (DEP)," *Lab Chip* **11**(6), 1118–1125 (2011).

19. S.-B. Huang et al., "High-purity and label-free isolation of circulating tumor cells (CTCs) in a microfluidic platform by using optically-induced-dielectrophoretic (ODEP) force," *Lab Chip* **13**(7), 1371–1383 (2013).

20. R. Gertler et al., "Detection of circulating tumor cells in blood using an optimized density gradient centrifugation," in *Molecular Staging of Cancer*, H. Allgayer, M. M. Heiss, and F. W. Schildberg, Eds., pp. 149–155, Springer, Berlin, Heidelberg (2003).

21. J. Che et al., "Classification of large circulating tumor cells isolated with ultra-high throughput microfluidic Vortex technology," *Oncotarget* **7**(11), 12748 (2016).

22. H.-S. Moon et al., "Continual collection and re-separation of circulating tumor cells from blood using multi-stage multi-orifice flow fractionation," *Biomicrofluidics* **7**(1), 014105 (2013).

23. E. E. van der Toom et al., "Technical challenges in the isolation and analysis of circulating tumor cells," *Oncotarget* **7**(38), 62754 (2016).

24. A. Kowalik, M. Kowalewska, and S. Góźdź, "Current approaches for avoiding the limitations of circulating tumor cells detection methods—implications for diagnosis and treatment of patients with solid tumors," *Transl. Res.* **185**, 58–84.e15 (2017).

25. D. Zou and D. Cui, "Advances in isolation and detection of circulating tumor cells based on microfluidics," *Cancer Biol. Med.* **15**(4), 335 (2018).

26. M. Pødenphant et al., "Separation of cancer cells from white blood cells by pinched flow fractionation," *Lab Chip* **15**(24), 4598–4606 (2015).

27. E. Ozkumur et al., "DT 45 Miyamoto, E. Brachtel, M. Yu, P.-i. Chen, B. Morgan and J. Trautwein," *Sci. Transl. Med.* **5**, 179ra47 (2013).

28. V. Gupta et al., "ApoStream™, a new dielectrophoretic device for antibody independent isolation and recovery of viable cancer cells from blood," *Biomicrofluidics* **6**(2), 024133 (2012).

29. V. Murlidhar, L. Rivera-Báez, and S. Nagrath, "Affinity versus label-free isolation of circulating tumor cells: who wins?" *Small* **12**(33), 4450–4463 (2016).

30. T. B. Lannin, F. I. Thege, and B. J. Kirby, "Comparison and optimization of machine learning methods for automated classification of circulating tumor cells," *Cytometry A* **89**(10), 922–931 (2016).

31. C. M. Svensson, R. Hubler, and M. T. Figge, "Automated classification of circulating tumor cells and the impact of interobsever variability on classifier training and performance," *J. Immunol. Res.* **2015**, 573165 (2015).

32. C. M. Svensson et al., "Automated detection of circulating tumor cells with naive Bayesian classifiers," *Cytometry A* **85**(6), 501–511 (2014).

33. L. L. Zeune et al., "Deep learning of circulating tumour cells," *Nat. Mach. Intell.* **2**(2), 124–133 (2020).

34. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**(7553), 436–444 (2015).

35. S. Wang et al., "Label-free detection of rare circulating tumor cells by image analysis and machine learning," *Sci. Rep.* **10**, 12226 (2020).

36. C. L. Chen et al., "Deep learning in label-free cell classification," *Sci. Rep.* **6**, 21471 (2016).

37. A. Ciurte et al., "Automatic detection of circulating tumor cells in darkfield microscopic images of unstained blood using boosting techniques," *PLoS One* **13**(12), e0208385 (2018).

38. Y. Zhang et al., "Computational cytometer based on magnetically modulated coherent imaging and deep learning," *Light Sci. Appl.* **8**, 91 (2019).

39. S. Ren et al., "Faster R-CNN: towards real-time object detection with region proposal networks," in *Adv. Neural Inf. Process. Syst.* (2015).

40. D. K. Singh et al., "Label-free fingerprinting of tumor cells in bulk flow using inline digital holographic microscopy," *Biomed. Opt. Express* **8**(2), 536–554 (2017).

41. H. Hasinovic et al., "Janus emulsions from a one-step process; optical microscopy images," *J. Dispersion Sci. Technol.* **35**(5), 613–618 (2014).

42. J. Matas et al., "Robust wide-baseline stereo from maximally stable extremal regions," *Image Vision Comput.* **22**(10), 761–767 (2004).

43. D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," https://arxiv.org/abs/2010.16061 (2020).

44. M. Rizon et al., "Object detection using circular Hough transform," http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.104.8311 (2005).

45. K. He et al., "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.* (2016).

46. T. Ching et al., "Opportunities and obstacles for deep learning in biology and medicine," *J. R. Soc. Interface* **15**(141), 20170387 (2018).

47. G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.* **42**, 60–88 (2017).

48. C. Shao et al., "Detection of live circulating tumor cells by a class of near-infrared heptamethine carbocyanine dyes in patients with localized and metastatic prostate cancer," *PLoS One* **9**(2), e88967 (2014).

49. X. Yang et al., "Near IR heptamethine cyanine dye-mediated cancer imaging," *Clin. Cancer Res.* **16**(10), 2833–2844 (2010).

50. M. Kamal et al., "PIC&RUN: an integrated assay for the detection and retrieval of single viable circulating tumor cells (vol 9, 17470, 2019)," *Sci. Rep.* **10**, 2877 (2020).

51. A. Bonnomet et al., "Epithelial-to-mesenchymal transitions and circulating tumor cells," *J. Mammary Gland Biol. Neoplasia* **15**(2), 261–273 (2010).

52. A. Mitra, L. Mishra, and S. Li, "EMT, CTCs and CSCs in tumor relapse and drug-resistance," *Oncotarget* **6**(13), 10697–10711 (2015).

53. M. Yu et al., "Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition," *Science* **339**(6119), 580–584 (2013).

54. G. Kallergi et al., "Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients," *Breast Cancer Res.* **13**(3), R59 (2011).

55. A. Satelli et al., "Epithelial-mesenchymal transitioned circulating tumor cells capture for detecting tumor progression," *Clin. Cancer Res.* **21**(4), 899–906 (2015).

56. A. J. Armstrong et al., "Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers," *Mol. Cancer Res.* **9**(8), 997–1007 (2011).

Golnaz Moallem is a postdoctoral research fellow in the Personalized Integrative Medicine Laboratory, Department of Radiology, Stanford University. She received her BS degree in electrical engineering from Isfahan University of Technology, and her MS and PhD degrees in electrical engineering from Texas Tech University. Her research interests include machine learning and computer vision applications in clinical data analysis, particularly developing vision systems for medical image analysis, including image classification and object detection.

Adity A. Pore is a graduate research assistant in the Vanapalli Lab at Texas Tech University. She received her BS and MS degrees in chemical engineering from Gujarat Technological University and Texas Tech University, respectively. She is currently pursuing a PhD in chemical engineering. Her current project focuses on isolation and characterization of rare cancer associated cells from the blood of breast cancer patients which could be useful for predicting treatment outcomes.

Anirudh Gangadhar is currently a PhD candidate in chemical engineering from Texas Tech University. Previously, he has received his MS degree in the same major from the University of Florida. His research interests include digital holographic microscopy, label-free screening of cancer cells in blood and machine learning.

Hamed Sari-Sarraf received his PhD in electrical engineering from the University of Tennessee in 1993. He is currently a professor of electrical engineering at Texas Tech University. His area of interest is applied machine vision and learning.

Siva A. Vanapalli received his PhD in chemical engineering from the University of Michigan. He is currently a professor in chemical engineering and the Bryan Pearce Bagley Regents Chair in engineering at Texas Tech University. His research interests are in microfluidics, fluid dynamics, complex fluids, and machine learning, which are applied to biomedical applications in cancer and healthy aging.