Computers in Biology and Medicine 199 (2025) 111333

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Augmentation-based deep learning for identification of circulating tumor cells

Martina Russo [a], Giulia Bertolini [b], Vera Cappelletti [b], Cinzia De Marco [b], Serena Di Cosimo [b], Petra Paiè [c], Nadia Brancati [a,*]

[a] Institute for High Performance Computing and Networking-National Research Council of Italy (ICAR-CNR), Naples, Italy
[b] Fondazione IRCCS-Istituto Nazionale dei Tumori, Milan, Italy
[c] Politecnico di Milano, Milan, Italy

ARTICLE INFO

Keywords:
Circulating tumor cells
Cancer
Metastases
Deep learning
Augmentation
DEParray

ABSTRACT

Circulating Tumor Cells (CTCs) are crucial biomarkers in liquid biopsy, offering a noninvasive tool for cancer patient management. However, their identification remains particularly challenging due to their limited number and heterogeneity. Labeling samples for contrast limits the generalization of fluorescence-based methods across different hospital datasets. Analyzing single-cell images enables detailed assessment of cell morphology, subcellular structures, and phenotypic variations, often hidden in clustered images. Developing a method based on bright-field single-cell analysis could overcome these limitations. CTCs can be isolated using an unbiased workflow combining Parsortix® technology, which selects cells based on size and deformability, with DEPArray™ technology, enabling precise visualization and selection of single cells. Traditionally, DEPArray-acquired digital images are manually analyzed, making the process time-consuming and prone to variability. In this study, we present a Deep Learning-based (DL) classification pipeline designed to distinguish CTCs from leukocytes in blood samples, aimed to enhance diagnostic accuracy and optimize clinical workflows. Our approach employs images from the bright-field channel acquired through DEPArray technology leveraging a ResNet-based Convolutional Neural Network. To improve model generalization, we applied three types of data augmentation techniques and incorporated fluorescence (DAPI) channel images into the training phase, allowing the network to learn additional CTC-specific features. Notably, only bright-field images have been used for testing, ensuring the model's ability to identify CTCs without relying on fluorescence markers. The proposed model achieved an F1-score of 0.798, demonstrating its capability to distinguish CTCs from leukocytes. These findings highlight the potential of DL in refining CTC analysis and advancing liquid biopsy applications.

1. Introduction

Detection of cancer at an early stage significantly improves patient response to treatment and survival. However, nearly 50% of patients are still diagnosed at an advanced stage [1]. Recent advances in biological understanding and technological progress have led to the development of innovative approaches to improve early diagnosis. Single-cells sequencing and spatial transcriptomics shed new lights into the comprehension of tumor development and progression and in the interaction with tumor microenvironment. These data have improved precision therapy for cancer patients. However, the cost and time commitment are substantial, and the procedure requires a tumor tissue biopsy that is invasive and, in some cases, not feasible for certain patients [2]. In this context, the analysis and detection of Circulating Tumor Cells (CTCs) is crucial. CTCs are cells shed from the primary tumor or metastatic sites into the bloodstream that have emerged as a valuable biomarker in oncology via liquid biopsy, providing a real-time and non-invasive means of cancer detection, monitoring, and treatment [3]. Their extreme rarity and heterogeneity pose significant challenges for accurate identification and classification. Usually, techniques based on specific detection of markers expressed by CTC are performed, but they may fail to capture the full spectrum of the heterogeneous CTC population, resulting in an underestimation of CTC counts in other clinical settings [4]. The absence of a universally accepted method prevents the standardization and clinical validation of CTC-based biomarkers [5].

Consequently, there is a need for alternatives that can make this

* Corresponding author.
E-mail address: nadia.brancati@cnr.it (N. Brancati).

https://doi.org/10.1016/j.compbiomed.2025.111333
Received 6 March 2025; Received in revised form 19 May 2025; Accepted 19 November 2025
Available online 21 November 2025

process more efficient and robust, facilitating the high-throughput application of CTC-based technologies and enhancing the study of cancer patients and the progression of their disease. CellSearch is the only FDA-approved (Food and Drug Administration) system for enumerating CTCs in different kinds of cancer clinical settings as breast, colon, and prostate cancer. This technology is based on capture of CTC expressing the epithelial marker EpCAM and therefore loses information on CTC that are not characterized by an epithelial phenotype [6]. After the capture of CTC, a method to analyze images and classify them is required. The ACCEPT Software is an image analysis package for the automated CTC enumeration and phenotyping that is currently under active development for the EU Cancer-ID project [7]. The use of approaches of this kind has demonstrated that CTC size varies significantly across different tumor types and is generally smaller than tumor cell lines commonly used as Refs. [8]. This underscores the importance of morphological analysis in optimizing size-based isolation methods and the need to acquire single-cell images of CTC to finely study their morphological characteristics. The ACCEPT Software is often used in combination with CellSearch for automatic analysis of CTCs. However, it relies on predefined rules and morphological criteria, which can introduce human bias and limit its adaptability to diverse imaging conditions.

An alternative approach is a marker-independent strategies such as Parsortix technology [9] that exploit a microfluid cassette that can enrich CTC based on their physical properties such as size and deformability. A notable innovation in single cell characterization and isolation of CTC is represented by the cell sorting technology based on imaging developed by the Di Trapani group, known as DEPArray™ (Menarini Silicon Biosystems, S.p.A., Italy) [10].

Other paper reported the possible combination of Parosortix with microwell formats (Sievewell and CellCelector Nanowell) for efficient detection, labeling, and isolation of single CTCs from patient's blood samples, enabling downstream DNA and RNA [11]. However, the acquisition with DEPArray has an extremely high precision, even for ultra-rare cells (< 1 cell/ml), and is more suitable for molecular analysis of CTCs.

Even with label-free enrichment technologies such as Parsortix®, CTCs may still be overlooked or misclassified as leukocytes when relying on immunofluorescence based evaluation [12]. This highlights the need for an alternative approach capable of identifying these cells based on their intrinsic morphological characteristics in bright-field (BF) imaging, reducing dependence on fluorescence markers. Despite these technologies' advantages, manual classification of CTCs remains a bottleneck in clinical workflows. This process requires expert analysis of digital images acquired through DEPArray, which is prone to inter-observer variability and may introduce subjectivity into the classification. In this context, Machine (ML) and Deep Learning (DL) models can learn directly from data, enabling more robust and automated CTC detection. They may excel at distinguishing CTCs from other blood cells, even in complex or suboptimal images, and can generalize across different datasets without extensive manual adjustments. In particular, Convolutional Neural Networks (CNNs), have demonstrated strong performance in automating CTC detection [13,14], achieving high recall rates and reinforcing the potential of Artificial Intelligence (AI)-driven methods to enhance early cancer detection and monitoring in medical image analysis [15].

In this study, we propose a DL-based classification pipeline designed to distinguish CTCs from leukocytes in a liquid biopsy using BF images acquired with DEPArray technology. In detail, a pre-trained CNN has been trained on a private dataset and given the small number of acquired CTC images, augmentation data have been used in the training phase. Augmentation operation include both affine and color transformation of BF images and also images coming from the DAPI fluorescence channel, which allows for improving performance of the model. Importantly, only BF images are used during testing, ensuring that the model can identify CTCs without reliance on fluorescence markers.

2. Related works

Several approaches have been proposed in the literature about CTC analysis. Many methods use images containing clusters of cells, and the aim is the detection and quantification of the cells. Guo et al. have proposed a CNN to automatically detect CTCs in peripheral blood using immunofluorescence in situ hybridization (imFISH) images [13]. However, this technique is based on counting the copy number of chromosome 8 using CEP8 immunofluorescence labeling rather than analyzing cell morphology. BRIA (BReast cancer Imaging Algorithm), a fully automated ML-based pipeline designed to detect, segment, and classify metastatic breast CTC cells in multi-channel immunofluorescence images is proposed in [16]. Svensson et al. [17] presented a Naive Bayesian Classifier (NBC) to reliably and automatically detect and quantify CTCs. Cells are collected with a functionalized medical wire, stained for fluorescence microscopy, and the classification is performed by using RGB color histograms. In [18], the authors used a combination of computer vision and CNNs, that demonstrated strong performance in automating CTCs detection from multi-channel clustered fluorescence images, achieving high recall rates. Application on multi-channel images greatly improves the quality of the analyzed information, but the limited test dataset may not guarantee optimal generalization.

Given the scarcity and rarity of CTCs, Liang et al. [19] used a novel data generation with Segment Anything Model (SAM) in combination with copy-paste to increase the number of images, which allows for improved generalization over existing models due to a new loss function. Despite the promising approach, the unreliability of synthetic data and computational cost remain challenges to consider.

A single-cell method is proposed in [20], where an auto-encoder feed by fluorescent images of blood samples has been proved to have an accurate identification of CTCs.

In [21], Miccio et al. demonstrated that label-free approaches eliminate the need for sample staining, reducing preparation time, potential alterations to cell properties, and dependence on specific dyes or markers. In [22] is presented a DL method that uses Whole Slide Images (WSI) for the automatic detection and enumeration of CTCs in microscopic blood images, reducing subjectivity and workload for cytologists. The high resolution of WSIs, however, poses great challenges in their analysis and management [23].

A tentative to merge information from fluorescence and BF images of single cells has been made in [24], where the authors have developed a proof-of-concept DL model that identifies cancer cells originating from cholangiocarcinoma in unlabeled microscopy images based on morphological differences, where each image contains 20–30 individual cells on average, emphasizing the importance of analysis on morphological features of cells that allow different cell types to be distinguished from each other.

Wang et al. [25] used immune microparticles to negatively mark white blood cells rather than CTCs, so that cancer cells can be directly distinguished in the BF channel of microscopy. In this way, all heterogeneous cancer cells and their phenotypic properties can be preserved for further cancer studies. However, applying such methodologies to clusters of cells could lead to the loss of crucial information.

Akashi et al. [26] have demonstrated that, despite training their CNN-based detection system using KYSE520 cell lines with high EpCAM expression, the model has been still able to accurately identify KYSE30 cell lines, which are EpCAM-negative. This suggests that DL does not rely solely on traditional biomarkers, but can detect previously unknown or subtle features, such as cell morphology or minute differences in nuclear structure, to differentiate CTCs from blood cells.

The need to label the sample to obtain contrast agent effects greatly limits the generalization of methods based on fluorescence images, which cannot be applied to datasets from different hospital structures. On the other hand, images containing cell clusters can make it difficult to examine individual cell morphology, subcellular structure, and phenotypic variations; instead, this analysis is much simplified in the case

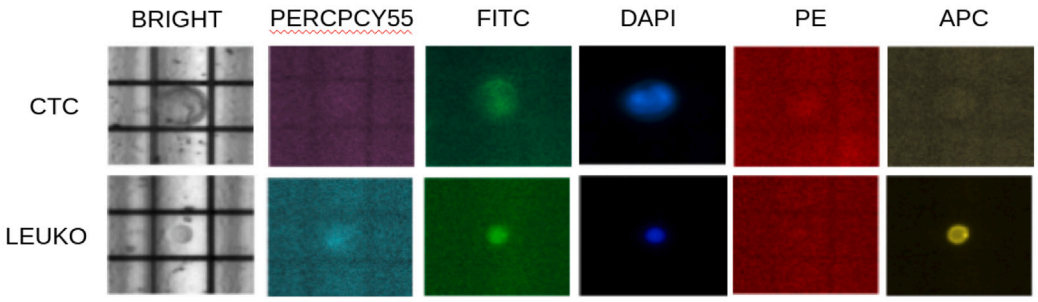


Fig. 1. Different types of images produced by DEParray.

of images containing single cells. Taking into account these limitations of existing methods in the literature, we propose a method based on the analysis of BF images (without the use of fluorescence), acquired with a single-cell technique.

3. Methodology

3.1. Isolation and characterization of CTCs

Isolation and characterization of CTCs were performed according to a previously validated and described protocol [27]. Briefly, Peripheral Blood Mononuclear Cell (PBMC) were separated using SepMate™ PBMC Isolation Tubes from 10 ml blood of two Healthy Voluntaries (HV), thirteen patients with Non-Small Cell Lung Cancer (NSCLC) and two with cholangiocarci- noma.

CTC analysis was performed for patients who provided written informed consent to participate in the CTCNeo study, approved by the Independent Ethics Committee of the Istituto Nazionale dei Tumori of Milan (CTCNeo INT76/21).

For spike-in experiments, about 500 single cells from A549 and Calu3 NSCLC cell lines were spiked into PBMC of HV and processed as detailed below. PBMC from HV and patients were processed with Parsortix system to enrich for CTC. Recovered cells were stained with fluorescent antibodies against specific CTC markers and immune marker CD45, counter-stained with DAPI (4',6-diamidino-2-phenylindole), used for nuclear staining which identifies cells with intact rounded nuclei. Fluorophore conjugated antibodies (detected in FITC/ PE/ PerCP-Cy5.5 channels) were used to analyze the expression of specific antigens potentially expressed by CTC and finally APC channel was used to identified CD45+ leukocytes. Representative images of spike-in cells were captured randomly for each fluorescent channel and BF. Putative CTCs, defined by cell size and negativity for CD45 expression, were isolated and subjected to DNA amplification, LowPass sequencing, and copy number alteration (CNA) analysis to confirm their tumor origin, as detailed in Vismara et al. [27]. Different types of images, stained and in grayscale, are produced by DE- PArray, see Fig. 1. The grayscale image is the BF channel, while the remaining images are the different channels stained with the specific markers. Not all channels were used to create the dataset for our experiments, but only images from BF and DAPI channels. We aim to streamline healthcare professionals' workflow by reducing the dependency on extensive sample staining with multiple markers.

3.2. Dataset characteristics

The dataset comprises 529 images of tumor cell lines spiked into the blood of HV, resembling CTC, 52 CTC images from 13 patients with NSCLC, and 388 leukocyte images. The dataset was partitioned into three distinct subsets for the experiments: training, validation, and test sets. The training set was dedicated to learning the DL model, while the validation set served as an intermediate dataset to fine-tune hyperparameters during the model optimization process. The test set was reserved for evaluating the model's performance on previously unseen data, mimicking real-world scenarios. To enhance generalization during training and address the scarcity of real CTC images, spiked-in cell images were employed for both the training and validation phases, whereas actual CTC images were exclusively used for testing. For leukocytes, 15% of the total images were allocated to the test set. In detail, 529 spiked-in cell images and 332 leukocyte images were utilized for training and validation, with 10% of them randomly selected for validation. The final test phase included 52 images of CTCs and 56 images of leukocytes (Table 1).

Table 1
Number of dataset's images.

	CTC	LEUKO
Train	479	303
Augmented Train	2395	1515
Validation	50	29
Test	52	56
TOTAL	581	388

3.3. Training images

Image augmentation techniques were utilized to expand the dataset and improve the model's ability to generalize. Specifically, three augmentation operations were implemented, incorporating random transformations like rotations and flipping, along with adjustments to brightness and color. The choice to apply these three types of augmentation was motivated by the goal of increasing data diversity. Additionally, images from the DAPI channel were included to further augment data, enhancing generalization and allowing the network to learn additional morphological CTCs' features. This multi-channel approach helps capture intrinsic characteristics of CTCs that may not be evident in BF images alone. Starting from a total of 782 training images in the training set, a final number of 3910 images was obtained after augmentation and the addition of DAPI channel images. In Table 1 the original and augmented training data are reported.

3.4. Proposed method

The decision to rely on the BF channel is due to two main factors: (i) it is a standard imaging output of DEPArray technology, and (ii) the need to simplify the work of researcher operators by evaluating whether it is possible to avoid excessive staining of samples with different markers, which may also vary depending on both the type of tumor to be identified and the hospital structure. However, in BF images, background noise or artifacts may be more pronounced, particularly when the tissue or sample preparation is not ideal. This can interfere with the accurate interpretation of the image. Moreover, they often lack contrast, making it harder to distinguish subtle differences between cells, particularly when they have similar morphological characteristics. This can reduce the clarity and accuracy of cell identification. DAPI

staining is relatively simple and inexpensive compared to other staining techniques. It does not require complex protocols, making it a convenient choice for routine use in various biological and medical imaging applications. The DAPI channel is highly specific and highly effective for visualizing cell nuclei and distinguishing nucleated cells from debris or anucleated cells, regardless of the tumor type. Representative images of CTC and non-CTC cells in both DAPI and BF microscopy are shown in Fig. 2.

Given these considerations, we decided to use BF in combination with DAPI channel images for our experiments. In particular, to exploit the DAPI property, we augment our training dataset by introducing DAPI images. In this way, we have a dual advantage: (i) we increase the number of training images and (ii) introduce new features (those from the DAPI images) that aid the learning process of the networks.

On this dataset, we applied a DL approach, leveraging AI advancements to enhance classification and analysis of complex data. Specifically, we used CNNs, which are a type of neural network that is especially well-suited for image processing. These networks are structured to simulate the visual processing in the human brain, using convolutional layers to extract features from an image through specific filters. In other words, CNNs are capable of automatically learning patterns and relevant features from visual datasets without the need for manual feature extraction. In our study, we used several advanced CNNs architectures, including ResNet, [28], EfficientNet [29], and DenseNet [30]. These models were chosen due to their strong performance in handling complex image classification tasks, attributed to their efficiency in feature learning and ability to generalize across diverse datasets. ResNet, for instance, features a deep architecture with residual connections that help address the vanishing gradient problem. EfficientNet is specifically designed to optimize the trade-off between network depth, width, and input resolution, and finally, DenseNet employs densely connected layers to enhance information flow throughout the network.

Each architecture offers multiple variants tailored to different computational and performance needs. ResNet, for example, ranges from the lightweight ResNet18 with 18 layers to the more advanced ResNet152 with 152 layers. This diversity allows for flexibility in choosing models suited for various image recognition tasks.

A comparative analysis of these CNNs architectures was conducted, selecting specific versions that balance accuracy and computational efficiency, particularly for medical applications. Specifically, ResNet34 and ResNet50 were chosen from the ResNet family, while EfficientNetB4 and DenseNet121 were selected from the EfficientNet and DenseNet families, respectively. All networks were trained and tested on the dataset outlined in Section 3.2, incorporating the data augmentation strategies detailed in Section 3.3.

The workflow of the proposed method is shown in Fig. 3.

4. Results and discussion

4.1. Comparative results

Several performance metrics were employed to evaluate the effectiveness of classification models, including Accuracy, F1-score, Precision, and Recall. These metrics rely on four key components: true positives (TP), which represent correctly identified positive cases; false positives (FP), where negative instances are mistakenly classified as positive; true negatives (TN), referring to correctly identified negative cases; and false negatives (FN), where positive instances are incorrectly classified as negative. Accuracy measures the proportion of correct predictions out of the total, offering a broad assessment of model performance. However, in cases of class imbalance, accuracy alone can be misleading. The F1-score, calculated as the harmonic mean of Precision and Recall, provides a balanced evaluation by considering both false positives and false negatives. This makes it particularly useful when dealing with imbalanced datasets where neither Precision nor

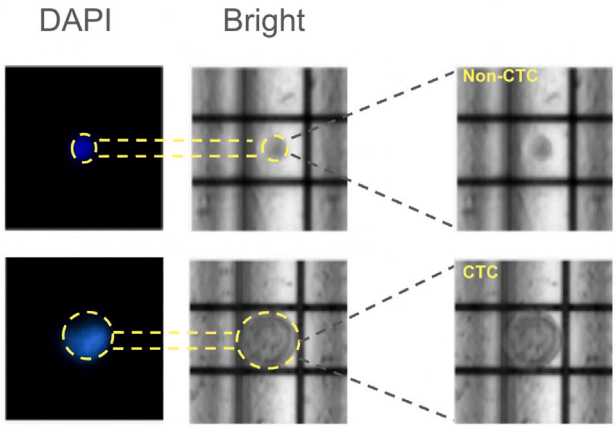


Fig. 2. Representative images of CTC and non-CTC cells in both DAPI and BF microscopy.

Recall should outweigh the other. Precision focuses on how many of the predicted positive cases are actually correct, which is critical in situations where false positives carry significant consequences, such as in medical screenings or fraud detection. Conversely, Recall measures how well the model identifies actual positive cases, making it especially important in scenarios where missing a positive instance, such as in disease diagnosis, could have serious repercussions.

We evaluated these metrics on several CNNs, including DenseNet121, EfficientNetB4, ResNet50, and ResNet34, all pre-trained on the ImageNet dataset. Considering the imbalance of the training dataset, the F1-score was chosen as the metric to evaluate the proposed method. So, for each architecture, the best model with the highest F1-score was chosen during the validation phase, and then this model was evaluated on a test set. The images were all resized to 148 × 148, which corresponds to the size of the smallest image in the dataset. Cross entropy loss was considered for the backpropagation and an AdamW optimizer [32] was adopted with an initial learning rate equal to 10^{-7}. For each experiment, the mean and standard deviation of the selected evaluation metrics on the test set using five random weight initializations were computed. The average performance metrics are summarized in Table 2. ResNet50 achieved the highest F1 score, 0.798, indicating that this model significantly outperformed the others in terms of the balance between precision and recall. A precision value of 0.828 indicates that about 17.2% of the leukocytes were misclassified as CTCs, resulting in false positives. On the other hand, the recall value of 0.804 indicates that nearly 20% of true CTCs were not identified, resulting in false negatives. This limitation can be addressed in future research, as false negatives in a diagnostic context could have significant clinical consequences. However, ResNet50 was particularly effective at learning the relevant features for the task as a whole. With an F1 score of 0.578, EfficientNetB4 performed better than DenseNet121 and ResNet34, but still lagged behind ResNet50. EfficientNet models are designed to be more parameter-efficient, but in this case, it seems that EfficientNetB4 did not generalize as well as ResNet50. Despite belonging to the same ResNet family, ResNet34 achieved an F1 score of 0.407, which is much lower than ResNet50. This suggests that the deeper architecture of ResNet50 allowed it to learn more complex features, leading to significantly better performance. The lowest F1 score (0.378) was observed with DenseNet121. While DenseNet architectures are known for efficient feature reuse, this result indicates that DenseNet121 struggled with the dataset, possibly due to overfitting, or difficulties in extracting discriminative features.

Unfortunately, a direct comparison with existing studies in the literature is not feasible, as most approaches focus on different types

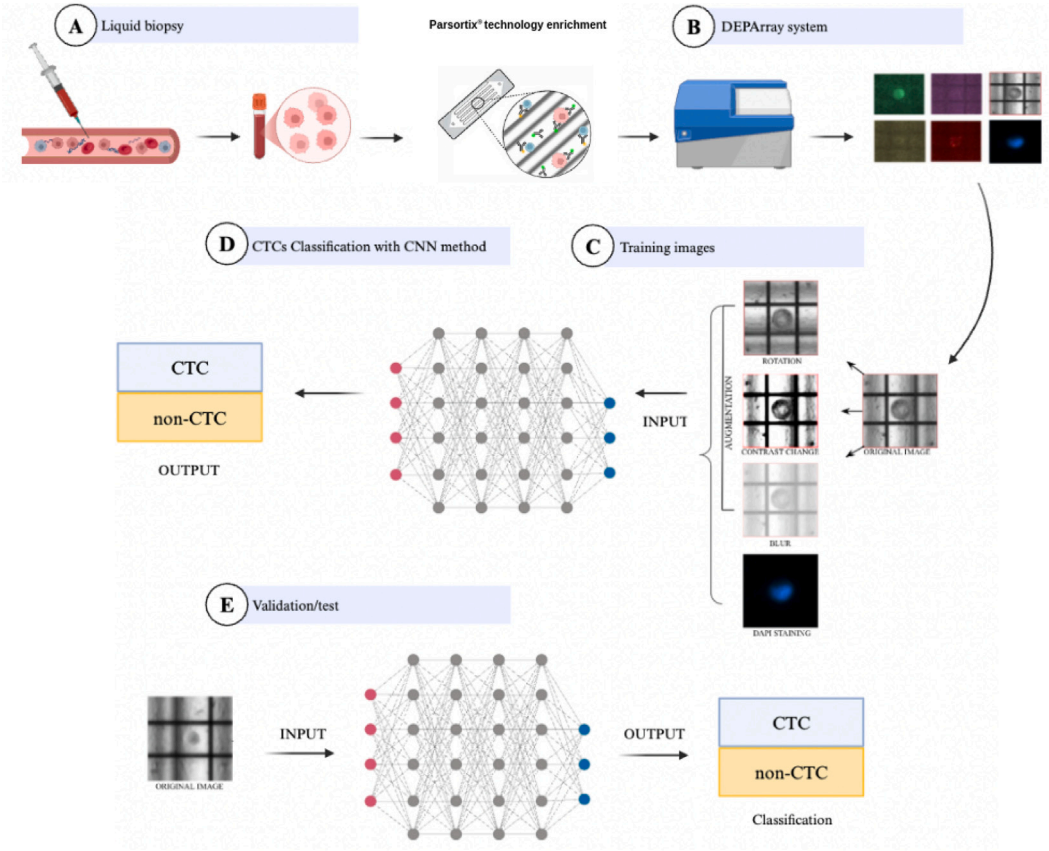


Fig. 3. Workflow of the proposed method. A. Liquid biopsy is a blood test that detects cancer cells or tumor DNA, avoiding invasive procedures. **B.** The combined Partortix™ [31] and DEPArray™ systems use electric fields to isolate and select single cells, like CTCs. **C.** Training image pre-processing consist of BF DEPArray images whose variability is increased through the use of augmentation operations, and fluorescent-field images of DAPI-labeled cells. **D.** The output of the CNN is the classification of images into CTC and non-CTC. **E.** BF images are used in the validation/test phase to identify CTCs.

Table 2
Results of the different CNN architectures. In bold the best results for each measure.

	Accuracy	Precision	Recall	F1-score
EfficientnetB4	0.579 ± 0.027	0.580 ± 0.024	0.580 ± 0.024	0.578 ± 0.027
Densenet121	0.535 ± 0.005	0.760 ± 0.000	0.518 ± 0.005	0.378 ± 0.011
Resnet34	0.549 ± 0.014	0.765 ± 0.005	0.533 ± 0.015	0.407 ± 0.029
Resnet50	**0.798 ± 0.003**	**0.828 ± 0.011**	**0.804 ± 0.005**	**0.798 ± 0.005**

of images, such as fluorescence-based imaging or cell clusters. Consequently, the architectures proposed in those studies are not directly applicable to our framework, which relies solely on BF images and single-cell analysis. On the other hand, it is not possible to validate our architecture on their datasets, as the corresponding images are not publicly available online.

4.2. Ablation study

To assess the importance of the introduction of DAPI images in the augmentation process, an ablation study was performed. In Table 3, the results of the following experiments are reported:

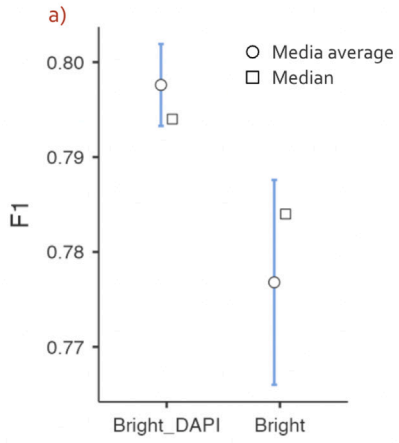
- AUG1: only one augmentation operation;
- AUG2: two augmentation operations;
- BF w/o DAPI: three augmentation operations, but without DAPI images;
- BF w/ DAPI no AUG: augmentation operation only with DAPI images;
- BF w/ DAPI: our proposed approach, with three augmentation operations and a further augmentation with DAPI images.

Observing the results, it is possible to note that a key element in the enhanced performance of the ResNet50 model was the implementation of carefully designed data augmentation strategies. These techniques substantially increased the diversity of the training set, addressing the challenges associated with the limited initial dataset. By introducing a broader range of variations, the model was better equipped to learn robust and generalizable features, ultimately improving its classification accuracy and reducing the risk of overfitting. In particular, the results improved as the number of augmentation operations introduced increased (AUG1, AUG2). Indeed, the application of three specific augmentation functions improved the robustness of the model (BF w/o DAPI). Furthermore, to understand the importance of augmentation operations during training, we attempted to train the model without them, inserting only DAPI images to increase the variability of the data (BF w/ DAPI no AUG). These findings indicate that a substantial number and variability of data are essential for optimal model performance. Finally, integrating an auxiliary dataset containing DAPI fluorescence channel images in the training set further enhanced performance across all evaluated metrics (BF w/ DAPI).

To better confirm the importance of BF image analysis and that DAPI images were used only for additional support, we also conducted experiments by reversing the roles of the two types of images:

Table 3
Results of ablation study.

	Accuracy	Precision	Recall	F1-Score
AUG1	0.722 ± 0.009	0.793 ± 0.015	0.730 ± 0.010	0.709 ± 0.010
AUG2	0.759 ± 0.009	0.793 ± 0.012	0.767 ± 0.006	0.757 ± 0.010
BF w/o DAPI	0.780 ± 0.012	0.806 ± 0.017	0.782 ± 0.013	0.777 ± 0.012
BF w/ DAPI no AUG	0.663 ± 0.017	0.690 ± 0.026	0.660 ± 0.010	0.641 ± 0.011
BF w/ DAPI	**0.798 ± 0.003**	**0.828 ± 0.011**	**0.804 ± 0.005**	**0.798 ± 0.005**
DAPI w/o BF	0.537 ± 0.009	0.560 ± 0.010	0.550 ± 0.010	0.510 ± 0.010
DAPI w/ BF	0.497 ± 0.019	0.507 ± 0.021	0.507 ± 0.021	0.474 ± 0.027



b) Normality test (Shapiro-Wilk)

	W	p
F1	0.811	0.020

c) Independent samples t-test

		Statistics	p
F1	U of Mann-Whitney	0.00	0.011

d) Homogeneity of variances test (Levene)

	F	gdl	gdl2	p
F1	2.20	1	8	0.176

Fig. 4. Statistical results of Mann–Whitney test (a) Statistical results with $p - value < 0.05$ demonstrate a statistically significant difference between the two groups considered (BF and BF with DAPI). (b) The data do not follow a normal distribution. (c) Mann–Whitney test. (d) Homogeneity test.

- DAPI w/o BF: DAPI images were used as main images for the analysis, and three augmentation operations were performed on them for the training phase. In the validation and test phase, only DAPI images were used for the final classification;
- DAPI w/ BF: for the training phase, DAPI images, three augmentation operations on them, and BF images were used. As before, only DAPI images were used for the final classification in the validation and test phase.

The results in this case were rather discouraging in both cases: for DAPI w/o BF, the F1-score decreased from 0.798 to 0.510, while for DAPI w/ BF the F1-score even decreased to 0.474. These results show that the utilization of images derived from the BF during the testing phase is of paramount importance. These images enable the model to analyze the morphological shape of the cell, facilitating its classification into the appropriate category. On the other hand, the utilization of images in the fluorescence field of DAPI was demonstrated to be pivotal during the model's training phase, as it enables the network to extract the crucial information necessary for the subsequent validation step.

4.3. Statistical analysis

A statistical analysis of the data has been performed. We used Jamovi software to conduct Student's t-test. The data were analyzed by checking the homogeneity of variance using Levene's test ($p - value = 0,176$), and the normality of the distributions using the Shapiro–Wilk test, which showed that the data did not follow a normal distribution since the p-value is lower than 0.05 ($p - value = 0,020$). So, considering this not normal distribution and the sample size was relatively small, the Mann–Whitney U test, a non-parametric test that does not assume normality, was employed. In particular, to evaluate the impact of including images in the DAPI channel on the model's performance during the training set, the Mann–Whitney U test was used to compare the F1 score obtained under two conditions: (1) the use of the dataset containing BF images with three augmentation operations, and (2) the use of the combined dataset containing both BF with three augmentation operations and DAPI channel images (BF w/o DAPI vs BF w/ DAPI). The F1-score was chosen as the evaluation metric since it balances precision and recall, providing an overall measure of the model's performance. The Mann–Whitney U test was performed to determine if the difference in F1-score between the two groups was statistically significant, with a significance level set at 0.05 ($p - value = 0,011$) (see Fig. 4). This result indicates that the inclusion of the dataset containing DAPI channel images as an augmentation operation is crucial for increasing data variability, ultimately enhancing the model's generalization capabilities. The observed statistical significance strongly supports the rejection of the null hypothesis, confirming that this improvement is not due to random variability but rather to the contribution of fluorescence-derived features. Moreover, training solely on BF images, even with augmentation, appears insufficient to achieve comparable performance, highlighting the necessity of leveraging fluorescence-enhanced training strategies. These findings reinforce the practical importance of incorporating additional feature-rich data sources during training, ensuring a more robust and reliable classification pipeline for CTC identification.

5. Conclusion

CTCs play a crucial role as biomarkers in liquid biopsy, providing a minimally invasive tool to monitor cancer progression and guide therapeutic strategies. Despite their clinical potential, the identification of CTCs remains challenging due to their extreme rarity and heterogeneity. To address these issues, an unbiased workflow combining Parsortix® technology, which isolates CTCs based on size and deformability, with DEPArray™ technology, which enables the precise visualization and selection of single cells, has been developed. Traditionally, the analysis of cell images obtained through DEPArray is performed manually by experienced researchers, a process that is both time-intensive and prone to variability. Automating the identification of CTCs within clinical workflows could streamline the detection of metastases, enhance the accuracy of therapeutic decisions, and ultimately

improve patient outcomes. In this study, we have presented a DL-based classification system designed to differentiate CTCs from leukocytes in liquid biopsy samples. The system leverages the ResNet architecture, a CNN widely recognized for its robustness in medical image analysis. Among the tested architectures, applied to images acquired with DEPArray technology, ResNet50 demonstrated superior performance, achieving an F1-score of 0.798, which was statistically significant (p-value < 0.05) compared to ResNet34 (0.407), EfficientNetB4 (0.578), and DenseNet121 (0.378). This result underscores the suitability of ResNet50 for addressing the inherent complexity of BF images of CTCs.

A pivotal element in achieving these promising results was the implementation of tailored data augmentation techniques, which increased the variability of the training data and compensated for the limited size of the initial dataset. Specifically, three augmentation functions were employed to improve the robustness of the model, and further augmentation was given by the inclusion of an additional dataset images of the DAPI fluorescence channel led to further improvements in all performance metrics. These findings highlight the potential of combining data augmentation strategies with multimodal imaging to overcome dataset limitations, advancing the application of DL in CTCs identification.

CRediT authorship contribution statement

Martina Russo: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation. **Giulia Bertolini:** Writing – review & editing, Writing – original draft, Validation, Formal analysis, Data curation. **Vera Cappelletti:** Writing – review & editing, Writing – original draft, Validation. **Cinzia De Marco:** Writing – review & editing, Writing – original draft, Validation, Formal analysis, Data curation. **Serena Di Cosimo:** Writing – review & editing, Writing – original draft, Validation, Formal analysis, Data curation. **Petra Paiè:** Writing – review & editing, Writing – original draft, Validation, Formal analysis, Conceptualization. **Nadia Brancati:** Methodology, Formal analysis, Conceptualization, Writing – review & editing, Writing – original draft, Validation, Supervision, Software.

Human and animal rights

The authors declare that the work described has not involved experimentation on humans or animals.

Funding

This work has been supported by: European Union – Next Generation EU, Missione 4, Componente 1, “Fondo PRIN 2022” – HIMALAYA – ID 20228MHWPZ - CUP B53D23002380006.

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

This work was funded by European Union – Next Generation EU, Missione 4, Componente 1, “Fondo PRIN 2022” – HIMALAYA – ID 20228MHWPZ - CUP B53D23002380006, whose support was instrumental in carrying out this study.

References

[1] D. Crosby, S. Bhatia, K.M. Brindle, L.M. Coussens, C. Dive, M. Emberton, S. Esener, R.C. Fitzgerald, S.S. Gambhir, P. Kuhn, et al., Early detection of cancer, Science 375 (6586) (2022) eaay9040.

[2] Y. Han, D. Wang, L. Peng, T. Huang, X. He, J. Wang, C. Ou, Single-cell sequencing: a promising approach for uncovering the mechanisms of tumor metastasis, J. Hematol. Oncol. 15 (1) (2022) 59.

[3] C. Alix-Panabières, K. Pantel, Liquid biopsy: from discovery to clinical application, Cancer Discov. 11 (4) (2021) 858–873.

[4] S. Wang, Y. Zhou, X. Qin, S. Nair, X. Huang, Y. Liu, Label-free detection of rare circulating tumor cells by image analysis and machine learning, Sci. Rep. 10 (1) (2020) 12226.

[5] S. Ju, C. Chen, J. Zhang, L. Xu, X. Zhang, Z. Li, Y. Chen, J. Zhou, F. Ji, L. Wang, Detection of circulating tumor cells: opportunities and challenges, Biomark. Res. 10 (1) (2022) 58.

[6] S. Riethdorf, L. O'Flaherty, C. Hille, K. Pantel, Clinical applications of the cellsearch platform in cancer patients, Adv. Drug Deliv. Rev. 125 (2018) 102–121.

[7] L. Zeune, G. van Dalum, C. Decraene, C. Proudhon, T. Fehm, H. Neubauer, B. Rack, M. Alunni-Fabbroni, L.W. Terstappen, S.A. van Gils, et al., Quantifying her-2 expression on circulating tumor cells by accept, PloS One 12 (10) (2017) e0186562.

[8] P.A. Mendelaar, J. Kraan, M. Van, L.L. Zeune, L.W. Terstappen, E. Oomen-de Hoop, J.W. Martens, S. Sleijfer, Defining the dimensions of circulating tumor cells in a large series of breast, prostate, colon, and bladder cancer patients, Mol. Oncol. 15 (1) (2021) 116–125.

[9] G. Wishart, A. Templeman, F. Hendry, K. Miller, A.-S. Pailhes-Jimenez, Molecular profiling of circulating tumour cells and circulating tumour dna: Complementary insights from a single blood sample utilising the Parsortix® system, Curr. Issues Mol. Biol. 46 (1) (2024) 773–787.

[10] M. Di Trapani, N. Manaresi, G. Medoro, Deparray™ system: an automatic image-based sorter for isolation of pure circulating tumor cells, Cytom. Part A 93 (12) (2018) 1260–1266.

[11] Liwen Yang, Mahdi Rivandi, Andr/'e' Franken, Maarten Hieltjes, Pieter Jan van der Zaag, Constantin Nelep, Jens Eberhardt, Stefan Peter, Dieter Niederacher, Tanja Fehm, et al., Implementing microwell slides for detection and isolation of single circulating tumor cells from complex cell suspensions, Cytom. Part A 101 (12) (2022) 1057–1067.

[12] M. Bates, B.M. Mohamed, M.P. Ward, T.E. Kelly, R. O'Connor, V. Malone, R. Brooks, D. Brooks, S. Selemidis, C. Martin, et al., Circulating tumour cells: The good, the bad and the ugly, Biochim. Biophys. Acta (BBA)-Reviews Cancer 1878 (2) (2023) 188863.

[13] Z. Guo, X. Lin, Y. Hui, J. Wang, Q. Zhang, F. Kong, Circulating tumor cell identification based on deep learning, Front. Oncol. 12 (2022) 843879.

[14] M. Vidlarova, A. Rehulkova, P. Stejskal, A. Prokopova, H. Slavik, M. Hajduch, J. Srovnal, Recent advances in methods for circulating tumor cell detection, Int. J. Mol. Sci. 24 (4) (2023) 3902.

[15] D. Painuli, S. Bhardwaj, et al., Recent advancement in cancer diagnosis using machine learning and deep learning techniques: A comprehensive review, Comput. Biol. Med. 146 (2022) 105580.

[16] E. Schwab, B. Annaldas, N. Ramesh, A. Lundberg, V. Shelke, X. Xu, C. Gilbertson, J. Byun, E.T. Lam, Fully automated ctc detection, segmentation and classification for multi-channel IF imaging, in: International Workshop on Medical Optical Imaging and Virtual Microscopy Image Analysis, Springer, 2024, pp. 55–65.

[17] C.-M. Svensson, S. Krusekopf, J. Lücke, M. Thilo Figge, Automated detection of circulating tumor cells with naive bayesian classifiers, Cytom. Part A 85 (6) (2014) 501–511.

[18] S. Calvo-Almeida, I. Serrano-Llabrés, V.M. Cal-González, P. Piairo, L.R. Pires, L. Diéguez, L. González-Castro, Multichannel fluorescence microscopy images ctc detection: A deep learning approach, in: AIP Conference Proceedings, vol. 3030, AIP Publishing, 2024.

[19] S. Liang, X. Bai, Y. Gu, Improving circulating tumor cell detection using image synthesis and transformer models in cancer diagnostics, Sensors 24 (23) (2024) 7822.

[20] L.L. Zeune, Y.E. Boink, G. van Dalum, A. Nanou, S. de Wit, K.C. Andree, J.F. Swennenhuis, S.A. van Gils, L.W. Terstappen, C. Brune, Deep learning of circulating tumour cells, Nat. Mach. Intell. 2 (2) (2020) 124–133.

[21] L. Miccio, F. Cimmino, I. Kurelac, M.M. Villone, V. Bianco, P. Memmolo, F. Merola, M. Mugnano, M. Capasso, A. Iolascon, et al., Perspectives on liquid biopsy for label-free detection of circulating tumor cells through intelligent lab-on-chips, View 1 (3) (2020) 20200034.

[22] Y. Liu, H. Si, Y. Chen, Y. Chen, Faster R-CNN based robust circulating tumor cells detection with improved sensitivity, Proc. 2nd Int. Conf. Big Data Technol. 2019, 252–256.

[23] N. Brancati, G. De Pietro, D. Riccio, M. Frucci, Gigapixel histopathological image analysis using attention-based neural networks, IEEE Access 9 (2021) 87552–87562.

[24] C. Piansaddhayanon, C. Koracharkornradt, N. Laosaengpha, Q. Tao, P. Ingrungruanglert, N. Israsena, E. Chuangsuwanich, S. Sriswasdi, Label-free tumor cells classification using deep learning and high-content imaging, Sci. Data 10 (1) (2023) 570.

[25] S. Wang, S. Hong, S. Cai, J. Lei, J. Chen, N. Zhang, Z. Ai, K. Liu, M. Tang, Negative depletion mediated brightfield circulating tumour cell identification strategy on microparticle-based microfluidic chip, J. Nanobiotechnol. 18 (2020) 1–9.

[26] T. Akashi, T. Okumura, K. Terabayashi, Y. Yoshino, H. Tanaka, T. Yamazaki, Y. Numata, T. Fukuda, T. Manabe, H. Baba, et al., The use of an artificial intelligence algorithm for circulating tumor cell detection in patients with esophageal cancer, Oncol. Lett. 26 (1) (2023) 1–9.

[27] M. Vismara, C. Reduzzi, M. Silvestri, F. Murianni, G. Lo Russo, O. Fortunato, R. Motta, D. Lanzoni, F. Giovinazzo, P. Miodini, et al., Single-cell phenotypic and molecular characterization of circulating tumor cells isolated from cryopreserved peripheral blood mononuclear cells of patients with lung cancer and sarcoma, Clin. Chem. 68 (5) (2022) 691–701.

[28] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[29] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in: International Conference on Machine Learning, PMLR, 2019, pp. 6105–6114.

[30] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

[31] C. Løppke, AM. Jørgensen, NT. Sand, RB. Klitgaard, G. Daugaard, Agerbæ k. MØ, Combined microfluidic enrichment and staining workflow for single-cell analysis of circulating tumor cells in metastatic prostate cancer patients, Sci Rep. 14 (1) (2024) 17501.

[32] I. Loshchilov, F. Hutter, Decoupled weight decay regularization, 2017, arXiv preprint arXiv:1711.05101.