





Article

Deep Active Learning for Automatic Mitotic Cell Detection on HEp-2 Specimen Medical Images

Asaad Anaam ^{1,†} , Mugahed A. Al-antari ^{2,†} , Jamil Hussain ³ , Nagwan Abdel Samee ⁴ ,
Maali Alabdulhafith ^{4,*} and Akio Gofuku ^{1,*}

- ¹ Graduate School of Interdisciplinary Science and Engineering in Health Systems, Okayama University, Okayama 700-8530, Japan
- ² Department of Artificial Intelligence, College of Software & Convergence Technology, Daeyang AI Center, Sejong University, Seoul 05006, Republic of Korea
- ³ Department of Data Science, College of Software & Convergence Technology, Daeyang AI Center, Sejong University, Seoul 05006, Republic of Korea
- ⁴ Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
- * Correspondence: mialabdulhafith@pnu.edu.sa (M.A.); gofuku-a@okayama-u.ac.jp (A.G.)
- † These authors contributed equally to this work.

Abstract: Identifying Human Epithelial Type 2 (HEp-2) mitotic cells is a crucial procedure in anti-nuclear antibodies (ANAs) testing, which is the standard protocol for detecting connective tissue diseases (CTD). Due to the low throughput and labor-subjectivity of the ANAs' manual screening test, there is a need to develop a reliable HEp-2 computer-aided diagnosis (CAD) system. The automatic detection of mitotic cells from the microscopic HEp-2 specimen images is an essential step to support the diagnosis process and enhance the throughput of this test. This work proposes a deep active learning (DAL) approach to overcoming the cell labeling challenge. Moreover, deep learning detectors are tailored to automatically identify the mitotic cells directly in the entire microscopic HEp-2 specimen images, avoiding the segmentation step. The proposed framework is validated using the I3A Task-2 dataset over 5-fold cross-validation trials. Using the YOLO predictor, promising mitotic cell prediction results are achieved with an average of 90.011% recall, 88.307% precision, and 81.531% mAP. Whereas, average scores of 86.986% recall, 85.282% precision, and 78.506% mAP are obtained using the Faster R-CNN predictor. Employing the DAL method over four labeling rounds effectively enhances the accuracy of the data annotation, and hence, improves the prediction performance. The proposed framework could be practically applicable to support medical personnel in making rapid and accurate decisions about the mitotic cells' existence.

Keywords: medical HEp-2 specimen images; HEp-2 mitotic cell detection; deep active learning (DAL); automatic data annotation; computer-aided detection (CAD)



Citation: Anaam, A.; Al-antari, M.; Hussain, J.; Abdel Samee, N.; Alabdulhafith, M.; Gofuku, A. Deep Active Learning for Automatic Mitotic Cell Detection on HEp-2 Specimen Medical Images. *Diagnostics* **2023**, *13*, 1416. <https://doi.org/10.3390/diagnostics13081416>

Academic Editor: Sameer Antani

Received: 11 February 2023

Revised: 9 April 2023

Accepted: 12 April 2023

Published: 14 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Connective tissue diseases (CTD) are autoimmune disorders initiated by generating abnormal autoantibodies that act against nuclear and cytoplasmic antigens and so-called anti-nuclear antibodies (ANAs). ANAs affect body tissues, causing chronic inflammatory disorders such as systemic lupus erythematosus, rheumatoid arthritis, systemic sclerosis, Sjögren's syndrome, and idiopathic inflammatory myopathies [1]. Diagnosing such diseases involves identifying and measuring the corresponding ANAs in the patients' serum [1]. In this regard, the indirect immunofluorescence using human epithelial type-2 cells (IIF HEp-2) substrates is considered the standard serological protocol for ANAs testing [2].

In the IIF HEp-2 protocol, several ANAs appear as distinct HEp-2 cell fluorescence patterns, and each one is associated with a particular disorder [3,4]. The IIF HEp-2 protocol involves several visual analysis steps at the fluorescence microscope, including determining

the positivity strength of the specimen (slide), recognizing the HEp-2 cells in the mitotic phase, and then classifying the staining patterns of interphase cells to identify the type of ANAs [4]. As it requires several steps for slide preparation and visual pattern interpretation, the IIF HEp-2 manual protocol is a subjective and semi-quantitative method that heavily depends on the expertise of the practitioner, which increases the diagnosing variability across observers or laboratories. In addition, the preparation time requirement reduces the throughput of this test. Accordingly, the limitations of the manual IIF HEp-2 protocol have motivated recent interest in developing reliable computer-aided diagnosis (CAD) systems to support the diagnostic procedure [4,5]. However, a small portion of this interest is directed to approaching the task of mitotic cell identification even though it is a crucial step in the IIF HEp-2 diagnostic procedure, and hence, for developing a feasible HEp-2 CAD system [4]. Recently, the CAD system based on deep learning has gotten much attention for automating the prediction process with a massive number of extracted deep features, improving detection performance, and reducing user intervention during the training process [6,7].

HEp-2 cell patterns could be classified based on the cell cycle phases into two broad categories, i.e., mitotic patterns (cell under division phase) and interphase patterns (non-mitotic phase). The HEp-2 cells undergo the mitotic phase through 10~15% of their entire lifecycle, causing their appearance in the HEp-2 specimen images to be very rare compared to the interphase cell patterns [4]. During the mitotic phase, HEp-2 cells exhibit distinct patterns as the cells' DNA has different concentrations and distributions compared to the interphase counterpart [8–11]. Figure 1 shows examples of mitotic cells occurring in different HEp-2 specimen interphase patterns and their corresponding binary masks. Mitotic cell identification is clinically important for validating the correctness of the specimen's preparation procedures and providing supportive information for discriminating between some ambiguous interphase patterns in the mixed patterns specimen images. This is because some mitotic patterns match certain interphase types [4].

HEp-2 mitotic cell detection using HEp-2 specimen images has barely been studied in the literature of HEp-2 image analysis [4]. Typically, the vast majority of these works addressed this task as a binary classification problem between the minority mitotic class against the remaining majority interphase class in the cell-level images [9–11]. However, addressing this task in the cell-level images imposes many challenges for building a practical HEp-2 CAD system. The most critical challenge is the requirement of cell segmentation step overhead. The HEp-2 individual cell segmentation is currently conducted using masks of the specimen images obtained through the DAPI binary mask [12]. DAPI is a fluorescent dye that delineates cells' DNA regions to obtain ground truth masks. However, this added extra cost, effort, and time to the data acquisition process [13] in addition to the concerns about the carcinogenic nature of DAPI materials [14]. Furthermore, the severe skewness of the acquired cell-level HEp-2 mitotic-interphase data towards the interphase majority class further complicates the classification task, requiring additional data balancing approaches.

Motivated by the outstanding success of the recent deep learning object detectors in various imaging applications, and contrary to the previous works, this work proposes a deep learning-based multiple mitotic cell detection framework that could be implemented directly on the microscopic HEp-2 specimen images without segmentation and DAPI channel step overhead. The proposed mitotic cell detection framework is trained based on a deep active learning strategy to refine the detection ground truth (GT) labeling iteratively during training. The main contributions of this work are further highlighted as follows:

- An end-to-end deep learning framework is proposed to accurately and rapidly detect multiple mitotic cells from the entire HEp-2 specimen images within seven different classes of the I3A dataset: Centromere, Golgi, Homogeneous, Nuclear Membrane, Speckled, Nucleolar, and Mitotic Spindle. This is the first work that proposes detecting mitotic cells from the microscopic specimen image regardless of its class type, which provides a practical framework that could be integrated with a microscopic imaging

system to support instant and direct detection of the mitotic cell from the captured specimen images.

- A deep active learning (DAL) strategy is involved to automatically annotate the mitotic cells for the detection task by adjusting the bounding boxes to surround each mitotic cell in the specimen images. Providing such annotation data could initiate further studies for developing mitotic cell detection approaches that could be applied to whole HEP-2 specimen images.
- A pre-processing step via contrast-limited adaptive histogram equalization (CLAHE) is employed to enhance the specimen image quality leading to improve detection performance.
- A comprehensive detection study is conducted by adopting two well-performed deep learning detectors (i.e., YOLO [15] and Faster R-CNN [16]) for detecting the HEP-2 mitotic cells to compare and select the best solution for this detection task.

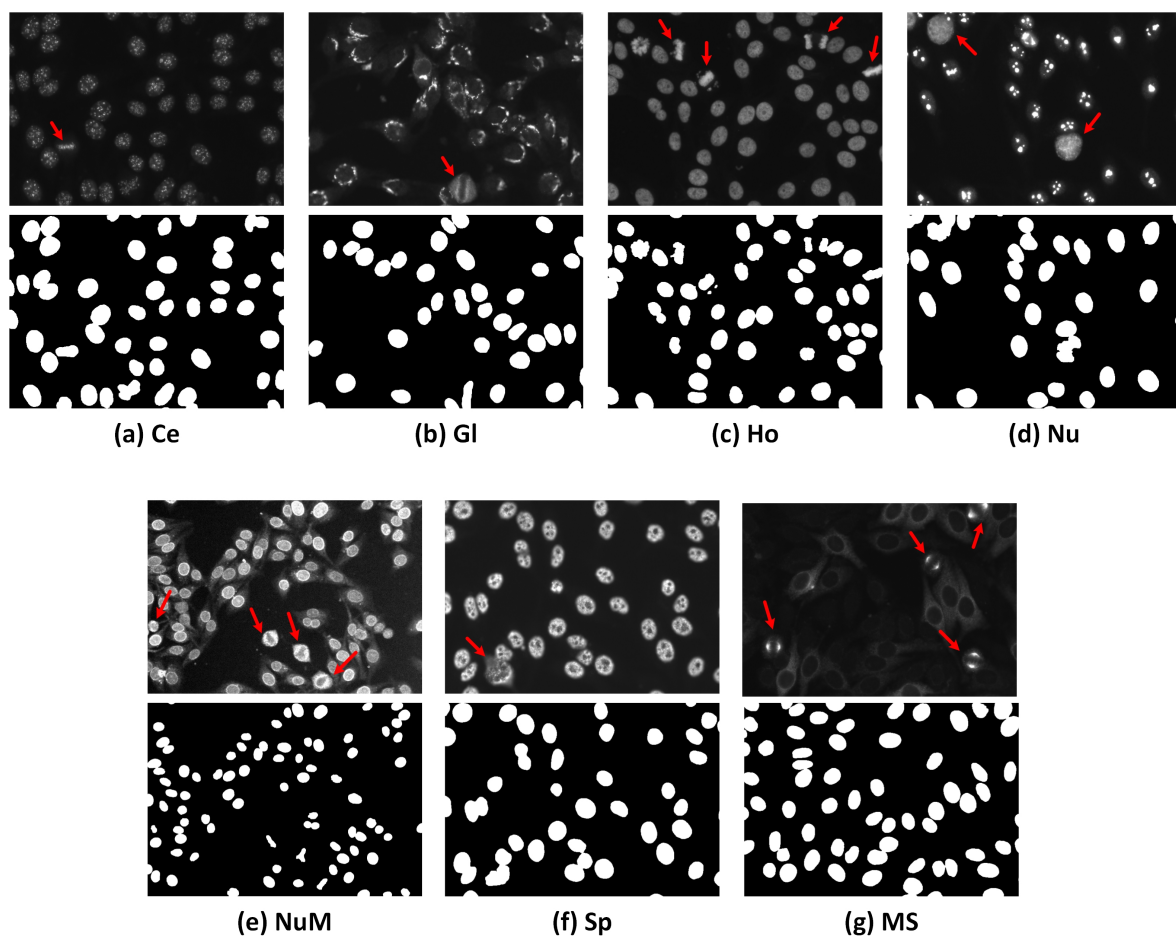


Figure 1. Examples of specimen images from the I3A Task-2 dataset. The first row shows the FITC images, while the second row shows their corresponding DAPI channel binary masks. Classes annotated as: (a) Centromere (Ce); (b) Golgi (Gl); (c) Homogeneous (Ho); (d) Nucleolar (Nu); (e) Nuclear Membrane (NuM); (f) Speckled (Sp); and (g) Mitotic Spindle (MS). The red arrows indicate the mitotic cells appearing in the HEP-2 specimen images.

The rest of this paper is organized as follows. A review of contemporary literature relevant to this study is presented in Section 2. Technical details of the proposed AI-based HEP-2 mitotic cell detection framework are provided in Section 3. The results and discussion of the experimental study are reported and discussed in Section 4. Finally, Section 5 summarizes the work findings and conclusion.

2. Related Works

Automatic detection of mitotic cells from various staining patterns of the entire HEp2 specimen images is challenging for building any CAD system for autoimmune disease [4,13,17]. The mitotic cell identification in the specimen images is a key to improving the precise slide preparation, indicating an important pattern similarity with specific interphase patterns that make the disease classification easier and present the rare unnoticed disorders during the microscopic inspection process by human experts [4]. The vast majority of previous works on HEp-2 image analysis have been proposed to solve the classification of interphase type HEp-2 staining patterns [4,13]. For instance, many initial studies attempted to approach this task using conventional learning methods, such as the works proposed in [12,18–21]. Alternatively, the later works proposed CNN-based approaches, which demonstrated superior performance over the conventional handcrafted-based learning methods [6,22–28].

2.1. HEp-2 Mitotic Cell Image Classification

Unlike the HEp-2 cell interphase patterns classification, a few research works have been conducted to solve the task of mitotic pattern detection from microscopic medical images for various diseases [10,13]. There are two kinds of deep learning approaches for predicting the mitotic cells: (1) HEp-2 single-cell image prediction, in which the individual cell images are segmented from the entire specimen image and processed to predict the class of each cell individually; and (2) HEp-2 specimen-image prediction, in which the entire specimen images are processed at once. For simplicity, the first and second approaches could be addressed as cell-level and specimen-level prediction scenarios, respectively. Figure 2 shows a schematic pictorial representation of both methods to process HEp-2 images. The cell-level prediction scenario is generally more accurate but needs exhausted user interventions regarding individual cell image extraction. Due to the availability of cell-level annotated datasets [13], cell-level-based approaches are commonly presented in the literature. Alternatively, the specimen-level scenario is preferable as a practical solution, especially with the scarcity of mitotic cell numbers in the HEp-2 samples. However, these methods have barely been studied in the literature due to the insufficiency of the annotated data [4]. The automatic detection of HEp-2 mitotic cells from the whole HEp-2 specimen images is important to develop reliable and feasible CAD systems to support the medical personnel in making rapid and accurate diagnosis decisions [4,13]. Moreover, improving the throughput capacity of the medical facilities reduces the variability among personnel/laboratories and minimizes the financial cost for such tests. In this section, we briefly describe the existing works for both cell-level as well as specimen-level prediction scenarios.

2.1.1. Cell-Level Prediction Scenario

Almost all existing works in the literature on HEp-2 mitotic cell detection attempted to address this task as a binary classification between mitotic and interphase cell-level images. Because the mitotic cells are a minority in all available datasets, most existing works proposed using conventional learning approaches for solving the unbalancing problem. For instance, Foggia et al. [9] manually collected a small balanced dataset comprising 126 images from both classes (i.e., mitotic and interphase). They proposed using textural and morphological descriptors with an AdaBoost classifier for this task, achieving a maximum classification accuracy of 86.5% in their dataset. Using a different imbalanced dataset, Percannella et al. [29] proposed a threshold selection technique called multi-objective optimization for selecting the best classification decision among different classifiers and using different training scenarios. Their classifiers were trained on a heterogeneous set of morphological, texture, and LBP features [30]. However, their performance evaluation demonstrated poor results. In another work, Iannello et al. [10] proposed using a hybrid multi-expert system (MES) and multi-objective optimization method by training several classifiers with different data balancing scenarios. However, their method showed a poor result for classifying the positive mitotic class, achieving a true positive rate of 51.7%.

Alternatively, Tonti et al. [31] suggested a completely unsupervised approach to differentiate the minority mitotic cell images from the remaining interphase cell images based on predefined rules on the morphological and texture GLCM [32] descriptors. The best classification accuracy from their method was achieved at 75.6%. Attempting to cope with data skewness, Miros et al. [11] proposed using both foreground fluorescent cell images with corresponding mask images to extract more representative features. Their classification algorithm was applied to a collection of shape, intensity, and texture features derived from the mask images.

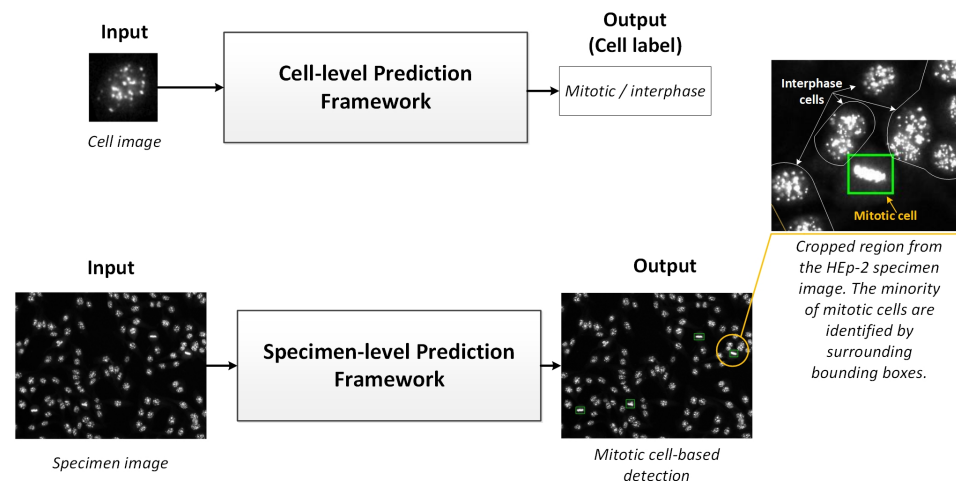


Figure 2. Mitotic cell prediction scenarios: Abstract view. The upper part shows the cell-level prediction framework, which requires cell image extraction overhead and then feeds them to the classification framework. In the lower pipeline, the prediction framework is designed to detect and classify mitotic cells from microscopic specimen images automatically and simultaneously.

Recently, this task was studied using various deep learning-based approaches. In [33], Gupta et al. extracted deep features from the different intermediate convolutional layers of the AlexNet. They pre-trained this network using the ImageNet dataset [34] and used the SVM for classification purposes. Afterward, Gupta et al. [35] proposed using a one-class SVM classifier to detect the nominal mitotic patterns among the majority of interphase patterns. In [36], Gupta et al. suggested using a Siamese deep network framework with a triplet loss function to address the unbalancing dataset by learning distance-based features. The SVM algorithm was used on the embedding-based features for classification. In [37], Gupta et al. recently proposed using a modified DCGAN generative model to generate new mitotic cell images for oversampling the minority class. Then, they extracted CNN-based and LM filter bank features from a mixture of real and manually selected GAN-synthesized mitotic images. For classification, they used the SVM algorithm. However, as shown by the evaluation results, the impact of adding the GAN-synthesized mitotic images was not clearly effective in improving the classification performance. In contemporary work, Anaam et al. [38] proposed an end-to-end deep learning model based on 1D-DCGAN on the feature space. In their work, a 1D-DCGAN model was trained to learn the feature distribution of the mitotic cells in the embedded space obtained by a CNN network. Then, the trained 1D-DCGAN generator was integrated within a CNN network during the training process to augment the mitotic sample features.

2.1.2. Specimen-Level Prediction Scenario

Unlike the cell-level scenario, the specimen-level scenario performs the prediction task over the entire HEP-2 specimen images [13]. A few works in the literature were introduced for approaching different tasks of the HEP-2 CADs pipeline using specimen-level images. For instance, Oraibi et al. [39] proposed a hybrid method that employs the VGG-19 CNN model [40] and the handcrafted feature of LBP [30] and joint motif

labels [41] to boost the discriminative capacity of CNN features. Cascio et al. [42] proposed a classification framework for the HEP-2 specimen images based on their individual cell images. Firstly, the cell images were extracted from the specimens using an active contours model algorithm (ACM). Then, a combination of intensity, shape, and texture features was used with the LDA selection method. Two-step classification was employed using class-aware binary SVMs and KNN. For specimen image segmentation and classification, Xie et al. [43] proposed a modified FCN [44] network that shares feature maps between the downsampling and upsampling parts of the network with a feature fusion strategy and additional skip connections. In a more recent work, Percannella et al. [45] proposed a U-Net-like model for jointly solving the tasks of specimen images intensity classification and segmentation on HEP-2 cell images in an end-to-end manner. A few works that proposed to solve segmentation and classification tasks of the interphase type of the HEP-2 specimen images [13].

However, to the best of our knowledge, the recent work presented by Gupta et al. [17] is the only existing work that partially addresses this task in specimen-level images. They proposed using a Faster R-CNN [16] algorithm to detect one type of mitotic cell called Mitotic Spindle to identify the Mitotic Spindle specimen images from the remaining interphase specimen classes. However, they ignore detecting other mitotic cells from the remaining different HEP-2 specimen classes [17]. Alternatively, for the first time, this work proposes an automatic deep learning-based framework for mitotic cell detection from all different classes of the HEP-2 specimen images: Centromere, Golgi, Homogeneous, Nuclear membrane, Speckled, Nucleolar, and Mitotic Spindle, which is essential to developing reliable and feasible HEP-2 CAD systems [4,13]. This work aims to contribute effectively toward developing such systems to support the medical staff in diagnosing rapidly and accurately, increasing laboratory throughput, and minimizing the cost of such tests.

2.2. GANs for HEP-2 Cell Image Analysis

Recently, increasing research has been conducted to study the use of GANs for different medical imaging applications involving detection tasks [46,47]. For example, Madani et al. [48] proposed a GAN-based semi-supervised learning approach to improve the detection performance of cardiac abnormality from X-ray imaging with less annotated data. Chen and Konukoglu [49] employed an adversarial auto-encoder model to detect lesion regions from brain MR images. Schlegl et al. [50] used an unsupervised fast AnoGAN (f-AnoGAN) to score the fitness of the unseen images to that of the normal images in the GAN-learned latent manifold to detect anomalies in the optical coherence tomography (OCT) images. For another imaging modality, Han et al. [51] proposed 3D Multi-Conditional GAN (MCGAN) to generate new volumetric nodules in lung CT images for augmentation purposes. Their MCGAN architecture was designed to conditionally optimize the generated nodules' position, size, and attenuation to enhance the performance of a cascaded 3D CNN-based object detector. Nonetheless, various other contemporary research suggested using GANs to boost the classification performance of different medical imaging modalities, such as works in [52–56].

Alternatively, a limited number of studies that used GANs to approach the HEP-2 cell image analysis tasks are found in the literature. For example, Li et al. [57] used a U-Net generator with a modified conditional pix2pix [58] network to improve the HEP-2 cell segmentation performance. Kastaniotis et al. [59] proposed using Teacher-network to guide the attention maps in the DCGAN [60] discriminator network to improve the quality of the generated HEP-2 cell images. In another work, Xie et al. [61] used the pix2pix [58] configuration to synthesize cell-level mask images to improve the performance of a cascaded CNN classification model. Later, Majtner et al. [28] employed a separated DCGAN model to synthesize new images of each HEP-2 interphase type class for augmentation purposes. In more recent work, Anaam et al. [6] studied the effectiveness of using different GAN models to generate new HEP-2 cell images for boosting the CNN classification performances. However, for the specific task of HEP-2 mitotic cell classification, the works

proposed by Gupta et al. [37] and Anaam et al. [38], introduced in Section 2.1.1, are the only published studies.

3. Materials and Methods

The end-to-end scenario of the proposed AI-based mitotic cell detection framework is depicted in Figure 3. This framework has four sequential steps to achieve the final prediction goal of mitotic cells from the entire HEP-2 specimen images. The execution processing steps include medical HEP-2 data collection and preparation, data pre-processing, mitotic cell annotation, and the detection step based on the AI models. The technical concept of each processing step is explained in detail in the following sections.

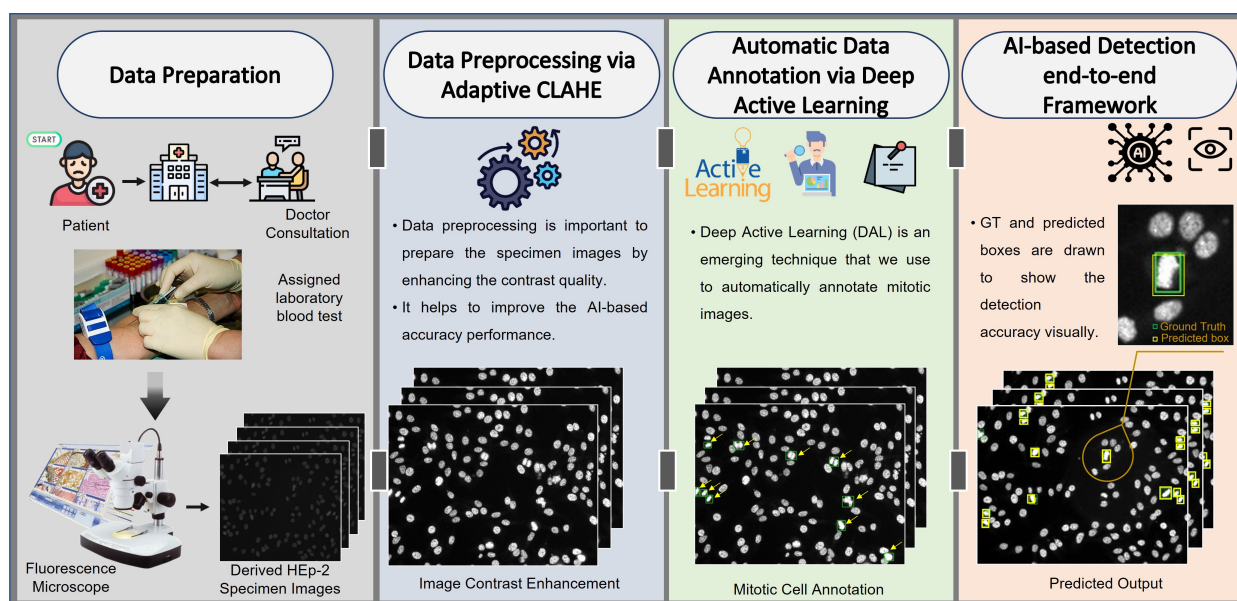


Figure 3. The proposed end-to-end AI-based framework for mitotic cell detection from the microscopic HEP-2 specimen images. The proposed framework involves four consecutive steps: data preparation, pre-processing, automatic labeling via deep active learning (DAL), and detection process via YOLO and Faster R-CNN.

3.1. Description of HEP-2 Medical Dataset

To achieve the goal of this work, the publicly available I3A Task-2 HEP-2 specimen dataset is used (Download link for the I3A dataset: <https://hep2.unisa.it/dbtools.html> (accessed on 20 March 2023)) [62]. This dataset was collected from 1001 cases with positive ANA samples at the Sullivan Nicolaides Pathology Laboratory in Australia. A monochrome camera mounted on a microscope with a Plan-Apochromat 20×/0.8 objective lens and Led illumination source was used for acquiring images with a resolution of 1388 × 1040 pixels from four distinct locations from each specimen slide. For each location, the imaging system acquires two different channel images: (1) the fluorescein-isothiocyanate (FITC) channel that carries the main contextual cell pattern information, which is typically used in the ANA tests; and (2) the 4',6-diamidino-2-phenylindole (DAPI) channel, which is a fluorescent stain used to delineate the HEP-2 cell nuclei and generate ground truth masks for the HEP-2 specimen images. In total, this dataset contains 1,008 specimen images consisting of seven classes: the first six classes are identified based on the majority of the interphase type patterns, which are Centromere (Ce), Golgi (Gl), Homogeneous (Ho), Nucleolar (Nu), Nuclear Membrane (NuM), and Speckled (Sp). However, the seventh Mitotic Spindle (MS) class is associated with the minority mitotic phase patterns that exist in those slides. Each image is provided with its corresponding DAPI-stained ground truth mask image. Each specimen image was assumed to have a single HEP-2 pattern (i.e., interphase type or the Mitotic Spindle class), and hence was given a single label.

Each specimen slide (i.e., four patches' images in the used dataset) should contain a few cells in the mitotic phase to validate the test preparation process. Thus, there is a need to manually annotate the mitotic cells across all specimen images to develop the proposed detection framework. Table 1 summarizes the data distribution over classes of the I3A Task-2 dataset. Figure 1 shows some examples of the HEp-2 specimen images with their corresponding DAPI binary masks.

Table 1. I3A Task-2 dataset description. Classes annotated as Centromere (Ce), Golgi (Gl), Homogeneous (Ho), Nucleolar (Nu), Nuclear Membrane (Num), Speckled (Sp), and Mitotic Spindle (MS).

Cell Phase	Interphase						Mitotic
HEp-2 Class	Ce	Gl	Ho	Nu	NuM	Sp	MS
No. of specimen images	200	84	212	204	40	208	60

As mentioned earlier, HEp-2 cells undergo the mitotic phase in a small portion of their lifecycle, making their appearance in the specimen images very rare. Moreover, depending on the staining patterns of the specimen, the associated mitotic cells can manifest in different fluorescence patterns, as illustrated in Figure 4. For instance, for the centromere and homogeneous staining patterns, the mitotic cells exhibit positive mitotic patterns in which the mitotic cell body is weakly/non-fluorescent. In contrast, the chromosome mass located in the middle part of the cell shows denser fluorescence, as shown in Figure 4a,b. It is unlikely that the mitotic cells that appear in the speckled, golgi, and nuclear membrane patterns exhibit opposite negative mitotic patterns. The negative mitotic pattern is characterized by a fluorescent cell body (i.e., peripheral part) and non/weak fluorescent in the collapsed chromosome mass located in the central part of the cell, as shown in Figure 4c–e. Moreover, mitotic cells associated with the nucleolar pattern could appear as negative patterns with diffuse cytoplasmic staining in some cases, as shown in Figure 4f, or might exhibit positive staining of the chromosomal region in other cases, as depicted in Figure 4g [63]. However, the Mitotic Spindle pattern is characterized by the stained spindle fibers that form a cone-shaped decoration for the poles of the mitotic cells, as shown in Figure 4h.

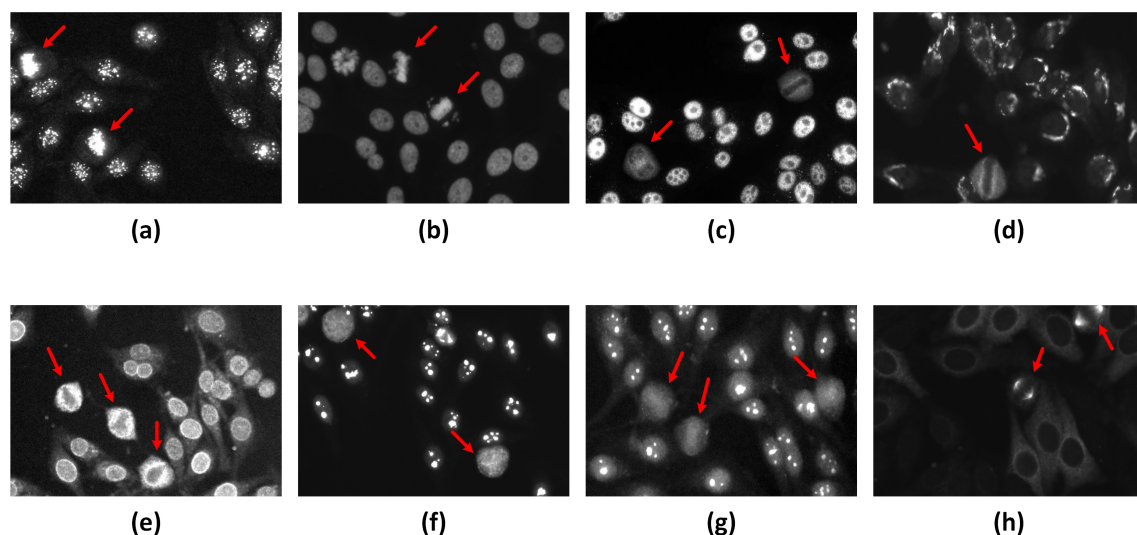


Figure 4. Examples of the different fluorescence patterns of the HEp-2 mitotic cells. The mitotic cells are annotated by red arrows. (a,b) Positive mitotic patterns that appear in Centromere and Homogeneous patterns, respectively. (c–e) Negative mitotic patterns appear in the Speckled, Golgi, and Nuclear membrane patterns, respectively. (f) Diffuse cytoplasmic mitotic patterns exhibited with Nucleolar pattern, while mitotic cells could appear with a positive pattern within the Nucleolar Specimen as shown in (g). (h) Mitotic Spindle pattern.

3.2. Medical Data Preparation

As it is known that the positive HEP-2 specimen images can be acquired with low contrast quality [12]. To improve the image contrast quality and help AI models for better prediction performance, contrast-limited adaptive histogram equalization (CLAHE) is applied [64,65]. Using CLAHE on the original HEP-2 specimen image, the histogram is stretched to cover the whole range of the intensity greyscale, enhancing the HEP-2 image contrast. Figure 5 shows an example of the pre-processing step based on the image processing histogram equalization technique. It is clearly shown that the image quality is enhanced by increasing the image contrast, and all cells become more distinguishable. Meanwhile, the structural similarity index (SSIM) indicates the quality of the image by improving its contrast with 92.65% and 89.13% for these two examples. As a part of the data pre-processing, the HEP-2 specimen images are normalized in the range of [0, 255] to improve the overall prediction performance of the AI-based models [65,66]. All images are resized using bi-cubic interpolation to scale their intensity pixels into the same range of 460×600 pixels.

3.3. Data Annotation via Deep Active Learning (DAL)

Deep active learning (DAL) is an emerging approach that automatically annotates unlabeled data to fine-tune the supervised learning algorithms accurately [67,68]. The concept of DAL is depicted in Figure 6. Indeed, data annotation is a time-consuming and expensive process when dealing with a large amount of unlabeled data. To avoid having manual annotation for the entire dataset instances, DAL could iteratively label the data samples that are assumed to maximize the deep learning model quality while minimizing the annotation efforts and time. In general, the importance of the DAL approach for automatic data labeling can be summarized as follows:

- **Efficient Use of Labeling Resources:** Labeling large datasets can be time-consuming and expensive. DAL can significantly reduce the manually labeled data needed to achieve high performance in a deep learning model, saving significant labeling resources.
- **Improved Model Performance:** DAL can lead to more accurate deep learning models. By selecting the most informative samples to label, the model can learn from high-quality data and improve its performance faster than it would with random sampling.
- **Active Learning:** DAL incorporates active learning, which is an approach that seeks to optimize the learning process by selecting the most informative samples to label. This can lead to faster convergence of the model and better generalization performance.
- **Human-in-the-Loop:** DAL can also incorporate human feedback into the labeling process. This can further improve the performance of the model by taking advantage of human expertise and intuition.

To detect the mitotic cells among all other cells in the same HEP-2 specimen image, the machine should learn the different patterns of mitotic cells of all the seven classes based on the ground truth (GT) labels. For the object detection task, the GT is the bounding box that surrounds the whole target object (i.e., mitotic cell) inside the HEP-2 specimen image. To automatically generate the GT labels for all HEP-2 specimen images, the DAL procedure includes the following strategies: initialization strategy, deep learning model selection, query strategy, and stopping criterion. For the initialization strategy, a sub-set of HEP-2 specimen images from all seven classes is carefully selected and manually annotated with the support of the GT binary DAPI masks shown in Figure 1. The GT bounding boxes are generated to embrace each mitotic cell inside a specific HEP-2 specimen image.

Three experts are involved in verifying the mitotic cell appearance in the specimen images to generate trustable GT labels. Two state-of-the-art deep learning detectors (i.e., YOLO and Faster R-CNN) are selected for the model selection as they are well-known and reliably used for detection tasks in the medical imaging domain [64,65,69,70]. Both detectors are pre-trained using the initial sub-set of annotated HEP-2 specimen images. Both detectors are evaluated during the iterative active learning process, and the best model is selected. Once the deep learning detector is pre-trained, the remaining unlabeled

specimen images are iteratively used to predict the potential mitotic cells. For the query strategy, the similarity between the predicted data in the first round of DAL and the initial subset is measured. Then, the specimen images with the highest similarity are selected by the machine, called the relevant or interesting sample images. These images are passed to the domain experts-in-loop to check, modify, and confirm the automated labeling process. They are able to manually adjust the generated bounding boxes' location, add missing bounding boxes for the unseeing objects, or even delete the wrong detected bounding boxes. For the stopping criterion, once all HEP-2 specimen images are correctly annotated, the DAL cycle is terminated. After dataset preparation and labeling steps, the dataset is randomly divided per class into 80% for training and 20% for testing. The validation set is represented by 10% of the training set. Since each patient has four different HEP-2 slices, we carefully group and include them in only one training, testing, or validation set to avoid any biases of the network's weights fine-tuning during the training time.

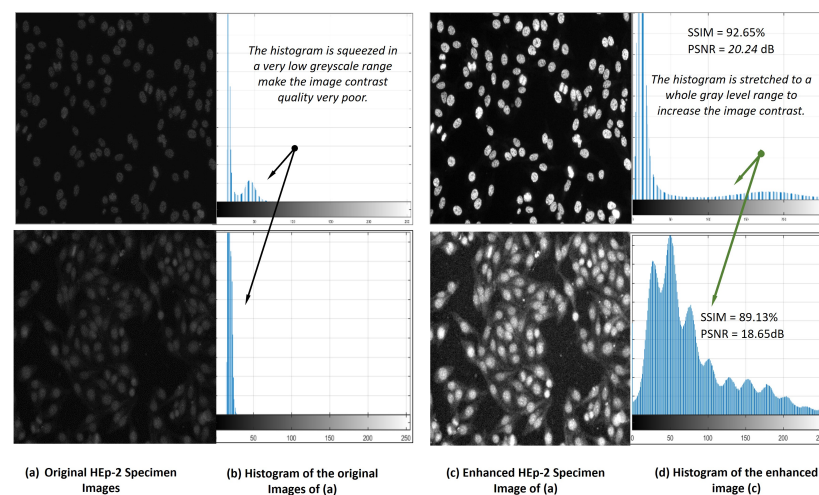


Figure 5. HEP-2 specimen image contrast quality improvement via contrast-limited adaptive histogram equalization (CLAHE) technique.

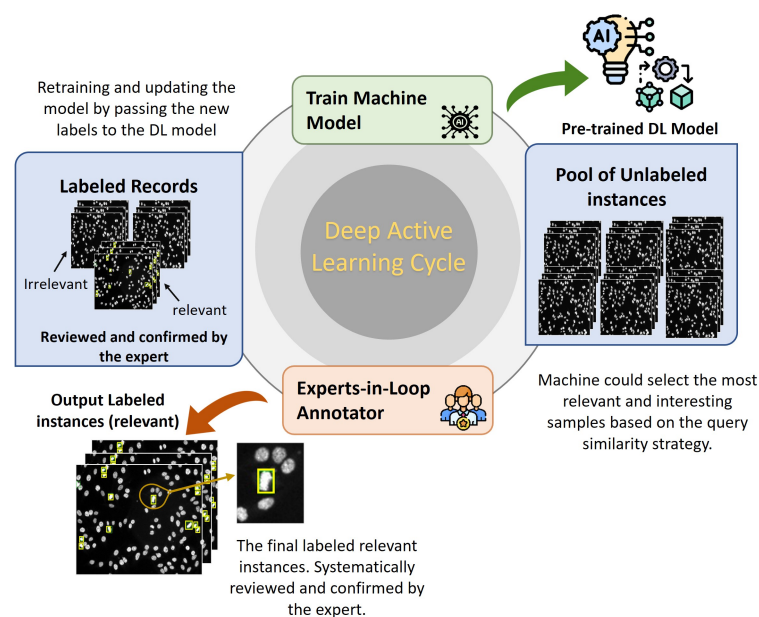


Figure 6. Deep active learning (DAL) strategy to automatically annotate the specimen HEP-2 images for the mitotic cell detection task.

3.4. Data Augmentation

We enlarged the size of the training dataset via various augmentation strategies to improve the prediction accuracy, avoid the over-fitting that may occur during the training time, and also to minimize the class imbalance problem [65]. We apply image photometric and geometric distortions, which are recent augmentation strategies used to increase the number of training instances [71]. For photometric distortion, the hue, saturation, and value of the HEP-2 specimen images are adjusted by 0.015, 0.7, and 0.4, respectively. To perform the geometric distortion, random scaling of 0.9, translation of 0.1, and rotation left-right of 0.5 are applied for training HEP-2 specimen images. In addition, Mosaic and MixUp augmentation methods are used with probabilities of 1 and 0.1, respectively. Figure 7 shows some augmented training HEP-2 specimen images.

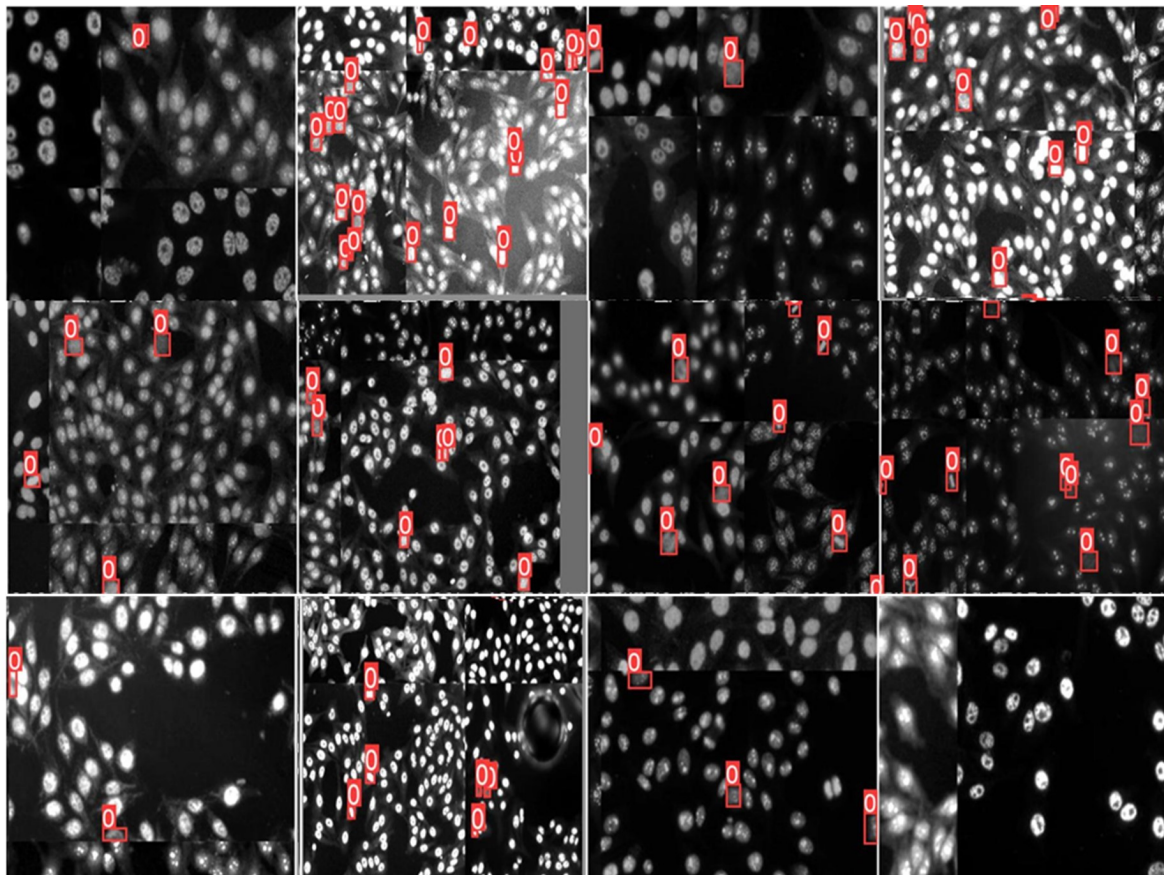


Figure 7. Examples of specimen images obtained by the used data augmentation techniques.

3.5. The Concept of the AI-based Mitotic Cell Detection

The detection of mitotic cells from the microscopic HEP-2 specimen images is achieved using robust deep learning detectors of YOLO [15] (i.e., YOLOv5) and Faster R-CNN [16]. These detectors are selected due to the following reasons. First, they have a high capability to detect multiple objects (i.e., mitotic cells) accurately and rapidly [65,70]. Second, simultaneously, they can estimate the confidence scores using the last neural network layer for each detected bounding box [72]. Third, these methods can detect objects with different sizes (predicts different bounding boxes' sizes), as they are equipped with a learnable bounding box coordinates regression function. Fourth, YOLO is known for its fast inference time, making it a good choice for real-time applications where speed is critical. Faster R-CNN also has a fast inference time compared to other R-CNN models. Fifth, YOLO, and Faster R-CNN have shown high accuracy in object detection tasks, with YOLO achieving higher mean average precision (mAP) on some datasets. YOLO and Faster R-CNN are known for their ability to detect objects at different scales and aspect ratios, making them good choices

for detecting objects in various environments. Both algorithms have well-established medical imaging implementations, demonstrating high performance in the detection tasks.

3.5.1. Mitotic Cell Detection Based on YOLO Predictor

Figure 8 shows the abstract view of the deep learning components of the YOLO predictor [15]. YOLO has a deep learning structure designed to directly detect the objects (i.e., mitotic cells in our case) from the input microscopic HEp-2 specimen images without user interventions [64,65,70]. Accurate and rapid object detection plays a crucial role in real-time AI applications. Thus, the one-stage detector based on the YOLO-v5 is selected to perform the mitotic cell detection for this research study. Generally, deep learning detectors involve two main stages [64,65,70]. The first stage is composed of a CNN-based backbone network to generate deep high-level features from the input image. The head predictor detects the bounding box coordinates surrounding the mitotic cells in the second stage. Recently, the YOLO deep learning structure was modified by inserting additional deep layers between the backbone and the head network to formulate a third part called the neck network [71], which plays a vital role in extracting millions of different features. This is the key behind the deep learning models to learn a huge number of features about the target objects and then detect them accurately.

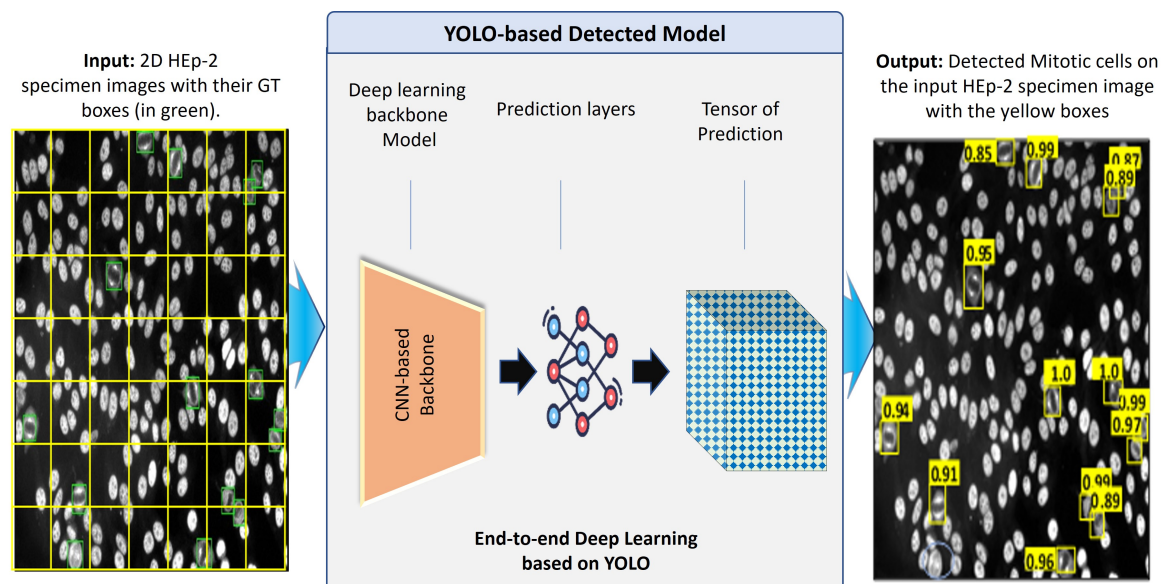


Figure 8. Deep learning YOLO detector for mitotic cell detection from the microscopic HEp-2 specimen images.

Several deep learning networks have been developed recently in the literature, such as VGG, ResNet, DenseNet, Swin Transformer, and CSP with SPP. Once the backbone network derives the deep features, the neck network interacts and plays its role between the backbone and heads to utilize the deep features effectively. The neck network has multiple bottom-up and top-down deep learning paths designed by different deep learning convolutional layers for reprocessing and rationally using the extracted features from the backbone network. Finally, the head layers are used to fine-tune and predict the bounding box coordinates of the potential objects in the input HEp-2 specimen image. In such cases, the output predicted boxes have several values, such as the bounding box coordinates: center (x,y), width (w), and height (h). Concurrently, the confidence score representing the probability of object existence is estimated for each box using the neural network layers. For each prediction output, a specific node in the network is responsible for optimizing and finding the best solution. At the end, all predictors are stored in the tensor of prediction, as shown in Figure 8.

3.5.2. Mitotic Cell Detection Based on Faster R-CNN

Faster Region-CNN (Faster-RCNN) [16] is an efficient deep CNN-based framework demonstrating excellent performance in object detection tasks. Recently, it has been widely applied for abnormality detection in medical imaging. For this work, the Faster R-CNN is adopted and used as a holistic framework for detecting mitotic cells from the entire HEP-2 specimen images. This algorithm takes the whole HEP-2 specimen images as an input and predicts the locations of mitotic cells in terms of bounding box coordinates. Simultaneously, it could predict the detection probability for each detected bounding box, indicating the object's existence probability.

Faster R-CNN deep architecture is composed of a backbone pre-trained network (CNN), region proposal network (RPN), region of interest (RoI) pooling, and head network. First, the backbone pre-trained CNN extracts the most representative deep feature maps from the input HEP-2 specimen images. This network takes the input images and yields their corresponding feature maps. Different CNN architecture could be employed as a feature generation component of the Faster R-CNN. In this work, we use a backbone of the ResNet-50 network [73], which is pre-trained on the large ImageNet dataset [34]. Then, the model is fine-tuned using our labeled dataset during the training process. Second, the region proposal network (RPN) is a simple network composed of two FC layers that take the feature maps extracted by the backbone network as input and generates class-agnostic region proposals (i.e., bounding boxes). In particular, RPN is trained to associate reference pre-defined fixed-sizes boxes placed uniformly throughout the image (known as anchors) with the truth bounding boxes by optimizing the regression error between them. The output of this network is a list of proposed bounding boxes' coordinates along with their probabilities of occurrence. In our implementation, we modified the RPN anchor boxes to be suitable to effectively cover the size range of the mitotic HEP-2 cells with the sizes of 32^2 , 64^2 , and 128^2 pixels. At the same time, the default settings of aspect ratios are kept intact as 1:1, 1:2, and 2:1, and the overlapping threshold of 0.7 is used. The RPN output is refined using a non-maximum suppression (NMS) algorithm to eliminate the duplicated or unwanted bounding boxes.

The region of interest (RoI) pooling layer uses the feature maps generated by the backbone network to extract features relevant to each proposed bounding box and resize them to a fixed size (i.e., $7 \times 7 \times C$, where C is the channel dimension of the feature map). Finally, the head network is a simple two fully connected layers used for final object classification and refinement of bounding boxes' coordinates. Deep learning components of the Faster R-CNN are trained using the labeled specimen images that provide the bounding boxes' coordinates for each mitotic cell inside the microscopic HEP-2 image. The abstract diagram of the deep learning Faster R-CNN architecture is depicted in Figure 9.

3.6. Experimental Setting

A multi-scale strategy is used during training to learn the mitotic cell prediction via multiple image spatial resolutions [74,75]. The hyper-parameters are experimentally optimized and selected to achieve the best mitotic cell detection accuracy. Whereas, a trial and error strategy is used to experimentally optimize and choose the training settings for both predictors separately [65,70].

For the YOLO detector, a mini-batch size of 32 with 100 epochs is selected to fine-tune the AI models. For optimization, the stochastic gradient descent (SGD) optimizer is selected with the initial learning rate (LR) of 0.01, the final one-cycle learning rate of 0.1, the momentum of 0.937, weight decay of 5×10^{-4} , warmup epochs of 3, warmup momentum of 0.8, and warmup initial bias learning rate of 0.1. For the Faster R-CNN detector, the implemented Faster RCNN model was trained using the SGD optimizer with an initial learning rate of 0.001 and a decay factor of 0.1 every 25 epochs. The model is trained for 100 epochs using the momentum of 0.9, weight decay of 0.0005, and batch size of 1. For both models, the gains of box loss, class loss, and object loss are adjusted to be

0.05, 0.3, and 0.7, respectively. The training IoU threshold and anchor-multiple threshold are selected as 0.2 and 4, respectively.

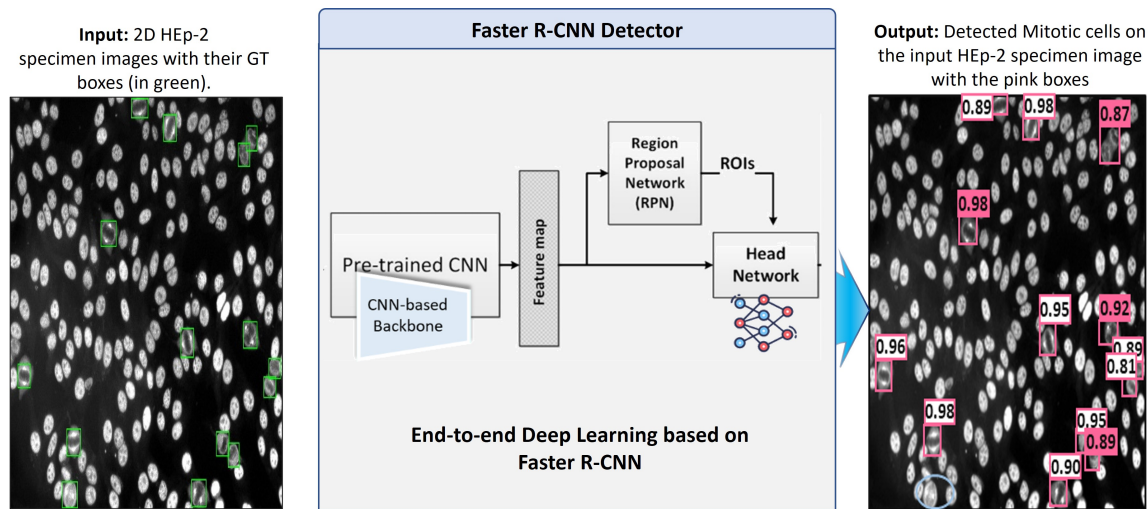


Figure 9. Deep learning Faster R-CNN detector for mitotic cell detection from the microscopic HEP-2 specimen images.

3.7. Performance Evaluation Strategy

The proposed deep learning detection framework is evaluated using a 5-fold cross-validation strategy using the testing sets. For each fold, the quantitative evaluation results of precision, recall, and mean average precession (mAP@0.5) are derived and reported, which are defined as follows:

- Precision: measures the ratio of the predicted true positives (TP) to all positive predictions (TP+FP):

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

- Recall: measures the ratio of the predicted true positive (TP) to the total positive (TP+FN):

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

- Mean Average Precision (mAP): measures the performance of object detection models by calculating the mean of average precision (AP) values over different recall values at a specific IoU threshold:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (3)$$

where N is the number of classes, and AP_i is the average precision of class i . AP represents a weighted average of precision at different recall values, and is calculated as follows:

$$AP = \sum_{k=0}^{n-1} [Recall(k) - Recall(k+1)] \times Precision(k), \quad (4)$$

where n is the number of thresholds, $recalls(n) = 0$, and $precisions(n) = 1$. During the calculations, the detection is considered TP if the IoU between the predicted and the GT bounding box is greater than the pre-specified threshold (i.e., 0.5 in our case).

Each evaluation metric is monitored during the training process to show the detection improvement over the training time. In addition, the training and validation bounding boxes regression and detection loss curves are also reported. Meanwhile, some qualitative results of the predicted mitotic cells on the entire HEp-2 specimen images are visually presented to show the potential predicted bounding boxes with their confidence probability scores.

3.8. Execution Environment

The experimental study is executed via a PC with a CPU of Intel(R) Core(TM) i7-10700KF @ 3.80 GHz, 32.0 GB RAM, six CPUs, and one GPU of NVIDIA GeForce RTX 3060.

4. Results and Discussion

4.1. Training/Validation Performance

The training and validation bounding box regression and object detection loss functions are recorded for each epoch during the fine-tuning optimization process of the deep learning models. This is to evaluate the deep learning parameter optimization progress during the training process. Figure 10 shows an example of the train/valid bounding box and object loss function curves over 100 epochs. Moreover, Figure 11 shows the evaluation metrics of precision, recall, and mAP that were recorded for the same training settings used to record the loss function curves of the validation set. It is clearly shown that the evaluation metrics improved with increasing the training epoch. This means that the deep learning detectors are learned well without any overfitting to the training data.

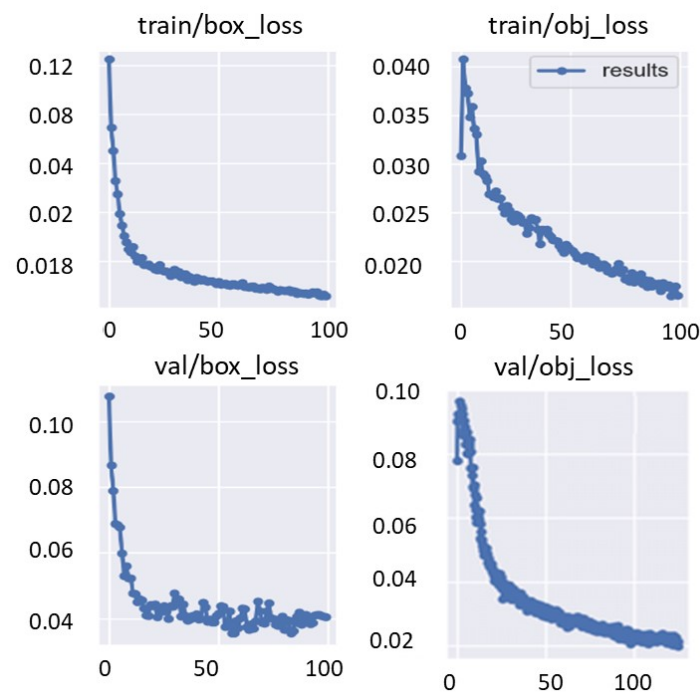


Figure 10. Training parameter optimization of train/valid bounding boxes regression and object detection losses functions. These curves are recorded using Fold-3 training/validation sets with YOLO detector.

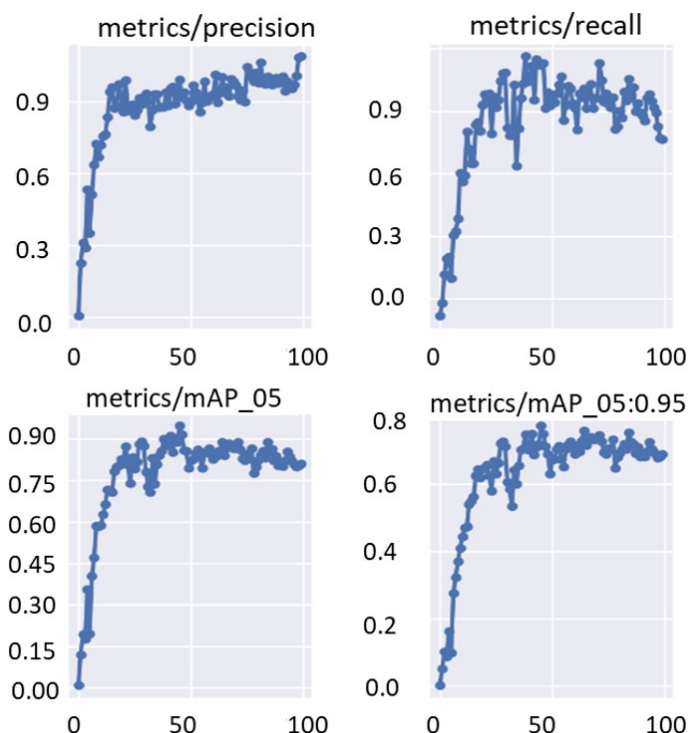


Figure 11. Validation performance in terms of precision, recall, and mAP. These curves are recorded using Fold-3 training/validation sets with the YOLO detector.

4.2. Evaluation of Data Labelling via DAL

For labeling HEP-2 mitotic cell bounding boxes, the DAL cycle is iteratively repeated four times. At each cycle, the evaluation of both detectors is checked and recorded. Figure 12 shows the capability of detectors to learn better in each round of the DAL cycle. The performance of both detectors is increased with each cycle because the number of annotated images is increased iteratively, as explained in Figure 6.

For the first DAL cycle, the subset of 200 HEP-2 specimen images is manually annotated by three experts as mentioned in Section 3.3. Both models of deep learning detectors are pre-trained using the selected 200 subset images at the first cycle. Using the rest of the unannotated images, both detectors are evaluated. After that, the highest similarity images with the first subset are passed and verified by the expert in the loop to check and modify the machine prediction and send them again for the second DAL cycle. This process is repeated four times until all HEP-2 specimen training images are totally annotated.

For the YOLO detector, the validation performance in terms of precision, recall, and mAP increased from 77.01%, 78.56%, and 72.98% in the first round into 90.26%, 89.53%, and 86.35% in the fourth round, respectively. Similarly, the validation performance of the Faster R-CNN during the DAL rounds increased from 70.15%, 69.89%, and 65.89% in the first round into 88.85%, 87.98%, and 85.61% in the fourth round in terms of precision, recall, and mAP, respectively. It is clearly shown that the YOLO detector could achieve a better detection evaluation over all DAL training cycles, as shown in Figure 12.

Figure 13 shows examples of the annotated HEP-2 specimen images for mitotic cells via the DAL strategy at the third cycle. The green boxes represent the machine prediction of the mitotic cells in each specimen image. In contrast, the expert-in-loop interactions with the machine to modify, remove, or add bounding boxes annotation for the mitotic cells are highlighted with yellow boxes inside circles in the figure. The experts in the loop remove the false positive (FP) cases that appear with cross signs in Figure 13. Furthermore, the expert added the missing boxes of the false negative (FN) cases that were missed by the machine but are TP. Such a case appears in the circle at the top of the middle image

in Figure 13. Moreover, the sizes of the detected boxes are adjusted to involve the whole object body, as shown in the circles in the left and right images of Figure 13. It is clearly shown that multiple mitotic cells from different classes could be effectively annotated via a machine with experts during the DAL cycles.

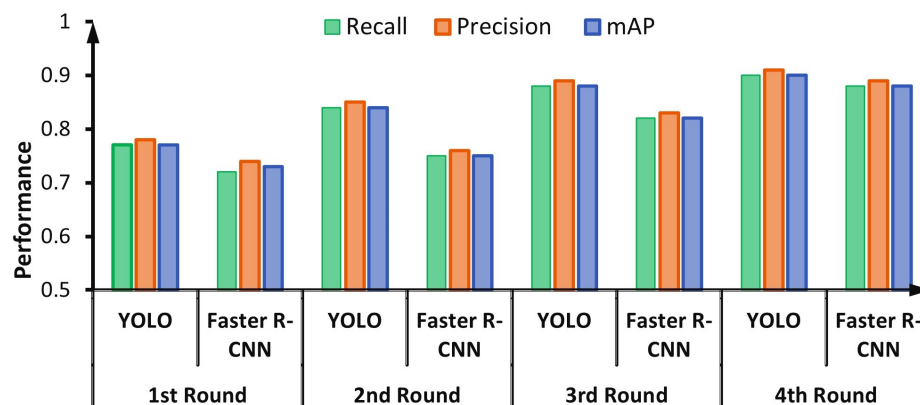


Figure 12. Iterative cycle detection performance for the deep active learning (DAL) during the mitotic cell automatic annotation process.

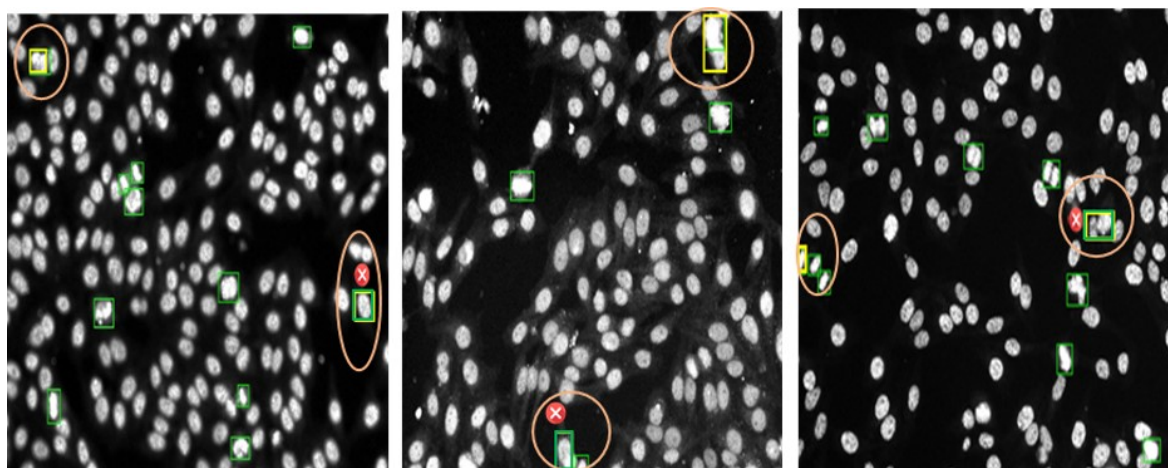


Figure 13. Examples of mitotic cell annotation process during the third DAL cycle. This is to show the Expert-in-Loop interaction with the machine to verify, modify, and confirm the labeling procedure. The left and middle images show the annotation using YOLO, while the right one is annotated via Faster R-CNN. The automated mitotic cell labels are depicted in the green boxes, while the adjusted labels by the expert-in-loop are superimposed by yellow boxes surrounded by orange circles. The falsely detected boxes (i.e., removed boxes) are annotated by small red circles with “x” signs.

4.3. Detection Results

The quantitative detection evaluation results using YOLO and Faster R-CNN detectors are summarized in Table 2. The mitotic cell detection results are recorded for each testing fold via the deep learning models trained using the same deep structure and training settings. For the YOLO detector, the average evaluation results in terms of recall, precision, and mAP are achieved with 90.01%, 88.30%, and 81.53%, respectively. For Faster R-CNN, the average detection performance for recall, precision, and mAP are 86.98%, 85.28%, and 78.50%, respectively. Indeed, both detectors could extract the mitotic cells from the different classes of the HEP-2 specimens: Centromere, Golgi, Homogeneous, Nuclear Membrane, Speckled, Nucleolar, and Mitotic Spindle. Using the YOLO algorithm for mitotic cell detection demonstrated relatively better results. As noted from Table 2, YOLO could correctly predict mitotic cells with average TP cases more than Faster R-CNN. Such results encourage adopting the YOLO-based framework as a practical solution for

supporting the detection of mitotic cells from the microscopic HEP-2 specimen images regardless of their interphase staining type.

Moreover, a paired *t*-test is used to statistically examine the significance of the differences between the results of both detectors. The null hypothesis (H_0) is that there are no significant differences between the results of both detectors, while the alternative hypothesis (H_a) is that their results are significantly different. Using a confidence level of $\alpha = 0.05$, *t*-test analysis demonstrated significant differences between the effects of YOLO and Faster R-CNN with p-values of 1.1467×10^{-16} , 4.6848×10^{-60} , and 1.0523×10^{-59} for the comparison between results of recall, precision, and mAP, respectively. As a result, based on the *t*-test analysis, the YOLO detector performed better than the Faster R-CNN on HEP-2 mitotic cell detection from the microscopic specimen images.

Table 2. Mitotic cell detection evaluation performance (%) using YOLO and Faster R-CNN detectors over a 5-fold testing protocol using the I3A Task-2 Dataset.

Fold Test	YOLO Detector			Faster R-CNN Detector		
	Recall	Precision	mAP	Recall	Precision	mAP
Fold-1	89.126	87.147	80.269	86.101	84.122	77.244
Fold-2	91.026	89.025	82.652	88.001	86	79.627
Fold-3	90.235	88.845	81.345	87.21	85.82	78.32
Fold-4	89.542	87.865	81.026	86.517	84.84	78.001
Fold-5	90.125	88.652	82.365	87.101	85.627	79.34
Average	90.011	88.307	81.531	86.986	85.282	78.506
SD	± 0.72	± 0.78	± 0.98	± 0.72	± 0.79	± 0.97

Both detectors demonstrate accurate cell object detection and bounding box coordinate prediction for the mitotic cells, as shown in the example depicted in Figure 14. The evaluation process shows a high overlap between the ground truth boxes and their corresponding predicted boxes, achieving an IoU score of larger than 0.9 for most of the detected mitotic cells for both YOLO and Faster R-CNN detectors, as shown in Figure 15a. However, the YOLO model could predict slightly more accurate boxes regarding the IoU counting 561 optimal detected cases compared to 551 for the Faster R-CNN. In contrast, the number of correctly detected mitotic cells at IoU = 90% is calculated to be 133 and 144 cases for YOLO and Faster R-CNN, respectively. In less than 50% of IoU, the Faster R-CNN predicts 10 cases, while YOLO detects only 2. Such cases are considered to be weakly detected since the IoU is very low. Meanwhile, the higher confidence score proves that deep learning models could predict the mitotic cells correctly with high objectness probability. Figure 15b shows the capability of both detectors to predict the mitotic cells with a high confidence score for most of the testing HEP-2 specimen images.

The pre-processed step for the HEP-2 specimen images using the CLAHE technique for contrast stretching improves the detection performance [65]. Specifically, such a pre-processing step supports the deep learning YOLO and Faster R-CNN detectors to achieve better evaluation results, as shown in Figure 16. The enhanced HEP-2 images gain improvements on their histogram in terms of the sharpness of the adjacent pixels to be more distinguishable, which improves the images' spatial contrast quality making the detection process easier. This might justify the improvement in the prediction performance using the CLAHE technique. For the YOLO detector, the detection evaluation results are improved by 0.08, 0.06, and 0.04 in terms of recall, precision, and mAP, respectively. Similarly, the detection results of the Faster R-CNN are improved by 0.08, 0.09, and 0.07 in terms of recall, precision, and mAP, respectively.

Figure 17 shows the impact of the adopted data augmentation techniques on the overall prediction performance of the proposed framework. The figure compares the performance of two dataset versions: original (i.e., without augmentation) and augmented datasets. It is

clearly noticeable that both detectors achieved better detection performance by augmenting the training dataset. For instance, the recall, precision, and mAP are improved by 0.10, 0.085, and 0.075, respectively, for the YOLO detector. Similarly, the Faster R-CNN yielded higher scores by 0.098, 0.092, and 0.083 regarding recall, precision, and mAP, respectively.

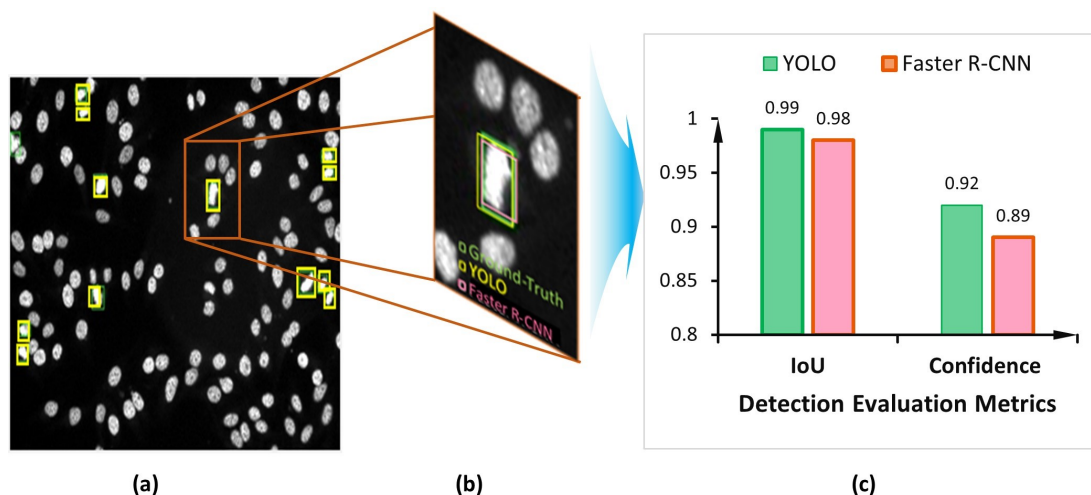


Figure 14. Example of the YOLO and Faster R-CNN qualitative detection results for the task of mitotic cell detection from the microscopic HEP-2 specimen images: (a) The output HEP-2 specimen image with the predicted mitotic cells; (b) Zoomed in a small region from the HEP-2 specimen image that contains a single mitotic cell; (c) Detection score comparison of both the YOLO and Faster R-CNN detectors.

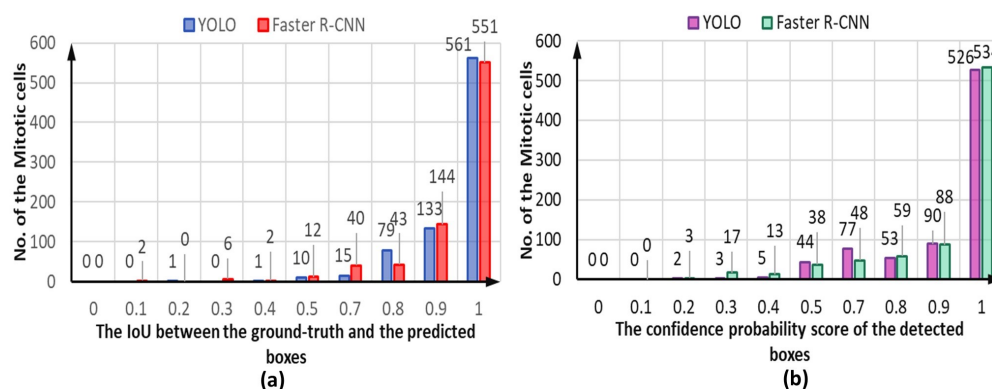


Figure 15. The capability of both YOLO and Faster R-CNN for accurate detection of the mitotic cells in terms of the (a) intersection over union (IoU) and (b) confidence probability threshold of the detected objects. These results are recorded from the third-fold trial.

Figure 18 shows examples of the HEP-2 mitotic cell detection results acquired using both YOLO and Faster R-CNN detectors. Both models could accurately detect most mitotic cells from different HEP-2 specimen images regardless of the class type. True positive (TP), false positive (FP), and false negative (FN) cases are annotated with different color circles, as shown in Figure 18. The TP cases represent the number of mitotic cells that are detected as mitotic. In contrast, the number of non-mitotic cells predicted as mitotic cells are the FP cases, while the FN cases mean the number of true mitotic cells the model does not detect. The falsely detected cases might be raised because multiple mitotic cells from different classes are involved in this study. The variability in the HEP-2 mitotic cell appearance is the top reason for detection failure. Whereas, HEP-2 mitotic cells manifest in different morphologies depending on the stage of the cell division process. Specifically, some mitotic cells show high morphological and textural similarity with other interphase patterns, which makes it challenging for the detection algorithm to identify all mitotic cells accurately.

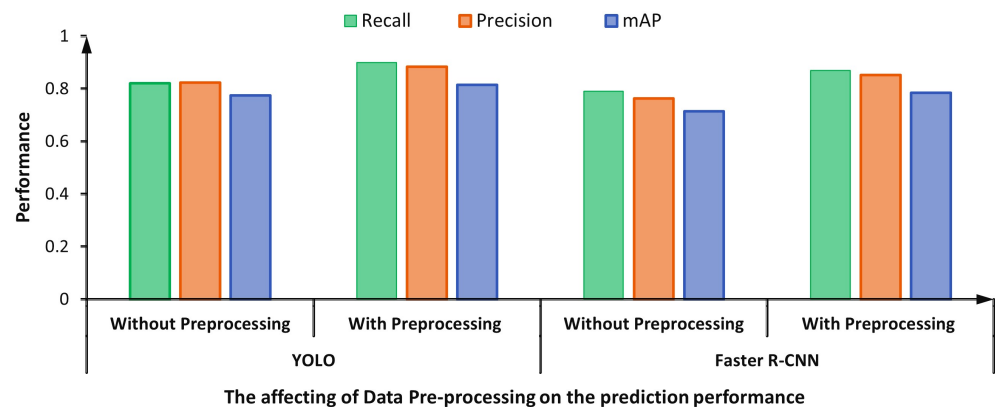


Figure 16. Effect of pre-processing on the mitotic cell detection performance: YOLO vs. Faster R-CNN.

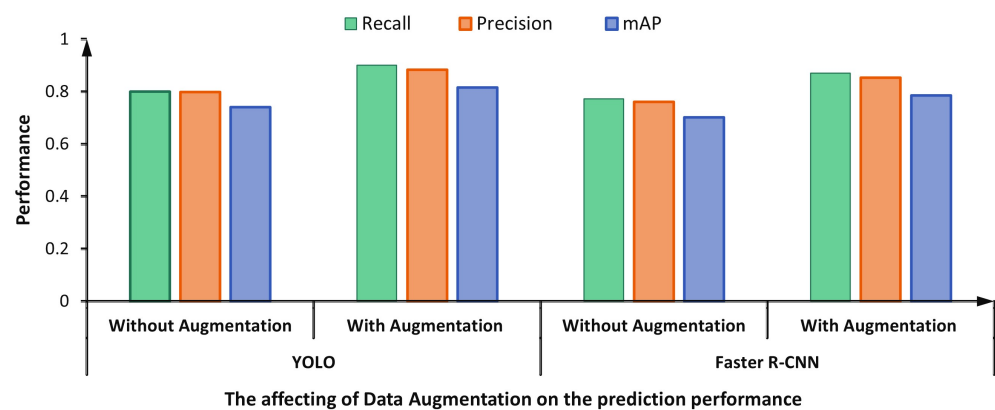


Figure 17. The prediction performance of the proposed framework via YOLO vs. Faster R-CNN using the original and augmented datasets.

The recent work of Gupta et al. [17] detects only the spindle mitotic cells ignoring the mitotic cells from other classes such as centromere, golgi, homogeneous, nuclear membrane, speckled, and nucleolar, which limits their prediction methodology to be applied as a practical solution. Alternatively, our proposed methodology aims to detect the mitotic cells from all HEp-2 staining classes. Such a comprehensive CAD system framework could be applicable in practical applications to minimize user intervention and labor concentration.

The execution time comparison for a single image detection is reported in Table 3. The YOLO detector needs 600 s to train per epoch, while the faster R-CNN requires around 1225 s. This is due to the deep learning structures of both detectors and the separated RPN network of the Faster RCNN model that has different capabilities to reach their training process in an epoch. For testing a single HEp-2 specimen image with the same number of mitotic cells, the YOLO algorithm outperforms the Faster R-CNN by 0.545 ms. The rapid detection strategy for YOLO makes it preferable for practical real-time applications. Meanwhile, YOLO could predict 5.56 frames per second, while Faster R-CNN predicts 3.03 specimen images per second regardless of the number of cells inside a single slide.

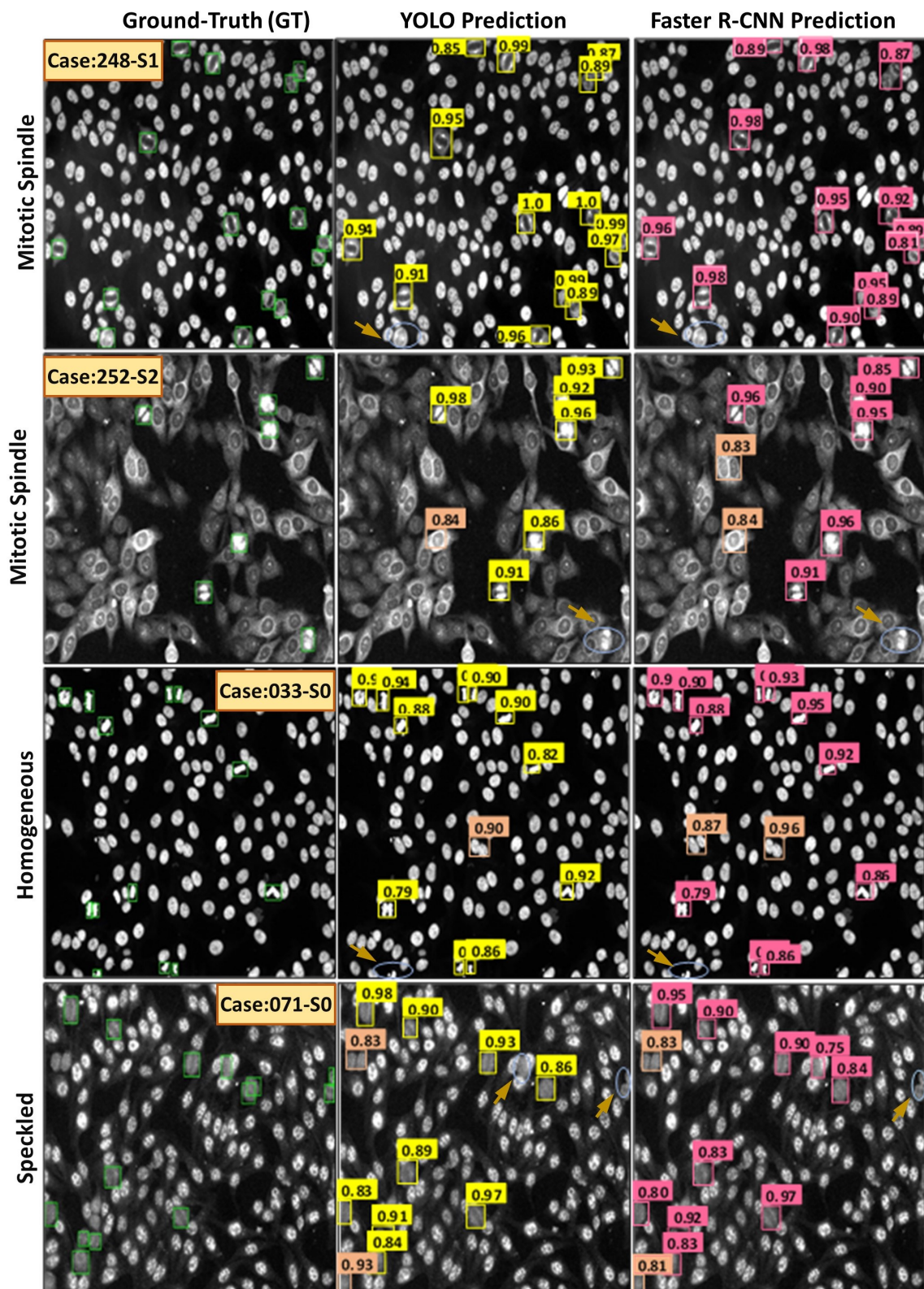


Figure 18. Some qualitative examples of detecting mitotic cells from the entire HEp-2 specimen images via YOLO and Faster R-CNN. The first column shows the ground truth bounding boxes in green. The true positive (TP) detected cases are labeled by yellow and pink boxes for YOLO and Faster R-CNN predictors, respectively. The false positive (FP) cases are highlighted by coral boxes for both detectors, while the false negative (FN) cases are highlighted by blue circles indicated with brown arrows.

Table 3. Execution time comparison of the adopted deep learning detection algorithms.

Deep Learning Detector	Dataset	Training Time/Epoch (Second)	Testing Time/a Single Specimen Image (ms)	FPS
YOLO	I3A Task-2	600	180	5.56
Faster R-CNN		1225	330	3.03

4.3.1. Comparison with Existing Works

Research studies for HEp-2 mitotic cell detection are limited in the literature. To the best of our knowledge, this is the first deep learning study to detect multiple mitotic cells directly from the microscopic HEp-2 specimen images of different HEp-2 interphase-type staining classes. As mentioned above, Gupta et al. [17] used the Faster R-CNN detector to only identify spindle mitotic cell types for classifying the Mitotic Spindle specimen images from other specimen image classes. However, they ignored the mitotic cells existing in the other specimen classes, making the comparison with their results unfair. Their model was trained to learn deep features from only the spindle mitotic cells, making their model biased toward specific features and ignoring other types of mitotic cells from the remaining classes.

On the other hand, all previous works that addressed the mitotic-interphase classification task were proposed on cell-level images. In particular, a small subset of carefully cropped mitotic cells was classified against another set of interphase cell images as provided in the I3A Task-3 dataset [62]. Accordingly, the results achieved by those methods are relatively higher as the binary classification problem is simple and straightforward. However, those methods require careful segmentation for all HEp-2 cell objects beforehand, which is not practically suitable. Therefore, the direct comparison between the results of our mitotic cell detection framework with those cell-level classification approaches is not fairly applicable. However, Table 4 provides a comparison summary between the proposed detection framework against other recent existing works.

4.3.2. Limitations and Future Work

The limitations of this work are associated with the scarcity of medical annotated datasets for models' training, which imposes the use of the DAL mechanism for automatic data annotation. As known, the DAL procedure requires intensive labor work, especially from the experts who are responsible for verifying and confirming the automatic labels of the machine. The similarity among the mitotic cells with some interphase cells confuses the deep learning detectors in some sense predicting FP and FN cases. Even so, we experimentally optimized the training parameters of both detectors to achieve the best detection performance and minimize the falsely detected cases. It is required to improve this framework to make detection for all cell types to develop a comprehensive one-shot HEp-2 cell detection framework, which is challenging due to the lack of an accurate annotation for all individual cell objects in each specimen image.

A prospective future extension of this work is to extend the annotation task to include the interphase cell patterns of the HEp-2 specimen images to develop an end-to-end trainable framework for simultaneously classifying mitotic and interphase cell types directly from the microscopic HEp-2 specimen images. In contrast to recent related works such as that of Xie et al. [61], which assign a single class label for each specimen image and cannot identify the mitotic cells among the interphase type slides, the prospective framework is proposed to instantly detect and classify all cell objects in the specimen image individually. Accordingly, specimen images with mixed patterns could be properly diagnosed. To achieve this goal, an intensive effort is required from field experts in the process of cell annotations and qualitative performance evaluation.

Table 4. Comparison between the proposed AI-based detection framework with other recent works. In table: BcA is the Balance class Accuracy.

Reference	Framework Description			BcA	mAP@0.5	
	Dataset	Task Definition	Features Extraction			Classifier
Tonti et al. (2015) [31]	Subset derived from MIVIA dataset	Binary classification (mitotic vs. interphase)	Morphological rules for selecting mitotic group + GLCM [32] for classification	Unsupervised K-means clustering	0.82	-
Miros et al. (2015) [11]	Subset derived from I3A dataset (cell-level images)		Combination of Shape (elliptical, roundness, etc.), size, intensity (i.e., histogram, mean, etc.) and texture (GLCM [32])	SVM	0.72	-
Gupta et al. (2019) [35]			CNN-based feature (using AlexNet [76]) + Decision-level fusion + BoW [77]	One-Class SVM	0.95	-
Gupta et al. (2019) [36]	I3A Task-3 dataset (cell-level images)		Siamese CNN- based feature (trained with triplet-loss [78])	SVM	0.82	-
Gupta et al. (2020) [37]			CNN-based feature (using AlexNet [76]) + BoW [77] + LM filter bank [79] + DCGAN [60] augmentation	SVM	0.98	-
Gupta et al. (2020) [17]	I3A Task-2 (Specimen images)	Detection of Mitotic spindle specimen type	Faster R-CNN [16]	0.98	-	
Proposed Framework *	I3A Task-2 (Specimen images (annotated using DAL))	Detection of mitotic cells from specimen images	End-to-end deep detection model	Recall: 0.90 Precision:0.88	0.815	

* The best evaluation results that derived using the proposed AI framework based on the YOLO predictor.

Inspired by the recent success of using GANs for medical data augmentation [56], a potential solution to mitigate the problem of annotated data scarcity is by optimizing conditional variants of GANs (e.g., pix2pix [58] or CycleGAN [80]) for augmenting the HEP-2 medical data [51]. Moreover, more recent detection algorithms such as EfficientDet [81] are planned to be implemented for this task in order to achieve the highest detection performance. Using the entire specimen image, all ANAs testing workflow steps (i.e., positive/negative classification, patterns intensity classification, mitotic cell detection, and HEP-2 cell patterns identification) could be integrated into a single practical end-to-end HEP-2 CAD system.

Moreover, due to the high variability of different HEP-2 cell patterns and the limitation of accurately annotating a large number of microscopic images, it is significant to provide an uncertainty quantification (UQ) measure for the DL detectors' predictions for a robust diagnosis [82,83]. Improving AI models to produce quantitative uncertainty estimates as additional information to their predictions is important to alert human experts about difficult or unknown prediction cases, which is critical for building reliable medical diagnosis systems [84,85]. Integrating such a UQ mechanism with the HEP-2 detection DL models are supposed to provide a real-world reliable HEP-2 CAD system that supports decision-making for autoimmune disease diagnosis.

5. Conclusions

Mitotic cell detection from the whole HEp-2 specimen images is a mandatory procedure to approve the correctness of the ANAs slide preparation and to support accurate classification for other interphase cell patterns. This paper presents an automatic deep learning framework to detect the mitotic cells directly from the microscopic HEp-2 specimen images. Due to the scarcity of the labeled dataset, the deep active learning (DAL) approach is involved in annotating the mitotic cells supporting the supervised deep learning models to achieve better detection performance. The experimental results demonstrated the superiority of the YOLO-based framework over the Faster R-CNN counterpart in this task. Data pre-processing via CLAHE improved the prediction performance of both YOLO and Faster R-CNN detectors by mAP of 4% and 7%, respectively. Moreover, the employed data augmentation techniques boosted the detection performance of the proposed HEp-2 mitotic cell detection framework by mAP of 7.5% and 8.3% for YOLO and Faster R-CNN, respectively. The proposed framework shows its capability and reliability to automatically detect multiple mitotic cells from diverse types of HEp-2 specimen images.

Author Contributions: Conceptualization, A.A., M.A.A.-a. and A.G.; methodology, A.A., J.H. and M.A.A.-a.; software, A.A. and J.H.; validation, M.A., N.A.S. and A.G.; formal analysis, A.A., M.A.A.-a. and J.H.; investigation, A.A. and M.A.A.-a.; resources, M.A. and N.A.S.; data curation, A.A. and M.A.A.-a.; writing—original draft, A.A. and M.A.A.-a.; writing—review and editing, A.A., M.A.A.-a., J.H., M.A., N.A.S. and A.G.; visualization, M.A.A.-a., A.A. and J.H.; supervision, A.G. and M.A.; project administration, M.A.; funding acquisition, N.A.S. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R407), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: To achieve this study, free public “I3A Task-2 HEp-2 specimen” dataset is used. Weblink: <https://hep2.unisa.it/dbtools.htmls> (accessed on March 2023).

Acknowledgments: The authors would like to express their gratitude to the funding projects of this work: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R407), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2022-00166402).

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

IIF	Indirect Immunofluorescence
HEp-2	Human Epithelial type 2 Cells
ANAs	Anti-Nuclear Antibodies
CTD	Connective Tissue Diseases
AI	Artificial intelligence
CAD	Computer-Aided Detection
FITC	Fluorescein-isothiocyanate imaging channel
DAPI	The 4',6-diamidino-2-phenylindole imaging channel
GT	ground truth labeling
DAL	Deep Active Learning
CLAHE	contrast-limited adaptive histogram equalization
MES	Multi-expert system

GLCM	Gray-Level Co-occurrence Matrix
DCGAN	Deep Convolutional Generative Adversarial Networks
CNN	Convolutional Neural Network
SSIM	Structural Similarity Index
RPN	Region Proposal Network
RoI	Region of Interest
NMS	Non-maximum Suppression
SGD	Stochastic Gradient Descent
mAP	Mean Average Precision
IoU	Intersection over Union
YOLO	You Only Look Once
TP	True Positive
FP	False Positive
FN	False Negative
UQ	Uncertainty Quantification

References

- Wiik, A.S. Anti-nuclear autoantibodies: Clinical utility for diagnosis, prognosis, monitoring, and planning of treatment strategy in systemic immunoinflammatory diseases. *Scand. J. Rheumatol.* **2005**, *34*, 260–268. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wiik, A.S.; Høier-Madsen, M.; Forslid, J.; Charles, P.; Meyrowitsch, J. Antinuclear antibodies: A contemporary nomenclature using HEp-2 cells. *J. Autoimmun.* **2010**, *35*, 276–290. [\[CrossRef\]](#) [\[PubMed\]](#)
- Meroni, P.L.; Schur, P.H. ANA screening: An old test with new recommendations. *Ann. Rheum. Dis.* **2010**, *69*, 1420–1422. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hobson, P.; Lovell, B.C.; Percannella, G.; Saggese, A.; Vento, M.; Wiliem, A. Computer Aided Diagnosis for Anti-Nuclear Antibodies HEp-2 images: Progress and challenges. *Pattern Recognit. Lett.* **2016**, *82*, 3–11. [\[CrossRef\]](#)
- Rigon, A.; Infantino, M.; Merone, M.; Iannello, G.; Tincani, A.; Cavazzana, I.; Carabellese, N.; Radice, A.; Manfredi, M.; Soda, P.; et al. The inter-observer reading variability in anti-nuclear antibodies indirect (ANA) immunofluorescence test: A multicenter evaluation and a review of the literature. *Autoimmun. Rev.* **2017**, *16*, 1224–1229. [\[CrossRef\]](#)
- Anaam, A.; Bu-Omer, H.M.; Gofuku, A. Studying the Applicability of Generative Adversarial Networks on HEp-2 Cell Image Augmentation. *IEEE Access* **2021**, *9*, 98048–98059. [\[CrossRef\]](#)
- Kearney, V.; Ziemer, B.P.; Perry, A.; Wang, T.; Chan, J.W.; Ma, L.; Morin, O.; Yom, S.S.; Solberg, T.D. Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks. *Radiology. Artif. Intell.* **2020**, *2*, e190027. [\[CrossRef\]](#)
- Bradwell, A.; Hughes, R.S.; Harden, E. *Atlas of Hep-2 Patterns and Laboratory Techniques*; Binding Site: Birmingham, UK, 1995.
- Foggia, P.; Percannella, G.; Soda, P.; Vento, M. Early experiences in mitotic cells recognition on HEp-2 slides. In Proceedings of the 2010 IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS), Bentley, Australia, 12–15 October 2010; pp. 38–43. [\[CrossRef\]](#)
- Iannello, G.; Percannella, G.; Soda, P.; Vento, M. Mitotic cells recognition in HEp-2 images. *Pattern Recognit. Lett.* **2014**, *45*, 136–144. [\[CrossRef\]](#)
- Miros, A.; Wiliem, A.; Holohan, K.; Ball, L.; Hobson, P.; Lovell, B.C. A Benchmarking Platform for Mitotic Cell Classification of ANA IIF HEp-2 Images. In Proceedings of the 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Adelaide, Australia, 23–25 November 2015; pp. 1–6. [\[CrossRef\]](#)
- Hobson, P.; Lovell, B.C.; Percannella, G.; Vento, M.; Wiliem, A. Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset. *Artif. Intell. Med.* **2015**, *65*, 239–250. [\[CrossRef\]](#)
- Rahman, S.; Wang, L.; Sun, C.; Zhou, L. Deep learning based HEp-2 image classification: A comprehensive review. *Med Image Anal.* **2020**, *65*, 101764. [\[CrossRef\]](#)
- Chazotte, B. Labeling nuclear DNA using DAPI. *Cold Spring Harb. Protoc.* **2011**, *2011*, pdb.prot5556. [\[CrossRef\]](#)
- Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
- Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [\[CrossRef\]](#)
- Gupta, K.; Bhavsar, A.; Sao, A.K. Identification of HEp-2 specimen images with mitotic cell patterns. *Biocybern. Biomed. Eng.* **2020**, *40*, 1233–1249. [\[CrossRef\]](#)
- Foggia, P.; Percannella, G.; Soda, P.; Vento, M. Benchmarking HEp-2 cells classification methods. *IEEE Trans. Med Imaging* **2013**, *32*, 1878–1889. [\[CrossRef\]](#)
- Ensafi, S.; Lu, S.; Kassim, A.A.; Tan, C.L. Accurate HEp-2 cell classification based on Sparse Coding of Superpixels. *Pattern Recognit. Lett.* **2016**, *82*, 64–71. [\[CrossRef\]](#)

20. Manivannan, S.; Li, W.; Akbar, S.; Wang, R.; Zhang, J.; McKenna, S.J. An automated pattern recognition system for classifying indirect immunofluorescence images of HEp-2 cells and specimens. *Pattern Recognit.* **2016**, *51*, 12–26. [\[CrossRef\]](#)
21. Ponomarev, G.V.; Kazanov, M.D. Classification of ANA HEp-2 slide images using morphological features of stained patterns. *Pattern Recognit. Lett.* **2016**, *82*, 79–84. [\[CrossRef\]](#)
22. Li, Y.; Shen, L. A deep residual inception network for HEp-2 cell classification. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 12–20.
23. Lu, M.; Gao, L.; Guo, X.; Liu, Q.; Yin, J. HEp-2 cell image classification method based on very deep convolutional networks with small datasets. In *Proceedings of the Ninth International Conference on Digital Image Processing (ICDIP 2017)*, Hong Kong, China, 19–22 May 2017; Volume 10420, pp. 864–869.
24. Lei, H.; Han, T.; Zhou, F.; Yu, Z.; Qin, J.; Elazab, A.; Lei, B. A deeply supervised residual network for HEp-2 cell classification via cross-modal transfer learning. *Pattern Recognit.* **2018**, *79*, 290–302. [\[CrossRef\]](#)
25. Shen, L.; Jia, X.; Li, Y. Deep cross residual network for HEp-2 cell staining pattern classification. *Pattern Recognit.* **2018**, *82*, 68–78. [\[CrossRef\]](#)
26. Li, Y.; Shen, L. HEp-Net: A smaller and better deep-learning network for HEp-2 cell classification. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2019**, *7*, 266–272. [\[CrossRef\]](#)
27. Vununu, C.; Lee, S.H.; Kwon, O.J.; Kwon, K.R. A Dynamic Learning Method for the Classification of the HEp-2 Cell Images. *Electronics* **2019**, *8*, 850. [\[CrossRef\]](#)
28. Majtner, T.; Bajić, B.; Lindblad, J.; Sladoje, N.; Blanes-Vidal, V.; Nadimi, E.S. On the effectiveness of generative adversarial networks as HEp-2 image augmentation tool. In *Proceedings of the Image Analysis: 21st Scandinavian Conference, SCIA 2019*, Norrköping, Sweden, 11–13 June 2019; pp. 439–451.
29. Percannella, G.; Soda, P.; Vento, M. Mitotic HEp-2 cells recognition under class skew. In *Proceedings of the Image Analysis and Processing—ICIAP 2011*; Maino, G., Foresti, G.L., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6979 LNCS, pp. 353–362.
30. Ojala, T.; Pietikäinen, M.; Harwood, D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.* **1996**, *29*, 51–59. [\[CrossRef\]](#)
31. Tonti, S.; Di Cataldo, S.; Macii, E.; Ficarra, E. Unsupervised HEp-2 mitosis recognition in indirect immunofluorescence imaging. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, Milan, Italy, 25–29 August 2015; Volume 2015, pp. 8135–8138. [\[CrossRef\]](#)
32. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. In *IEEE Transactions on Systems, Man, and Cybernetics*; IEEE: Piscataway, NJ, USA, 1973; pp. 610–621.
33. Gupta, K.; Bhavsar, A.; Sao, A.K. CNN based mitotic HEp-2 cell image detection. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies-BIOIMAGING*; SciTePress: Setúbal, Portugal, 2018; pp. 167–174.
34. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F.-F. ImageNet: A Large-Scale Hierarchical Image Database. In *Proceedings of the CVPR09*, Miami, FL, USA, 20–25 June 2009.
35. Gupta, K.; Bhavsar, A.; Sao, A.K. Detecting mitotic cells in HEp-2 images as anomalies via one class classifier. *Comput. Biol. Med.* **2019**, *111*, 103328. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Gupta, K.; Thapar, D.; Bhavsar, A.; Sao, A.K. Deep metric learning for identification of mitotic patterns of HEp-2 cell images. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, 16–17 June 2019; pp. 1080–1086. [\[CrossRef\]](#)
37. Gupta, K.; Thapar, D.; Bhavsar, A.; Sao, A.K. Effectiveness of GAN-based Synthetic Samples Generation of Minority Patterns in HEp-2 Cell Images. In *Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, Montreal, QC, Canada, 20–24 July 2020; pp. 1376–1379. [\[CrossRef\]](#)
38. Anaam, A.; Al-Masni, M.A.; Gofuku, A. 1D-DCGAN for Oversampling Minority Mitotic Patterns in HEp-2 Cell Images. In *Proceedings of the 2022 IEEE 4th Global Conference on Life Sciences and Technologies (LifeTech)*, Osaka, Japan, 7–9 March 2022; pp. 172–174.
39. Oraibi, Z.A.; Yousif, H.; Hafiane, A.; Seetharaman, G.; Palaniappan, K. Learning local and deep features for efficient cell image classification using random forests. In *Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, 7–10 October 2018; pp. 2446–2450.
40. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv Prepr.* **2014**, arXiv:1409.1556.
41. Prasath, V.S.; Kassim, Y.M.; Oraibi, Z.A.; Guiriec, J.B.; Hafiane, A.; Seetharaman, G.; Palaniappan, K. HEp-2 cell classification and segmentation using motif texture patterns and spatial features with random forests. In *Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 4–8 December 2016; pp. 90–95.
42. Cascio, D.; Taormina, V.; Raso, G. An automatic HEp-2 specimen analysis system based on an active contours model and an SVM classification. *Appl. Sci.* **2019**, *9*, 307. [\[CrossRef\]](#)
43. Xie, H.; Lei, H.; He, Y.; Lei, B. Deeply supervised full convolution network for HEp-2 specimen image segmentation. *Neurocomputing* **2019**, *351*, 77–86. [\[CrossRef\]](#)
44. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

45. Percannella, G.; Petruzzello, U.; Ritrovato, P.; Rundo, L.; Tortorella, F.; Vento, M. Joint Intensity Classification and Specimen Segmentation on HEp-2 Images: A Deep Learning Approach. In Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 21–25 August 2022; pp. 4343–4349.
46. Yi, X.; Walia, E.; Babyn, P. Generative adversarial network in medical imaging: A review. *Med Image Anal.* **2019**, *58*, 101552. [\[CrossRef\]](#)
47. Kazemini, S.; Baur, C.; Kuijper, A.; van Ginneken, B.; Navab, N.; Albarqouni, S.; Mukhopadhyay, A. GANs for medical image analysis. *Artif. Intell. Med.* **2020**, *109*, 101938. [\[CrossRef\]](#)
48. Madani, A.; Moradi, M.; Karargyris, A.; Syeda-Mahmood, T. Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 1038–1042.
49. Chen, X.; Konukoglu, E. Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders. *arXiv Prepr.* **2018**, arXiv:1806.04972.
50. Schlegl, T.; Seeböck, P.; Waldstein, S.M.; Langs, G.; Schmidt-Erfurth, U. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Med Image Anal.* **2019**, *54*, 30–44. [\[CrossRef\]](#)
51. Han, C.; Kitamura, Y.; Kudo, A.; Ichinose, A.; Rundo, L.; Furukawa, Y.; Umemoto, K.; Li, Y.; Nakayama, H. Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection. In Proceedings of the 2019 International Conference on 3D Vision (3DV), Québec, QC, Canada, 16–19 September 2019; pp. 729–737.
52. Xue, Y.; Ye, J.; Zhou, Q.; Long, L.R.; Antani, S.; Xue, Z.; Cornwell, C.; Zaino, R.; Cheng, K.C.; Huang, X. Selective synthetic augmentation with HistoGAN for improved histopathology image classification. *Med Image Anal.* **2021**, *67*, 101816. [\[CrossRef\]](#)
53. de Souza, L.A., Jr.; Passos, L.A.; Mendel, R.; Ebigbo, A.; Probst, A.; Messmann, H.; Palm, C.; Papa, J.P. Assisting Barrett's esophagus identification using endoscopic data augmentation based on Generative Adversarial Networks. *Comput. Biol. Med.* **2020**, *126*, 104029. [\[CrossRef\]](#)
54. Ju, L.; Wang, X.; Zhao, X.; Bonnington, P.; Drummond, T.; Ge, Z. Leveraging regular fundus images for training UWF fundus diagnosis models via adversarial learning and pseudo-labeling. *IEEE Trans. Med Imaging* **2021**, *40*, 2911–2925. [\[CrossRef\]](#)
55. Barile, B.; Marzullo, A.; Stamile, C.; Durand-Dubief, F.; Sappey-Marini, D. Data augmentation using generative adversarial neural networks on brain structural connectivity in multiple sclerosis. *Comput. Methods Programs Biomed.* **2021**, *206*, 106113. [\[CrossRef\]](#)
56. Chen, Y.; Yang, X.H.; Wei, Z.; Heidari, A.A.; Zheng, N.; Li, Z.; Chen, H.; Hu, H.; Zhou, Q.; Guan, Q. Generative adversarial networks in medical image augmentation: A review. *Comput. Biol. Med.* **2022**, *144*, 105382. [\[CrossRef\]](#)
57. Li, Y.; Shen, L. cC-GAN: A robust transfer-learning framework for HEp-2 specimen image segmentation. *IEEE Access* **2018**, *6*, 14048–14058. [\[CrossRef\]](#)
58. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
59. Kastaniotis, D.; Ntinou, I.; Tsourounis, D.; Economou, G.; Fotopoulos, S. Attention-aware generative adversarial networks (ATA-GANs). In Proceedings of the 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Zagorochoria, Greece, 10–12 June 2018; pp. 1–5.
60. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. In Proceedings of the 4th International Conference on Learning Representations, (ICLR), San Juan, Puerto Rico, 2–4 May 2016; pp. 1–16.
61. Xie, H.; He, Y.; Lei, H.; Kuo, J.Y.; Lei, B. Segmentation guided HEp-2 cell classification with adversarial networks. In Proceedings of the 2019 Computing, Communications and IoT Applications (ComComAp), Shenzhen, China, 26–28 October 2019; pp. 374–379.
62. Hobson, P.; Lovell, B.C.; Percannella, G.; Vento, M.; Wiliem, A. Classifying anti-nuclear antibodies HEp-2 images: A benchmarking platform. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 3233–3238.
63. Buchner, C.; Bryant, C.; Eslami, A.; Lakos, G. Anti-nuclear antibody screening using HEp-2 cells. *J. Vis. Exp. JoVE* **2014**, *88*, 51211.
64. Al-Antari, M.A.; Al-Masni, M.A.; Choi, M.T.; Han, S.M.; Kim, T.S. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *Int. J. Med. Informatics* **2018**, *117*, 44–54. [\[CrossRef\]](#)
65. Al-Antari, M.A.; Han, S.M.; Kim, T.S. Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Comput. Methods Programs Biomed.* **2020**, *196*, 105584. [\[CrossRef\]](#)
66. Zerouaoui, H.; Idri, A. Reviewing machine learning and image processing based decision-making systems for breast cancer imaging. *J. Med. Syst.* **2021**, *45*, 8. [\[CrossRef\]](#)
67. Feng, D.; Haase-Schütz, C.; Rosenbaum, L.; Hertlein, H.; Glaeser, C.; Timm, F.; Wiesbeck, W.; Dietmayer, K. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 1341–1360. [\[CrossRef\]](#)
68. Ren, P.; Xiao, Y.; Chang, X.; Huang, P.Y.; Li, Z.; Gupta, B.B.; Chen, X.; Wang, X. A survey of deep active learning. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–40. [\[CrossRef\]](#)

69. Al-Masni, M.A.; Al-Antari, M.A.; Park, J.M.; Gi, G.; Kim, T.Y.; Rivera, P.; Valarezo, E.; Choi, M.T.; Han, S.M.; Kim, T.S. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Comput. Methods Programs Biomed.* **2018**, *157*, 85–94. [[CrossRef](#)] [[PubMed](#)]
70. Al-Antari, M.A.; Hua, C.H.; Bang, J.; Lee, S. Fast deep learning computer-aided diagnosis of COVID-19 based on digital chest x-ray images. *Appl. Intell.* **2021**, *51*, 2890–2907. [[CrossRef](#)] [[PubMed](#)]
71. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 2778–2788.
72. Al-Antari, M.A.; Hua, C.H.; Lee, S. Rapid and accurate AI-based diagnosis to distinguish COVID-19 from eight other lung respiratory diseases. In *Artificial Intelligence Strategies for Analyzing COVID-19 Pneumonia Lung Imaging, Volume 1: Characterization Approaches*; IOP Publishing: Bristol, UK, 2022.
73. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; Volume 2016, pp. 770–778. [[CrossRef](#)]
74. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
75. Li, Z.; Tian, X.; Liu, X.; Liu, Y.; Shi, X. A two-stage industrial defect detection framework based on improved-yolov5 and optimized-inception-resnetv2 models. *Appl. Sci.* **2022**, *12*, 834. [[CrossRef](#)]
76. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of the Advances in Neural Information Processing Systems*; Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Sydney, Australia, 2012; Volume 25, pp. 1–9.
77. Yang, J.; Jiang, Y.G.; Hauptmann, A.G.; Ngo, C.W. Evaluating bag-of-visual-words representations in scene classification. In Proceedings of the International Workshop on Workshop on Multimedia Information Retrieval, Bavaria, Germany, 24–29 September 2007; pp. 197–206.
78. Bromley, J.; Guyon, I.; LeCun, Y.; Säckinger, E.; Shah, R. Signature verification using a “siamese” time delay neural network. *Adv. Neural Inf. Process. Syst.* **1993**, *6*. [[CrossRef](#)]
79. Leung, T.; Malik, J. Representing and recognizing the visual appearance of materials using three-dimensional textons. *Int. J. Comput. Vis.* **2001**, *43*, 29–44. [[CrossRef](#)]
80. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
81. Tan, M.; Pang, R.; Le, Q.V. Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 10781–10790.
82. Leibig, C.; Allken, V.; Ayhan, M.S.; Berens, P.; Wahl, S. Leveraging uncertainty information from deep neural networks for disease detection. *Sci. Rep.* **2017**, *7*, 1–14. [[CrossRef](#)]
83. Abdar, M.; Pourpanah, F.; Hussain, S.; Rezazadegan, D.; Liu, L.; Ghavamzadeh, M.; Fieguth, P.; Cao, X.; Khosravi, A.; Acharya, U.R.; et al. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Inf. Fusion* **2021**, *76*, 243–297. [[CrossRef](#)]
84. Abdar, M.; Fahami, M.A.; Rundo, L.; Radeva, P.; Frangi, A.F.; Acharya, U.R.; Khosravi, A.; Lam, H.K.; Jung, A.; Nahavandi, S. Hercules: Deep Hierarchical Attentive Multilevel Fusion Model With Uncertainty Quantification for Medical Image Classification. *IEEE Trans. Ind. Inform.* **2022**, *19*, 274–285. [[CrossRef](#)]
85. Shamsi, A.; Asgharnezhad, H.; Jokandan, S.S.; Khosravi, A.; Kebria, P.M.; Nahavandi, D.; Nahavandi, S.; Srinivasan, D. An uncertainty-aware transfer learning-based framework for COVID-19 diagnosis. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 1408–1417. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.