**Research Article**

Yasemin ARDICOGLU AKISIN*, Nejat AKAR and Mert Burkay COTELİ

**Decision support system for the classification of Downey cells as a pre-diagnostic tool**

https://doi.org/10.1515/tjb-2023-0035
Received February 9, 2023; accepted November 22, 2023; published online December 19, 2023

**Abstract**

**Objectives:** Epstein–Barr virus (EBV) is a member of the herpes virus that causes infectious mononucleosis (IM). Downey cell is the atypical lymphocyte of IM and can be seen in various conditions. Peripheral blood smear (PBS) microscopic evaluation is used to identify Downey cells. A lack of experienced professionals or professional errors may obstruct early and accurate diagnostics for the microscopic evaluation. The main objective of this study is to create a decision support system by digitizing the PBS samples. A general tool providing an inexpensive and measurable solution is envisioned to analyze the PBS samples in detail to give alerting flags to prevent missing Downey cells in manual analysis.

**Methods:** The PBS dataset collected was split into Downey positives and negatives. The negative set consisted of 5 leucocyte subtypes. Mantiscope, a cloud-based slide scanner system, was used to collect images from the physical PBS samples. Clinically and laboratory-conﬁrmed 35 IM patients and 124 healthy PBS slides were selected for this procedure. A number of cell counts were obtained after the application of annotation and augmentation methods, and a partially balanced dataset was created for the artificial intelligence (AI) network training. The veriﬁcation steps included the calculation of sensitivity, speciﬁcity, and Cohen’s kappa metrics from the partitioned testing set that was not used during training. A validation process was also performed over the manually identiﬁed PBS samples to measure whether the algorithm noticed the samples or not.

**Results:** After testing this setup, we have observed 98 % sensitivity and 99 % speciﬁcity for Downey cells. According to the validation procedure of Downey positive and negative samples that were carried out by the physicians, a sensitivity of 57 %, speciﬁcity of 100 %, and Cohen’s kappa value of 0.5 were observed. Besides, the accuracy was found to be 66 % according to the physicians’ evaluations employing the digital images which were identiﬁed by Mantiscope.

**Conclusions:** Decision support systems can alert the physician for Downey cells and increase the rate of true diagnosis in PBS evaluation. A higher sensitivity and speciﬁcity for the detection of Downey cells would be achieved. However, the variance over the dataset is a constraint for effective diagnosis. As the annotation and AI development process continues to collect more data from patients, the model can be updated for future releases.

**Keywords:** Epstein–Barr virus; pre-diagnostics tool; artiﬁcial intelligence; Downey cells; peripheral blood smear

**Introduction**

Epstein–Barr virus (EBV), a member of the human herpesvirus family, is a linear, double-stranded DNA virus that was initially isolated from a cultured Burkitt lymphoma cell line by Epstein et al. in 1964. EBV primarily replicates in lymphocytes but may also replicate in the epithelial cells of the pharynx and parotid duct. The infection with EBV, which leads to infectious mononucleosis (IM), is more frequent during childhood, and it spreads primarily by saliva (oral transmission). The incubation period is four to eight weeks. Typical features of IM include fever, pharyngitis, adenopathy, malaise, and an atypical lymphocytosis. Splenomegaly, hepatomegaly, jaundice, and splenic rupture can occur in patients with IM, but these complications are rare [1–3].

Diagnosis of IM is made by clinical examination, revealing the classic triad of fever, lymphadenopathy, and pharyngitis, and by laboratory based on the findings,
including the presence of atypical lymphocytosis in Peripheral Blood Smear (PBS), heterophile antibodies, and EBV-specific antibodies (anti-VCA, anti EBN, and anti-EA to EBV-associated antigens) [4]. Atypical lymphocytes, called Downey’s cells, are larger than mature lymphocytes with activated blue cytoplasm adherent around the erythrocytes. Most are CD8+ cytotoxic T cells, and their presence in the PBS helps in the diagnosis of IM [5]. McKinlay and Downey published the description of these cells in 1923, and it has become a standard laboratory test for the diagnosis of infectious mononucleosis [6]. Besides IM, they can be seen in EBV, TORCH, hepatitis B, Syphilis, HIV, and SARS-CoV-2 infections, autoimmune disorders, Hodgkin’s disease, sarcoidosis, drug-induced reactions, and immune reactions [7, 8].

PBS is a laboratory procedure that involves the cytology of blood cells smeared on a slide. It is crucial for diagnosing various clinical diseases and revealing peripheral blood cell morphology, aiding in the early detection of primary and secondary blood-related issues [9]. In some of the disorders, the gold standard is the microscopic evaluation of PBS. Although complete blood count with differential (CBC-D) is another frequently used and one of the most basic tests guiding physicians, it cannot replace the morphological findings obtained by peripheral smear examination [10].

From the view of recent advances in digital hematology, data-driven AI techniques open new perspectives by providing rapid and reliable techniques with the help of a large amount of annotated and clear data collected from clinical studies. The improvement of early diagnostic tools and examination methods via microscopy is vital to accelerate the diagnostics processes.

Current infectious disease diagnostic processes conducted worldwide require expensive and scalable medical devices, consumables, and trained personnel. Even a small reduction in efficiency in these processes can lead to high health-related and resource costs. Medical diagnostics processes may differ depending on the intra- and inter-observer differentials [11]. Using microscopy of PBS in the diagnostics processes of infectious diseases is time-consuming for physicians. Unlike diagnostics with these limitations, a PBS examination in which the entire sample is analyzed with the help of technological advances and more objective evaluations will provide results closer to ideal. It will be used to identify the patients at risk for infectious diseases. A general tool providing an inexpensive and measurable solution is envisioned to analyze the samples in detail outputs alerting flags to be used in the diagnostics processes by physicians.

AI techniques aim to create decision support systems while helping physicians increase their sensitivity and specificity for diagnostics. The first usage of AI in hematology starts with examining laboratory tests to assist physicians in education and diagnostics workflows. These first three systems were created and installed in European hospitals 1995 for peripheral blood interpretation, flow cytometry immunophenotyping, and bone marrow reporting [12]. While these knowledge-based systems were created with the old machine learning techniques, recent technological advances offer deep learning (DL) based object detection and classification, which are totally dependent on the data volume and consistencies. Therefore, the main motivation is to collect a large variety of clear annotated data to assist the neural network (NN) models to train.

A DL-based blood cell image analyzer is used to attain a decision support system that picks the relevant cell types. Output reconciliation is a pre-diagnostic tool that would create abnormality labels associated with specific diseases. The main purpose of this study is to create a decision support system (using AI) that can alert the physician of the presence of Downey cells to increase the pre-diagnostics rate.

Materials and methods

The PBS sample collection process has been implemented in XXXX University Faculty of Medicine Department of Pediatrics and Clinical Biochemistry for the first 6 months of 2021. These PBS samples were prepared manually via smearing and staining [May Grünwald’s eosin–methylene blue (Sigma-Aldrich, Merck KGaA, Germany) for 5 min and Giemsa’s azur eosin methylene blue (Sigma-Aldrich, Merck KGaA, Germany) for 15 min].

The hardware used in this work is a slide scanner. The Mantiscope Slide Scanner was designed and manufactured for health organizations, demanding digitalization of the physical samples to share them with health professionals [13]. The electronic digital microscope device can autofocus on the slide samples. The user can move over the physical samples manually, tag different types of cells on images via AI, and store and share the images. Mantiscope’s cloud-driven system contains a large variety of cell morphology analytics (47 types of cells), a decision support system developed by hematology professionals.

Ethics committee approval was received for this study from the XXXX Clinical Research Ethics Committee (Ethics Committee Decision No. 095) on 16 March 2022.

Data collection and annotation

A dataset (PBS slides) was collected from 35 patients with clinically and serologically confirmed IM and 124 healthy controls. These samples were examined via manual microscopy for diagnostics. The negative set comprising 5 subtypes (Basophil, Neutrophil, Lymphocyte, Monocyte, Eosinophil) was selected from the WBC types. During this work, Mantiscope, a cloud-based slide scanner system, has been used to collect images from the physical PBS samples in 100× optical and 10× digital magnification with immersion oil [13, 14].

We have selected an analyzable region to collect 15 and 12 images in width and height, respectively, on equally spaced places from each
slide. Physicians’ annotations over the digitized sample images were collected through the cloud-based software. Each Downey cell annotation was verified by the joint decision of two expert physicians. According to the reconciliation, if there was a conflict between the physicians’ evaluation, the cell was disregarded and remarked as a monocyte or lymphocyte. As a result, the collected dataset is formed of only clear and visually recognizable cells. The preliminary diagnostics results for the evaluation were obtained from manual microscopy. Approximately 28,800 images were controlled to create a balanced cell dataset comprising the positive and negative sets. The workflow used in the annotation process for the procedure can be seen in Figure 1.

According to the annotation results, we have obtained an unbalanced condition in terms of the cell counts due to the different frequency of occurrence of the cell types. Therefore, augmentation methods were utilized to create a balanced dataset for DL network training. Augmentation is a routine process for DL-based training. It is used not only for imbalanced data but also for cases to increase the feature diversity of the dataset. In this work, transferring the physician’s expertise into the DL network effectively is the main objective. Accordingly, subsets of the overall data are generated to satisfy the balanced condition. Principal Component Analysis (PCA) creates equality for the number of cell types to feed into training [15]. It is a dimensionality reduction of the features of the full set. The cell counts obtained after annotation and augmentation are as follows (Table 1).

### Neural network architecture

The state of art for object detection using NN contains segmentation and classification tasks. Region proposals for the classification require extra effort over the processing units due to the possibility of having many regions to be classified. R-CNN, and then Faster R-CNN have data-adaptive region proposal forms by using extra computing power [16, 17]. Despite Faster R-CNN’s efficient object proposals and faster computations, it would not be preferable for smaller objects. Therefore, YOLO NN architecture is selected for the detection of blood cells because of having a region proposal model in pyramidal search that would promote different sizes of object detection and faster computations [18]. Yolov4-tiny is selected as the reference model by inserting a spatial

---

**Table 1**: WBC annotation counts and total number of cells in each cluster.

<table>
<thead>
<tr>
<th>Cell type (positive)</th>
<th>Annotations</th>
<th>w. augmentation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basophil</td>
<td>166</td>
<td>664</td>
<td>830</td>
</tr>
<tr>
<td>Neutrophil</td>
<td>7,087</td>
<td>28,348</td>
<td>35,435</td>
</tr>
<tr>
<td>Lymphocyte</td>
<td>7,162</td>
<td>28,648</td>
<td>35,810</td>
</tr>
<tr>
<td>Monocyte</td>
<td>776</td>
<td>3,104</td>
<td>3,880</td>
</tr>
<tr>
<td>Eosinophil</td>
<td>522</td>
<td>2,088</td>
<td>2,610</td>
</tr>
<tr>
<td>Downey</td>
<td>172</td>
<td>688</td>
<td>860</td>
</tr>
</tbody>
</table>

---

**Figure 1**: The workflow from scanning via mantiscope to testing.
attention module that provides extra information about cell-type positions over the images [19]. For spatial attention, average-pooling and max-pooling operations are applied along the channel axis, and they are concatenated to generate another feature extractor for the cell types [20]. Spatial attention is crucial for blood cell classification due to cell eliminations at the boundaries and making a comparison between the dimensions of Red Blood Cells (RBC). For example, physician examinations for the RBC would differ according to their dimensions when compared to the other ones.

The parameters such as number of classes (cell types), filter size, batch size, and subdivisions used for training are 6, 81, 64, and 16, respectively. Training starts with a balanced dataset (having equal or similar numbers of annotation counts for each cell) with a fixed learning rate of 0.1 up to 800 epochs, and a trained network from the balanced set is used for the initialization of the training process using step policy learning rate 0.01 with the unbalanced set. Tesla V100 GPU (Graphics Processing Unit) is used for training, and it lasts for approximately 240 h.

### Measurement methods

The method to measure the objective metrics, sensitivity, and specificity depends on updating the probability threshold and Intersection over Union (IOU) parameters, which were used to identify and exhilarate the detected object with the ground truth by the DL algorithm. The ground truth was partitioned into test (%20) and training (%80) annotated sets for the verification and training steps. The sensitivity, specificity, and Cohen's kappa metrics were calculated manually using the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) variables (sensitivity=TP/(TP+FN), specificity=TN/(TN+FP), Cohen’s kappa=[2*(TP*TN–FP*FN)]/[(TP+FP)*(FP+TN)+(TP+FN)*(TN+FN)] [21]. In addition, compatibility between the physician and AI decision makers based on Downey cells presence is determined using Cohen’s kappa analysis. Cohen suggested the Kappa result (calculated by using Jamovi Version 2.3) be interpreted as follows: values ≤0.4 as indicating no agreement and 0.41–0.60 as fair, 0.61–0.80 as moderate, 0.81–1.00 as almost perfect agreement [22].

### Results

The trained network was tested using the randomly selected set as %20 of the ground truth from the annotated dataset. Sensitivity and specificity were measured for each cell type, and they are reported in Table 2. IOU and probability threshold used for the measurements were selected as 0.5, and 0.25, respectively. If the measured IOU observed with the ground truth was greater than the selected value, and their cell types were the same, it was labeled as TP. TP of the negatives associated with each cell type are aggregated to calculate the measured TN for each cell. It was observed from the evaluation of WBCs that 98% sensitivity and 99% specificity is obtained. The validation procedure was also applied on Downey positive and negative samples, which were identified by the physicians through manual evaluation over microscope. The set comprises 11 Downey-positive and 20 Downey-negative PBS samples. These samples were digitized through Mantiscope system and evaluated using the developed method. We observed a sensitivity of 57%, specificity of 100% and a Cohen’s kappa value of 0.5 in the first control. These clinically remarked physical samples are processed through the proposed method, and the measured metrics are presented in Table 3. The second control was about the accuracy of the Downey cell identification through the digitized images. The physicians were asked whether the identified cell was Downey or not. From the identified 83 Downey cells, the physicians confirmed that 55 were true Downey cells and 28 could not be identified as Downey cells due to the pixelization error or wrong classification. The observed accuracy was 66%.

### Discussion

The manual microscopy process consists of classifying blood cells over PBS from human perception. The common classified features of WBC are cytoplasm/nucleus ratio, cytoplasm colors, and segmentation type of the nucleus. While manual microscopy is the tool used to identify these features objectively, digital magnification and pixelization effects have a negative effect over the observation for the identification of the cells from the digitized images. Therefore, the negative set used to measure the specificity was selected from only the WBC types in which the physicians can get confused to annotate.

<table>
<thead>
<tr>
<th>Cell type</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Cohen's kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eosinophil</td>
<td>314</td>
<td>40</td>
<td>74</td>
<td>10,351</td>
<td>81</td>
<td>99.6</td>
<td>89</td>
<td>99</td>
<td>0.84</td>
</tr>
<tr>
<td>Basophil</td>
<td>125</td>
<td>19</td>
<td>19</td>
<td>10,540</td>
<td>87</td>
<td>99.8</td>
<td>87</td>
<td>99</td>
<td>0.87</td>
</tr>
<tr>
<td>Neutrophil</td>
<td>4,644</td>
<td>634</td>
<td>951</td>
<td>6,021</td>
<td>83</td>
<td>90.5</td>
<td>88</td>
<td>86</td>
<td>0.74</td>
</tr>
<tr>
<td>Monocyte</td>
<td>470</td>
<td>80</td>
<td>110</td>
<td>10,195</td>
<td>81</td>
<td>99.2</td>
<td>85</td>
<td>98</td>
<td>0.82</td>
</tr>
<tr>
<td>Lymphocyte</td>
<td>4,981</td>
<td>799</td>
<td>949</td>
<td>5,684</td>
<td>84</td>
<td>87.7</td>
<td>86</td>
<td>85</td>
<td>0.72</td>
</tr>
<tr>
<td>Downey</td>
<td>131</td>
<td>48</td>
<td>3</td>
<td>10,534</td>
<td>98</td>
<td>99.5</td>
<td>73</td>
<td>99</td>
<td>0.83</td>
</tr>
</tbody>
</table>

TP, True Positive; TN, True Negative; FP, False Positive; FN, False Negative; PPV, Positive predictive value; NPV, Negative predictive value.
The Downey cell is a reactive lymphocyte that plays an important role in the immune response. Blood analysis is moving toward automated differentials, which may notice atypical lymphocytes. The presence and number of Downey cells are useful indicators and, in some cases, can be used to help diagnose specific disease states [7].

Today, sophisticated serological and molecular tests for confirming IM have high sensitivities; however, PBS remains an important diagnostic tool. Feder et al. indicated that performing a complete blood count with a manual differential is a rapid test since 1923 that can be helpful for the diagnosis of infectious mononucleosis [6]. Van der Meer et al. stated that recognizing abnormal lymphocytes in PBS can be used to rapidly diagnose various diseases [8].

PBS and bone marrow aspiration (BMA) evaluations are not standardized due to inconsistencies in preparation and physicians’ expertise variabilities. NN has become popular in recent years while solving unstandardized problems with the help of a large variety of data to transfer knowledge. In digital hematology, standardizing the PBS and BMA evaluation tests using DL-based techniques can be initiated [23]. These studies observed sensitivity and specificity metrics greater than 90 percent via measurements with the selected test dataset. While these studies only measure the classification probability for the individual blood cells, the objective measurements that cover the whole slide are required to define the percentages for each blood cell type for a detailed PBS examination. Blood cell-specific diseases for pre-diagnostics are also studied using DL techniques, and they are intended for automatically analyzing the samples for lymphoblastic leukemia [24–26]. By observing sensitivity and specificity metrics, the authors claim that this NN would identify risky and non-risky patients in values greater than 90 percent. Another DL approach evaluates BMA samples and discloses 95 percent sensitivity by providing an analysis tool [27]. Blood cell segmentation is also imperative for the detection networks to create a highly sensitive algorithm while it scores each segmented cell to be associated with a cell group. A segmentation algorithm before the NN is also proposed in Celebi et al.’s study that observes an accuracy greater than 90 % [28].

In the literature, studies mostly concentrate on the WBC morphologies due to easy detection and segmentation from the blood film images with their color, nucleus, and cytoplasm. With the conventional standard for examining PBS samples, it is also required to analyze RBC and platelets. Some types of RBC, such as sickle cells and target cells, also have the exact features that would be differentiated from other cell types. DL-based techniques concentrating on the sickle cells also reveal an accuracy greater than 90 percent for the detection [29, 30]. Annotation-based digital evaluation of DL algorithms would disclose a percentage in sensitivity and specificity by comparing the AI predictions with the physicians’ annotations. On the other hand, morphology-based diagnostics for the blood smear samples by manual microscopy vary among observers, resulting in unpredictable intra- and interobserver variability. The standard evaluation steps comprise counting the blood cells in an evaluation window with different morphological features and estimating each blood cell type percentage. While it is a manual process, counting the cells accurately using machine-learning techniques in more regions would improve the results by increasing the number of images collected through the samples. The indicators of infectious diseases would result in a small number of specific cells in the PBS evaluation. Therefore, analyzing a wide region of physicians through the microscope greatly affects accuracy. There are automatic slide scanner systems, i.e., Mantiscope and Sysmex XN series, for scanning the PBS slides [13, 14].

In this study, Downey cells were differentiated from other cells with a sensitivity and specificity of 98 and 99 %, respectively. Having an NPV of 99 %, false Downey cell diagnosis was minimized. A value of 0.83 was reached for the Cohen’s kappa analysis performed to evaluate the reconciliation between the physicians and AI decision-makers in terms of Downey cell presence. The Kappa value between 0.81 and 1.0 was accepted as an indicator of high compliance. It should also be noted that the physical samples categorized as Downey positive and negative could also be identified with 57 % sensitivity and 100 % specificity.

It should also be emphasized that this study’s lack of external validation is a limitation. Additionally, the ML model sensitivity and specificity for Downey positive and negative samples were lower than the performance evaluation for WBCs. This may be because the monolayer areas

<table>
<thead>
<tr>
<th>Table 3: Validation of the system due to physicians’ evaluation.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Downey (physicians)</td>
</tr>
<tr>
<td>Normal (physicians)</td>
</tr>
</tbody>
</table>

AI: Artificial Intelligence.
were not set exactly by the slide scanner during the digitalization of physical samples. At the same time, the examination of AI cannot associate any cell object to Downey from these digitized images. On the other hand, the physician can examine any pieces of the PBS slide over the manual microscope and can identify the Downey cells placed in any part of the PBS slide. Unfortunately, the imbalanced data and low sample size were the additional limitations of the study.

As a result of these evaluations, a machine learning algorithm was created, and the algorithm can be applied to digitized images to speed up the actual diagnosis. It can provide a helpful preliminary diagnosis to the physician. The accurate detection of Downey cells with AI might warn the physician to prevent missing these cells during the PBS evaluation.

Conclusions

This algorithm for identifying Downey cells to facilitate the accurate diagnosis can be applied over the digitized images collected from the slide scanners and can provide a helpful preliminary diagnosis to the physician. The verification and validation phases are performed to show the effectiveness of this process.

Artificial intelligence-based systems in hematology will help the physicians diagnose and treat diseases and their success rates will increase day by day. While the annotation and AI development process continues to collect more patient data, we would obtain higher sensitivity and specificity for the associated Downey cells. New training techniques and DL structures will also be discussed to get more accurate classification results in future work.

Acknowledgments: None.

Research ethics: Ethics committee approval was received for this study from XXXX Clinical Research Ethics Committee (Ethics Committee Decision No. 095) on 16th March, 2022.

Informed consent: None.

Author contributions: Ardicoglu Akisin Y and Coteli MB contributed to the conception and design of the study; Ardicoglu Akisin Y and Akar N interpreted the results; Ardicoglu Akisin Y and Coteli MB performed the statistical analysis; Ardicoglu Akisin Y and Coteli MB drafted the manuscript; and Akar N supervised the study. All authors read and approved the final manuscript.

Competing interests: Authors state no conflict of interest.

Research funding: None.

Data availability: All data supporting the findings of this study are available within the paper.

References


