This is the pre-peer reviewed version of the article "Differential diagnosis of hereditary anemias from a fraction of blood drop by digital holography and hierarchical machine learning" which has been published in the final form in ***Biosensors and Bioelectronics, 201, 113945, https://doi.org/10.1016/j.bios.2021.113945***.

Differential diagnosis of hereditary anemias from a fraction of blood drop by digital holography and hierarchical machine learning

Pasquale Memmolo,[1] Genny Aprea,[1] Vittorio Bianco,[1,*] Roberta Russo,[2,3] Immacolata Andolfo,[2,3] Martina Mugnano,[1] Francesco Merola,[1] Lisa Miccio,[1] Achille Iolascon[2,3] and Pietro Ferraro[1]

[1] *Istituto di Scienze Applicate e Sistemi Intelligenti "Eduardo Caianiello" (ISASI-CNR), via Campi Flegrei 34, 80078 Pozzuoli, Napoli, Italy.*

[2] *Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università Federico II di Napoli, Italy*

[3] *CEINGE—Biotecnologie Avanzate, Napoli, Italy*

[*] *corresponding author: vittorio.bianco@isasi.cnr.it*

Abstract

Anemia affects about the 25% of the global population and can provoke severe diseases, ranging from weakness and dizziness to pregnancy problems, arrhythmias and hearth failures. About 10% of the patients are affected by rare anemias of which 80% are hereditary. Early differential diagnosis of anemia enables prescribing patients a proper treatment and diet, which is effective to mitigate the associated symptoms. Nevertheless, the differential diagnosis of these conditions is often difficult due to shared and overlapping phenotypes. Indeed, the complete blood count and unaided peripheral blood smear observation cannot always provide a reliable differential diagnosis, so that biomedical assays and genetic tests are needed. These procedures are not error-free, require skilled personnel, and severely impact the financial resources of national health systems. Here we show a differential screening system for hereditary anemias that relies on holographic imaging and artificial intelligence. Label-free holographic imaging is aided by a hierarchical machine learning decider that works even in the presence of a very limited dataset but is enough accurate for discerning between different anemia classes with minimal morphological dissimilarities. It is worth to notice that only a few tens of cells from each patient are sufficient to obtain a correct diagnosis, with the advantage of significantly limiting the volume of blood drawn. This work paves the way to a wider use of home screening systems for point of care blood testing and telemedicine with lab-on-chip platforms.

This is the pre-peer reviewed version of the article "Differential diagnosis of hereditary anemias from a fraction of blood drop by digital holography and hierarchical machine learning" which has been published in the final form in ***Biosensors and Bioelectronics, 201, 113945, https://doi.org/10.1016/j.bios.2021.113945***.

Keywords: hereditary anemias; digital holography; machine learning; blood cells; point-of care diagnostics;

Introduction

Red Blood Cells (RBCs) in healthy state are rich in hemoglobin, a large iron content protein that allows RBCs delivering oxygen throughout human body and expelling carbon dioxide by the lungs (Elgsaeter et al., 1986). The term “anemia” refers to any condition that impairs the transfer of an adequate amount of oxygen to the tissues (Andolfo et al., 2016, 2018a, 2018b; Cappellini et al., 2020; Donker et al., 2016; Elgsaeter et al., 1986; Iolascon et al., 2019, 2020, 2021; King et al., 2015; Muckenthaler et al., 2017; Russo et al., 2014, 2020). This can be due to several concurring factors mostly associated with a shortage of healthy RBCs, e.g. iron-deficiency, B-12 vitamin deficiency, infections, serious chronic inflammatory diseases, cancer-like disorders such as leukemia, bleeding, among the others. Besides, hereditary anemias can cause either reduced half-life or impaired production of erythrocytes (Andolfo et al., 2016, 2018a, 2018b; Donker et al., 2016; Iolascon et al., 2019, 2020, 2021; King et al., 2015; Russo et al., 2014, 2020). The World Health Organization (WHO) estimated that anemia affects globally 1.62 billion people, i.e. the 24.8% of the entire world population, the 47.4% of them being preschool-age children (de Benoist et al., 2008). Of note, about 10% of these patients are affected by rare anemias of which 80% are hereditary. In sickle cell anemias, low oxygen tension promotes RBCs sickling and thus their ability to move along the blood vessels gets impaired. These abnormal cells can obstruct the vessel and tend to die early. Iron-Refractory Iron-Deficiency Anemia (IRIDA), Thalassemias, Hereditary Spherocytosis (HS), Congenital Dyserythropoietic Anemias (CDAs), Dehydrated Hereditary Stomatocytosis (DHS) are among the most spread and well-characterized forms of hereditary anemias (HA) (Andolfo et al., 2018a, 2018b; Donker et al., 2016; Iolascon et al., 2020; King et al., 2008; Russo et al., 2014). Anemia can provoke weakness and severe fatigue, abnormal swelling, dizziness, pale skin, headaches and chest pain, leg cramps, shortness of breath, cardiac arrhythmias and hearth failures, premature births in anemia-affected pregnant women, and even death in the case of acute bleeding (King et al., 2008; Qaseem et al., 2013; Russo et al., 2014). Additionally, one of the most harmful complication is the chronic iron overload that is often associated either to anemia with ineffective erythropoiesis, such as thalassemias and CDAs, or to HS. Regular blood transfusions may be needed in the case of thalassemia-affected patients as well as in approximately 20% of patients with CDA type II. Splenectomy is the most effective surgical treatment for HS, although it might not be necessary for

a patient with CDA II, or even worst, it can be contraindicated in DHS because of the risk of severe thrombotic events. Incorrect diagnosis can critically impact on the follow-up and therapy of these patients. Additionally, if proper medical treatments are followed, most of the anemia-related symptoms/complications could be avoided or at least mitigated (Qaseem et al., 2013). Thus, early detection of anemia is very vital for properly treating the patients and to associate the required iron-based and/or vitamin-based diet. Even more important is to achieve differential diagnosis through the development of effective modalities to set specific cures well-tailored to the anemia type each patient is affected by. In fact, although the diagnostic workflow for HA is part of normal clinical practice (Cappellini, M.D. et al., 2020; Iolascon et al., 2019, 2020), it is often very difficult to obtain differential diagnosis and classification as there is a wide range of unspecific phenotypes and overlapping phenotypes in patients with different genetic backgrounds (Russo et al., 2020). Therefore, achieving rapid, automatic and objective differential screening of RBCs from blood sample is strongly required to reduce analysis time and costs, and to minimize test errors. The use of very small sample volumes would also result in minimally invasive tests for patients. Single-cell analysis approaches have been proposed to automate the screening based on specific morphological features extracted from microscopy images. However, differential diagnosis requires more information, other than the mere morphology, to distinguish between different anemia types that share similar morphological traits (Akrimi et al., 2013; Hoppe, 2013; Khan et al., 2019). Conversely, label-free quantitative phase imaging methods extract, from the phase-contrast maps of RBCs, clinically relevant parameters that are related to their hemoglobin content and, once added to morphological information, can help distinguishing between healthy and sick RBCs (Ahmadzadeh et al., 2017; Jaferzadeh and Moon, 2016; Jaferzadeh et al., 2019; Memmolo et al., 2014; Mir et al., 2011; Moon et al., 2012; Mugnano et al., 2018; Park et al., 2010; Pham et al., 2013; Rappaz et al., 2009; Seo et al., 2010; Shaked et al., 2011; Yi et al., 2015; Zhu et al., 2013). The dynamic evaluation of the phase-contrast changes over time has been successfully used to map RBCs' membrane fluctuations, which are related to cells healthiness and ageing states (Jaferzadeh et al., 2019; Rappaz et al., 2009) or even to detect Covid-19 infection (O'Connor et al., 2021). In particular, Digital Holography (DH) in transmission microscopy configuration is one of the best candidates to provide label-free quantitative phase-contrast information from samples in Petri dishes or flowing along microfluidic channels in Lab-on-a -Chip (LoC) environment (Ahmadzadeh et al., 2017; Jaferzadeh and Moon, 2016; Jaferzadeh et al., 2019; Memmolo et al., 2014; Moon et al., 2012; Mugnano et al., 2018; Rappaz et al., 2009; Seo et al., 2010; Yi et al., 2015; Zhu et al., 2013). DH offers flexible

numerical refocusing of cells imaged outside of their focus plane, which can be very beneficial in microfluidics where cells flow within the chip with high throughput. In addition, portable DH microscopes have been developed to bring diagnostics to the patient's home, thus pushing the rapidly growing field of telemedicine (Bianco et al., 2017; Cacace et al., 2020; O'Connor et al., 2019; Patel et al., 2018; Seo et al., 2010; Zhu et al., 2013) that allows for low-cost diagnosis. DH allows enlarging the set of features that can be extracted from a RBC, adding 3D parameters to the conventional 2D descriptors (Ahmadzadeh et al., 2017; Jaferzadeh and Moon, 2016; Moon et al., 2012; Moon and Lee, 2015). However, morphometric features or their combination might be not distinctive enough for a differential diagnosis since these parameters can be dependent each other (Hejna et al., 2017; Moon et al., 2012). A more robust decision pipeline needs to be adopted to reduce the False Positive (FP) and False Negative (FN) rates that generate in the case of similar morphologies. Deep learning architectures have been proposed to face the problem. In fact, a neural network can learn distinctive attributes directly from the image appearance, being practically able to separate "similar" populations (Chen et al., 2016; Jo et al., 2019; Miccio et al., 2020; Mitani et al., 2020; Moen et al., 2019; Moon et al., 2013; Singla and Srivastava, 2020; Xu et al., 2017; Zhang et al., 2018). However, deep learning requires a huge amount of training data, and is not robust to misalignments or changes occurring in the recording system configuration or its key optical parameters. Recently, an all-optical biomarker has been proposed for label-free detection and differential classification of anemias. RBCs are modeled as biological lenses, and the corresponding wavefronts at the exit pupil described by Zernike polynomials are thus used as a classification basis (Mugnano et al., 2018). This approach improves the accuracy in separating sick cells from the healthy population, and can classify IRIDA from Thalassemia with residual overlap up to 38%. However, simultaneous differentiation between all the main inherited anemia classes was not feasible relying on the sole single-cell analysis (Mugnano et al., 2018).

Here we propose a holographic imaging differential screening of anemias-affected patients based on artificial intelligence. We demonstrate that by using a hierarchical machine learning decider and in the presence of a very limited dataset it is possible to accomplish a comprehensive clinical screening. In fact, the hierarchical decision pipeline proposed here is capable to significantly improve the performance of a differential classifier joining DH imaging with machine learning. We considered healthy RBCs and five classes of inherited anemias, namely IRIDA, α-thalassemia, HS, CDAs, and DHS1. We defined a set of 16 features well-tailored to separate different anemia classes and with minimal correlation between them (Bianco et al., 2020; Hejna et al., 2017). Then, we use

these features for training a cascade of two different classifiers following a hierarchical approach. The former is called to assess whether the sample belongs to the healthy or disease populations (binary prescreening stage). Once cells are found to belong to the disease class, a second classifier is employed to associate each single cell to an anemia phenotype (i.e. differential stage). We show that the best classifier for both stages is independently determined during training among 23 machine learning models, i.e. the two stages are flexibly decoupled. This approach is capable to distinguish each population with its own associated accuracy, FP and FN rates. As expected, for some types of anemia, the FP and FN rates are not low enough for reliable screening of populations in the case of single cell classification. Therefore, we correlate the decisions of the RBCs extracted by the same patient by applying a max-voting criterion (Kleiber et al., 2020; Natarajan and D. Sarkar, 2019; Singh, 2018). This has been found to hone the decision maker consistently and very effectively, so that even the low-performance differential classification based on single-cell analysis results in a reliable diagnostic response. Remarkably, we prove that only a few tens of cells from each patient are needed for achieving a correct diagnosis, and the training stage required a small dataset as well.

2. Materials and methods

2.1 Patients and preparation of red blood cell samples

Overall 8 patients with clinical and molecular diagnosis of CDAI, CDAII, HS, DHS1, IRIDA, and α-thalassemia were included in the study. Seven healthy donors were also included in the study. The university Ethical Committees approved the collection of the patient data from the Medical Genetics Ambulatory in Naples (DAIMedLab, “Federico II” University). Fresh blood samples from the patients and the healthy subjects were obtained after signed informed consent, and according to the Declaration of Helsinki. Sample preparation cells was performed as previously described. Briefly, 3 ml EDTA fresh blood sample was collected from each subject. Blood was centrifuged at 4°C at 2500 rpm for 15 min in order to separate RBCs, precipitated at the bottom of the sterile centrifugation tube, from the plasma and buffy coat. After centrifugation, the plasma and buffy coat were discarded and the RBC pellet was washed with same volume of a saline solution of 0.90% (w/v) sodium chloride (NaCl) in sterile water, and recentrifuged at 4°C at 2500 rpm for 15 min. After the second centrifugation, the supernatant fraction was removed, and an aliquot of isolated RBCs (<100 µL) was diluted in 10 mL of the solution of 0.90% (w/v) NaCl in sterile water, with a final osmolarity

of 308 mOsm/L in order to maintain the osmotic pressure of the RBCs' membranes. For the experiments, a final volume of diluted RBCs (<100 µL) was used.

2.2 Digital holographic recording and processing

The imaging system is a coherent-light Digital Holographic microscope exploiting Mach-Zehnder interferometer geometry in transmission mode. Cell population is placed on petri dish and illuminated by laser light at wavelength of 532nm and 3mW power (Torus - Laser Quantum). Light passing through the sample (object beam) is collected by a microscope objective (Olympus - Water immersion - 60X magnification and NA=1.2) and imaged on a 2048 × 2048 CMOS camera (USB 3.0 U-eye, from IDS). In order to obtain DH off-axis geometry, a second beam (reference beam) from the same light source is let interfere with the object one by means of a Beam Splitter (BS) that allows the fine tuning of the off-axis angle (see Fig. 1). For each donor, up to ten independent digital holograms of RBCs are recorded with the aim to collect only hundreds of RBCs for each pathological condition. The latter will be used to create the training and the validation sets for the machine learning model that will be employed for anemia diagnosis purpose.

Once an off-axis hologram is recorded, the twin image and the zero-th order of diffraction terms are filtered out in the Fourier spectrum obtaining the complex object wavefront. If the sample is imaged out of focus, numerical propagation (Bianco et al., 2017; Cacace et al., 2020; Memmolo et al., 2014; Mugnano et al., 2018; Shaked et al., 2011) is performed to provide its refocused complex amplitude, $C = P_{z_f}\{H\}$, where $P_z\{\dots\}$ is the propagation operator, and z_f is the best focus distance. From C, we obtain for each object its amplitude image, $|C|$, its phase-contrast signature, ψ, along with an estimate of the object support, S. Such information is used to extract a set of features allowing us to characterize each single cell in the first classification stages and to formulate a diagnosis related to the patient in the final stage. To this aim, we adopt a hierarchical max-voting approach, as sketched in Fig. 1. For the sake of clarity, Fig.1 reports a complete workflow referring to a diseased blood sample from IRIDA-affected patient. Let L be the number of RBCs belonging to the same subject and imaged through the DH microscope. The features extractor receives the couple $\{\psi, S\}$ for each of the L objects and calculates the 16-features vector, F.

This is the pre-peer reviewed version of the article "Differential diagnosis of hereditary anemias from a fraction of blood drop by digital holography and hierarchical machine learning" which has been published in the final form in ***Biosensors and Bioelectronics, 201, 113945, https://doi.org/10.1016/j.bios.2021.113945****.*

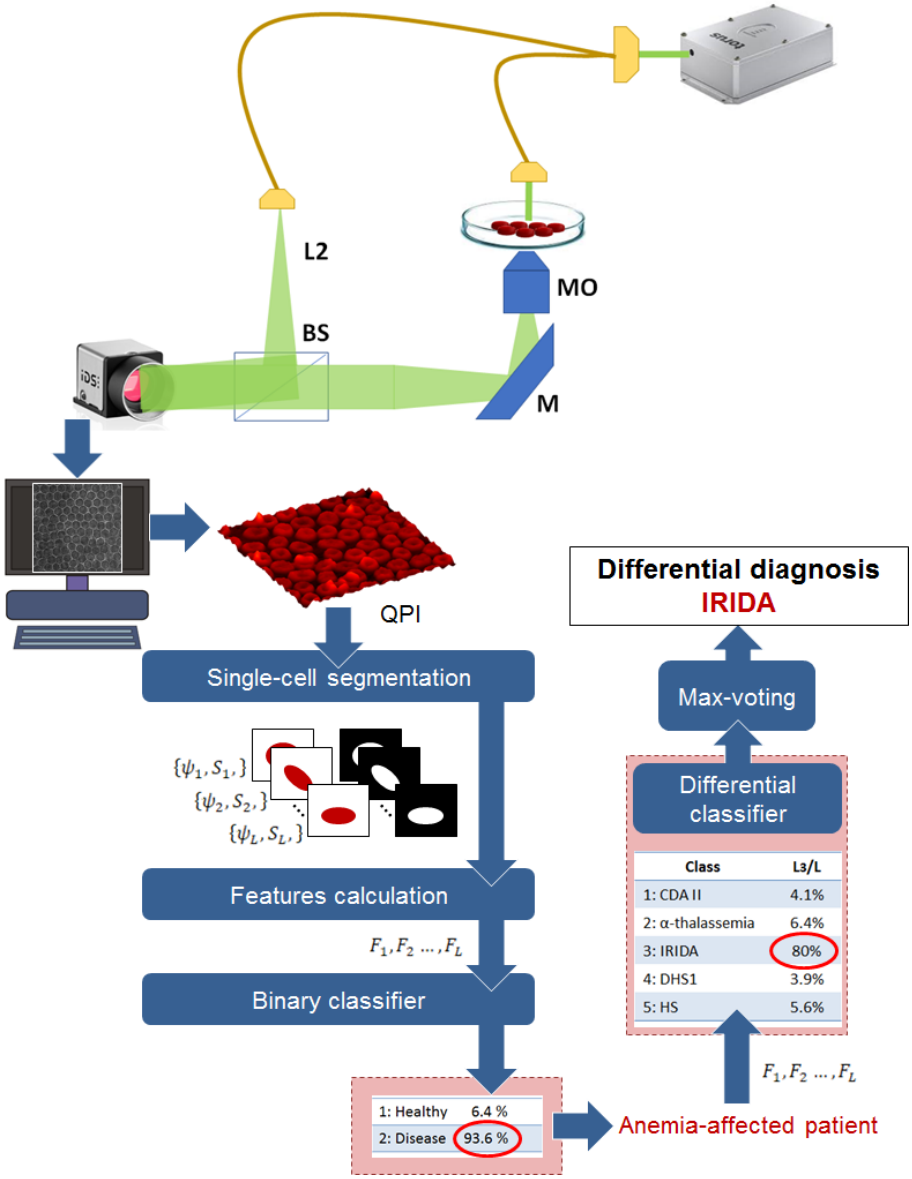

Fig. 1. Experimental setup for holographic recording and processing pipeline implementing the proposed hierarchical max-voting decider.

First, we divide the RBCs populations into healthy and sick cells. A binary classifier is employed to discern between these two classes, based on the set of features F measured from each sample (binary stage). Once one RBC is classified as “Healthy” or “Disease”, then the next RBC from the same patient’s blood drop is examined until all the L RBCs are classified. Then, a first decision is taken, according to a binary condition, detailed in the following. At this stage, the patient can be judged to be “Healthy” or “Anemia-affected”. In the latter case, the corresponding features vectors, $F_1, \ldots, F_L$ are received by a pre-trained differential classifier. This is committed to associate each RBC to one of the five inherited anemias that we have considered (differential stage). Differential classification is repeated for all the L RBCs belonging to the same subject/sample. The output is a table linking each type of anemia to the percentage of associated RBCs. The binary and the

differential classifiers are individually pre-trained, and the Matlab® Classification Learner (CL) tool is used in both cases to determine the best classifier-kernel configuration among 23 effective options available in the toolbox. This choice ensures high flexibility in the selection of the optimal configuration for both the specific classification tasks. The detailed description of the image features set is reported in the supplementary file. It is worth pointing out that, to achieve the final pathological condition decision, only few tens of RBCs (i.e. *L*) belonging to the same digital hologram need to analyzed within the hierarchical max-voting decider. These represent a very small fraction of the blood drop prepared for the experiments. In other words, after validating the proposed model, the differential diagnosis of hereditary anemia can be achieved by using one single digital hologram.

3. Results and discussion

3.1 Dataset analysis

The proposed hierarchical scheme is based on a cascade of two classifiers, i.e. a binary one devoted to establish if a sick condition occurs, and a multi-class decider used to identify the specific anemia. Therefore, two different datasets are created for the training and the validation of candidate classifiers. The first one, used to screen samples healthiness, consists of two classes of 1000 elements, labeled as “Healthy” and “Disease” to indicate if a normal or pathological condition occurs. The “Healthy” elements are obtained by randomly selecting 200 RBCs from 5 healthy donors and calculating the corresponding features. The “Disease” elements are obtained by selecting 200 RBCs from each pathological condition and using images of the samples coming from 1 donor per condition. In the case of CDA, we use the RBCs of the CDA II condition only. We refer to this dataset as the Binary dataset. Instead, the second classifier is employed if and only if the first one provides the “Disease” alert, so that the specific pathological condition has to be identified among the aforementioned 5 anemias. In this case, the training and validation dataset is composed of 1000 RBCs per anemia’s class, collected from the same donors used for the healthiness screening. We refer to this as the Differential dataset. The remaining RBCs samples of 2 healthy, 1 IRIDA, 1 CDA I and 1 α-thalassemia donors are used only in test, along with other RBCs of DHS1 and HS coming from the same donors used in the training and validation stages. In order to provide a visualization of the complexity of the addressed classification problem in terms of objects similarity, we display the t-distributed Stochastic Neighbor Embedding (t-SNE) maps for

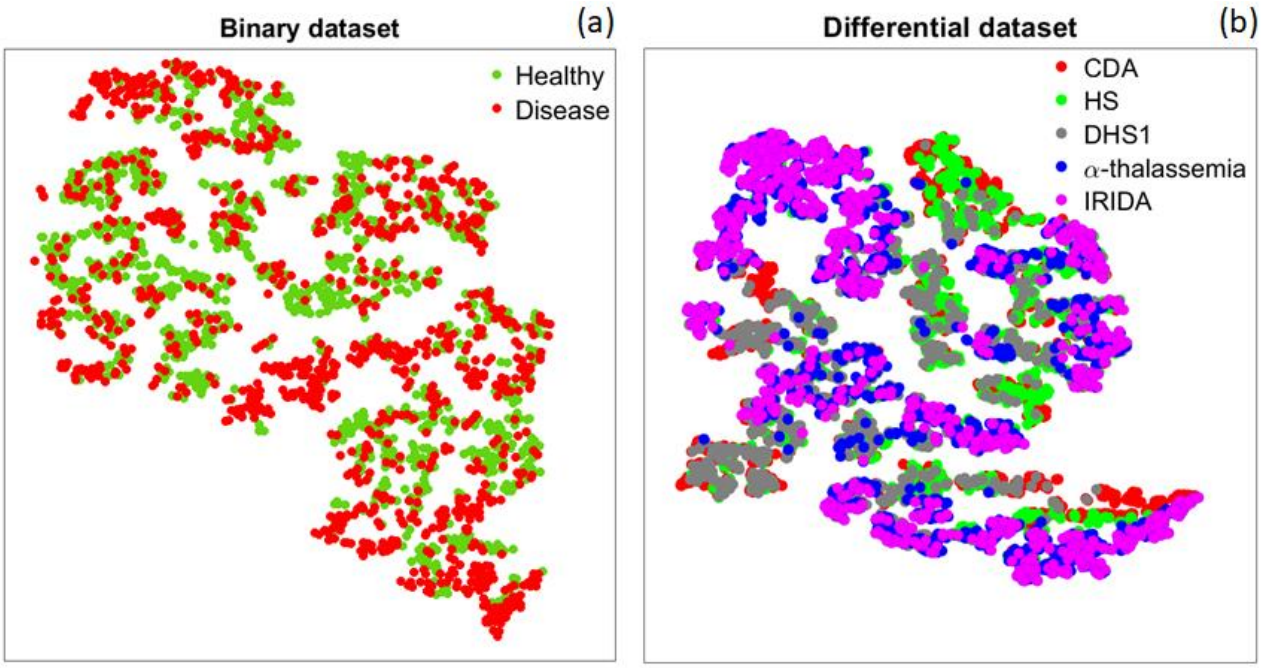

Fig.2. t-SNE visualization of training datasets for the (a) binary and (b) differential classifiers.

both the binary and differential datasets in Figs.2(a,b) respectively.

3.2 Identification of pathological conditions

The RBCs populations we considered are very similar each other in terms of morphometric traits, thus making very challenging the identification of a suitable classifier to perform the RBCs discrimination at the single cell level, as shown in Fig.2. Intuitively, it is not correct to tackle this classification problem on a single-cell basis, since each blood sample is representative of a unique pathological condition. Starting from this consideration, we design both binary and differential classifiers by using 1000 different RBCs belonging to the same sample, with a max-voting based decider that infers the classification output. In this way, anemia identification can be obtained by using few RBCs imaged in a single holographic snapshot, thus allowing the differential diagnosis from just a small portion of blood drop. We employ the Classification Learner toolbox of MATLAB® to compare the performance of 23 classifiers, trained with both binary and differential datasets. Among them, the best ones are selected in terms of overall accuracy with respect to the single-cell classification task. Fig. 3(a,b) report this analysis in the case of the binary classifier. The highest accuracy (84.3%) is reached by the cubic Support Vector Machine (SVM) as shown in Fig.3(a) and the corresponding confusion matrix obtained in validation is reported in Fig.3(b). Moreover, the two columns on the far right in (b) report the percentages of all the cases predicted to belong to each class that are correctly and incorrectly classified, respectively. These metrics are often called

This is the pre-peer reviewed version of the article "Differential diagnosis of hereditary anemias from a fraction of blood drop by digital holography and hierarchical machine learning" which has been published in the final form in ***Biosensors and Bioelectronics, 201, 113945, https://doi.org/10.1016/j.bios.2021.113945****.*

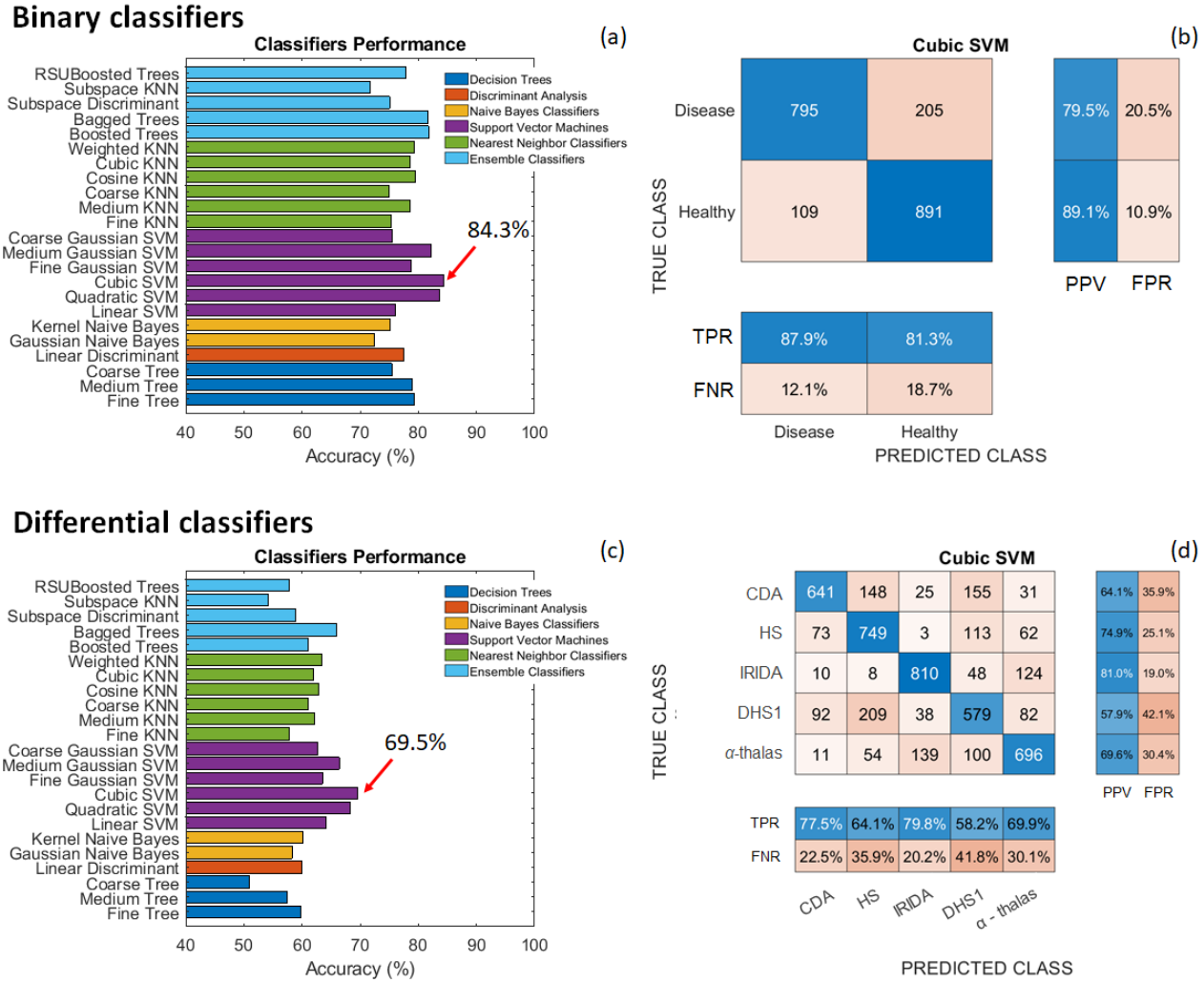

Fig.3. Performance of classifiers within the hierarchical scheme. (a,c) report the overall accuracies in validation obtained for the 23 classifiers trained with the binary and differential datasets, respectively. The cubic-SVM outperforms all the others in both cases, with accuracy of 84.3% for the binary dataset and 69.5% for the differential one with 5 classes corresponding to pathological conditions. (b,d) show the confusion matrices obtained in validation by using the cubic-SVM classifier for the binary and differential cases, respectively.

the Positive Predictive Value (PPV) or precision, and False Positive Rate (FPR), respectively. The rows at the bottom of Fig. 3(b) show the percentages of all the cases belonging to each class that are correctly and incorrectly identified, respectively. These metrics are often called the True Positive Rate (TPR) or recall, and False Negative Rate (FNR), respectively.

The same analysis is repeated for the differential classification with the aim to identify the specific anemia condition among the 5 diseases described above, i.e. CDA, HS, IRIDA, DHS1 and α–thalassemia. Also in this case, the cubic SVM provides the highest overall accuracy (69.5%), Fig. 3(c). The confusion matrix along with PPV, FPR, TPR and FNR are reported in Fig.3(d). In summary,

the proposed hierarchical decider consists of the cascade of two cubic SVM classifiers in which the second one is employed only if the first decision is for a diseased sample, as highlighted in Fig.2. The max-voting criterion is used to infer the output of the hierarchical decider when multiple RBCs belonging to the same sample are tested. Hence, the final output of the proposed scheme is the patient's condition, i.e. healthy/disease in the first stage and, if the disease status is revealed, the type of anemia. The hierarchical decider is tested by using 10 blood samples collected from 2 healthy donors and 8 donors having diagnosed anemia. For each sample, we acquire just one digital hologram and, after the process of holographic phase reconstruction, tens of RBCs are used for the condition identification.

Table 1 reports the results of the hierarchical classification to screen the donors' pathological conditions. In all cases, the binary decision is found to be correct, even with RBCs sample collected from different donors (e.g. IRIDA donor 2 and α-thalassemia donor 2) and with different phenotypes of the same anemia condition (CDA I when the classifier is trained with samples of CDA II), as highlighted in Table 1 with light blue color. Of course, tests are performed with RBCs never "seen" by the classifier even if collected from the same donors used in the training and validation step. Then, the 8 test cases of pathological conditions revealed by the first healthiness screening become the input of the second classifier called to perform the differential diagnosis. In the latter case, the max-voting based decision permits to correctly infer the donor's condition from the analysis of multiple RBCs belonging to the same sample, thus largely improving the

Table 1. testing results for the hierarchical classifier

	BINARY		DIFFERENTIAL	
True condition	Accuracy (%)	Healthiness	Accuracy (%)	Max-voting
Healthy (donor 6)	89.3	Healthy	---	---
Healthy (donor 7)	93.1	Healthy	---	---
CDA II	81.0	Disease	51.2	**CDA**
IRIDA (donor 1)	88.0	Disease	80.4	**IRIDA**
HS	58.7	Disease	58.5	**HS**
DHS1	84.9	Disease	71.4	**DHS1**
α – thalassemia (donor 1)	73.8	Disease	76.0	**α – thalassemia**
CDA I	51.2	Disease	50.0	**CDA**
IRIDA (donor 2)	91.4	Disease	84.3	**IRIDA**
α – thalassemia (donor 2)	83.7	Disease	69.0	**α – thalassemia**

diagnostic result with respect to the single-cell based classification approaches. More details about the test cases analysis are reported in the supporting information.

3.3 Integrated holographic diagnostic workflow of hereditary anemias

In order to show how the diagnosis could be simplified by our method we discuss in this section the current diagnosis protocol versus an alternative workflow that can benefit from the use of DH technology. The conventional workflow for diagnosis of HAs starts as first line of investigation with positive personal and familial history, complete blood count (CBC), and peripheral blood smear observation (PBSO), as illustrated in Figure 4, left column. These methods are effective and high-throughput, but they require skilled and trained personal. Particularly, this is an issue for PBSO. Indeed, although automated image-based recognition of RBCs, white blood cells and platelets has been proposed (Briggs et al., 2009; Ceelie et al., 2007; Da Costa, 2015), PBSO under the optical microscope is not automated for sick cell screening and subject to human errors since its performance strictly depend on the training and experience gained by the pathologists and hematologists. In view of the large number of worldwide cases, blood screening for anemia through CBC is a common practice by most of the population of the richest and middle-income Countries, which severely impacts the financial resources of National Health Systems (NHS) (Smith, 2010; Spence, 2010). Moreover, CBC and PBSO cannot automatically discern between different types of anemia, so biochemical assays are required. Indeed, the diagnostic workflow for HA requires specialized biochemical tests as EMA test, SDS-PAGE analysis and ektacytometry, and eventually bone-marrow aspirate, as second line of investigation (Cappellini, M.D. et al., 2020; Iolascon et al., 2019, 2020). However, biochemical tests are not free from errors and misdiagnosis can occur, e.g., between α-thalassemia or HS and CDA II, which might cause unnecessary surgery treatments. Finally, genetic testing serves as the confirmatory test. Currently, genetic testing is used early in the diagnostic workflow of HA, which removes the need for some of the specialized tests, especially when the clinical data for the patients are not informative, or when the patient is transfusion dependent (Russo et al., 2020). Nevertheless, genetic testing is not always conclusive; indeed, it has been estimated that the current diagnostic yields range between 38–87%, which depends on how many and what types of genes are included in the genetic analysis, and on the depth of the phenotypic assessment (Russo et al., 2018, 2020). In the right column in Figure 4 is illustrated the workflow that can be implemented by using the intelligent holographic imaging-based differential screening approach.

This is the pre-peer reviewed version of the article "Differential diagnosis of hereditary anemias from a fraction of blood drop by digital holography and hierarchical machine learning" which has been published in the final form in ***Biosensors and Bioelectronics, 201, 113945, https://doi.org/10.1016/j.bios.2021.113945****.*

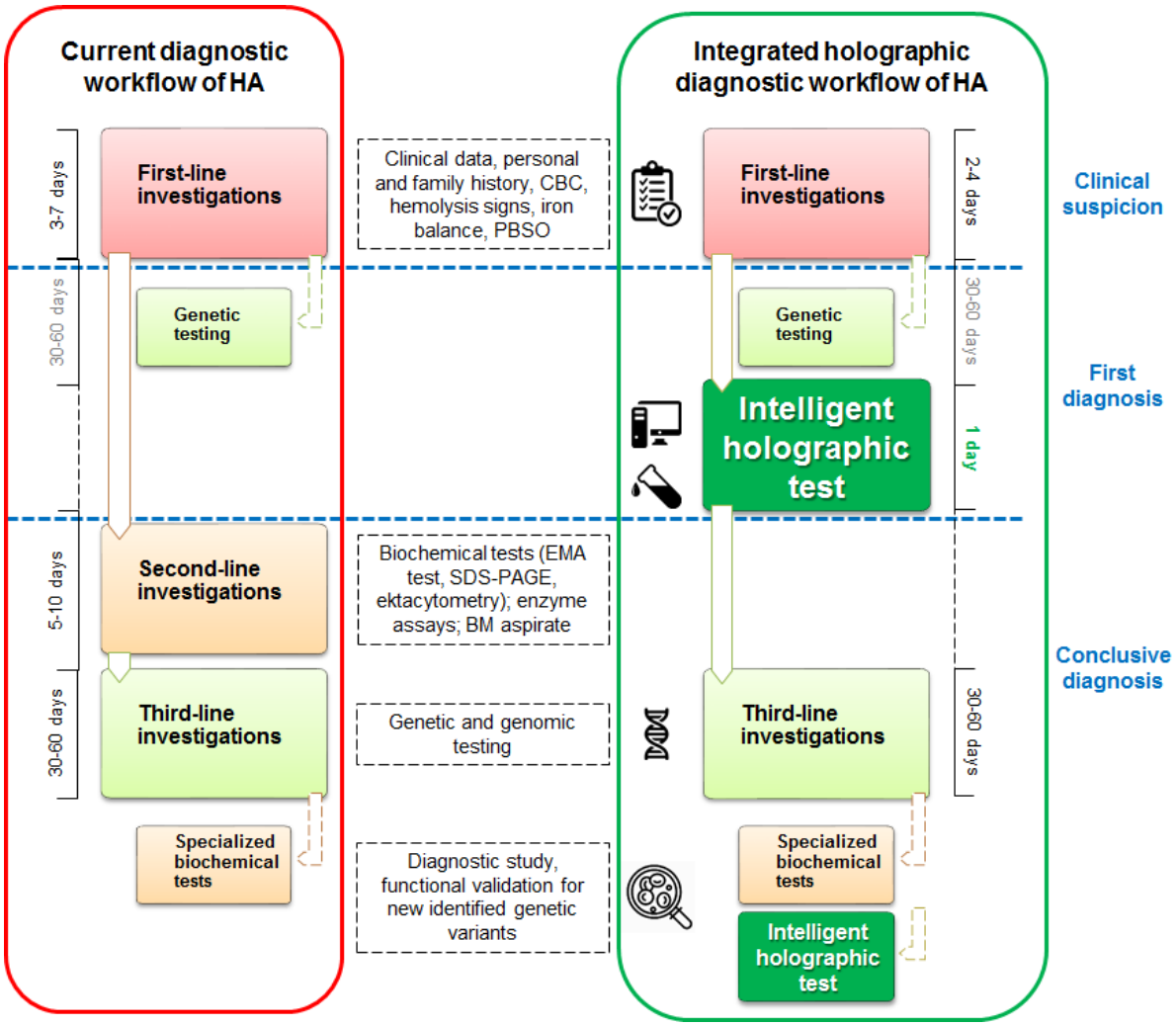

Fig. 4. Diagnostic workflow of HA. On the left, the current diagnostic workflow is shown. This conventional approach takes a turnaround time that ranges from 38 to 77 working days. On the right, the new proposed workflow. The new approach can be integrated into the routine pipeline as a complementary specialized test. The integrated workflow will take a turnaround time that ranges from 33 to 67 working days. Of note, using this approach, a first diagnosis could be achieved with a turnaround time of 2-4 days by avoiding the PBSO. For both approaches, early genetic testing in the diagnostic workflow has been considered for peculiar types of patients, such as those with severe transfusion-dependent anemia. The dashed arrows illustrate the phase of the diagnostic study that is needed to functionally validate and assess the pathogenicity of the newly identified genetic variants.

Our method could be usefully integrated in the conventional diagnostic workflow as a specialized test to improve or to fully replace in the next future the second-line investigation step, as shown in Figure 4. In particular, the first line investigation can be speeded up (from 3-7 days to 2-4 days) since, by using the proposed approach, the PBSO test can be skipped. It is worth noticing that the max-voting results for the CDA condition, shown in the last column of Table 1, does not discriminate between different phenotypes (e.g. CDA I vs. CDA II). This further differential screening could be

achieved by applying the same intelligent holographic testing framework but training the model with the proper phenotypes as distinct classes. Hence, to highlight this further possibility, we included a second Intelligent holographic test in the third line-investigation of Fig. 4.

To the best of our knowledge, a similar work based on a learning approach to identify hematologic disorders has been recently proposed (Kim et al., 2019). In that study, the profiling of individual RBCs obtained from 3D refractive index tomograms and time-lapse phase maps, has been used for training a suitable neural network from input features consisting of six properties of the RBC. Then, during the test stage, the prediction of pathophysiological conditions is obtained by averaging the prediction of profiling data for multiple RBCs. Three remarkable differences can be identified between this deep learning based approach (Kim et al., 2019) and the one we propose, i.e. (i) features are extracted from 3D refractive index tomograms and time-lapse phase maps of each RBC, while we use only one single phase image containing multiple RBCs, which makes the setup simpler and more prone to field portable lab-on-a-chip implementations for home screenings and telemedicine; (ii) we design a hierarchical decision architecture, by first determining if a disease condition occurs, then making a disorder diagnosis. We demonstrate that this choice improves the classification accuracy with respect to a classifier trained with both healthy and sick cells, as in (Kim et al., 2019); (iii) we employ the max-voting decision stage instead of averaging output predictions. This choice is related to the possibility to identify the anemia disease by using RBCs belonging to only one single phase image. In fact, our results demonstrate that few tens of cells are enough to allow an accurate diagnostic decision.

4. Conclusions

About one quarter of the world population is found to be affected by anemia-related blood disorders. However, the diagnosis about such complex disease is not trivial. If left untreated, anemia can provoke severe symptoms, especially in adult age. Associating a proper diet to medical course of treatments, well-tailored to the specific anemia type, is beneficial for most of the patients. Due to the large number of subjects involved, the currently adopted CBCs, genetic tests and PBSOs represent a considerable cost for any NHS, and are not fully accurate in differential classification. Moreover, unaided PBSO under the optical microscope require the presence of trained personal and is subject to human errors. Hence, automated and accurate differential screening of anemias is a highly pursued goal. We have shown that through a holographic microscope and a new processing

workflow based on machine learning applied to quantitative phase-contrast images, it is possible to obtain a correct and rapid diagnosis in a simple way.

Our strategy is based on a max-voting hierarchical architecture that can serve as a decision support tool for hematologists, being able to accurately discern between healthy RBCs and five classes of anemias in label-free manner and using a very limited amount of blood sample.

The hierarchical structure has been found to be helpful in improving the performance of all the tested classifiers. In other words, the cascade of a binary classifier that discriminates healthy vs. sick RBCs, and then a differential classifier devoted to discern between the different anemia types, works better than using one single global decider, independently on the specific classifier or kernel used. We defined a 16 features subspace and, for each hierarchical stage, we extensively searched for the best performing classifier for the populations under analysis. This has been found to be the cubic SVM for both the binary and the differential datasets. We have obtained an accuracy in the binary stage up to 84.3% in validating the cubic SVM model but we have reached the 51.2% in test as the worst case in identifying the CDA I. Nevertheless, when the max-voting approach is adopted, the healthiness screener accomplishes the correct prediction any time, even in the CDA I borderline case. The positive effect of the max-voting is much more evident in the differential classification assays, where the anemia condition is identified correctly in all tested cases, even with data coming from different donors, i.e. never “seen” by the hierarchical decider. An intelligent holographic assay could be embedded in the standard diagnostic workflow as a specialized test to achieve a fast first diagnosis, with the future scope of replacing the second line investigation step, Fig. 4. We believe the proposed DH-based architecture could pave the way to the development of low-cost and even smart-phone based anemia prescreening tools to be employed at home by any single subject. This would dramatically attenuate the economic burden of blood screening and would also make anemia assays more objective and accurate. Noteworthy, the approach we proposed only requires few tens of cells, corresponding to a small fraction of blood drop that can be taken from the patient in minimally invasive way.

Transparency declaration

Authors report no conflict of interest relevant to this publication.

Declaration of competing interest

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

References

1. Ahmadzadeh, E., Jaferzadeh, K., Lee, J., Moon, I., 2017, J. of Biomed. Opt. 22(7), 076015.
2. Akrimi, J.A., Ahmad, A.R., George, L.E., 2013, International Journal of Science and Research (IJSR) 2(3), 140-142.
3. Andolfo, I., Russo, R., Gambale, A., Iolascon, A., 2016, Haematologica 101(11), 1284-1294.
4. Andolfo, I., Russo, R., Gambale, A., Iolascon, A., 2018a, Am J Hematol. 93, 107-121.
5. Andolfo, I., Russo, R., Rosato, B.E., Manna, F., Gambale, A., Brugnara, C., Iolascon, A., 2018b, Am J Hematol 93(12), 1509-1517.
6. de Benoist, B., McLean, E., Egli, I., Cogswell, M., eds. Worldwide prevalence of anaemia 1993-2005. WHO Global Database on Anaemia Geneva, World Health Organization, 2008.
7. Bianco, V., Mandracchia, B., Marchesano, V., Pagliarulo, V., Olivieri, F., Coppola, S., Paturzo, M., Ferraro, P., 2017, Light: Science & Applications 6(9), e17055-e17055.
8. Bianco, V., Memmolo, P., Carcagnì, P., Merola, F., Paturzo, M., Distante, C., Ferraro, P., 2020, Advanced Intelligent Systems 2(2), 1900153.
9. Briggs, C., Longair, I., Slavik, M., Thwaite, K., Mills, R., Thavaraja, V., Foster, A., Romanin, D., Machin, S.J., 2009, Int J Lab Hem. 31(1), 48-60.
10. Cacace, T., Bianco, V., Mandracchia, B., Pagliarulo, V., Paturzo, M., Ferraro, P., 2020, Biomed. Opt. Exp.
11. Cappellini, M.D., Russo, R., Andolfo, I., Iolascon, A., 2020, Hematology Am Soc Hematol Educ Program. 1, 465-470.
12. Ceelie, H., Dinkelaar, R.B., van Gelder, W., 2007, J Clin Pathol. 60(1), 72-79.
13. Chen, C.L., Mahjoubfar, A., Tai, L.C., Blaby, I.K., Huang, A., Niazi, K.R., Jalali, B., 2016, Sci. Report 6, 21471.
14. Da Costa, L., 2015, Clin Lab Med. 35(1), 105-122.
15. Donker, A.E., Schaap, C.C.L., Novotny, V.M.J., Smeets, R., Peters, T.M.A., van den Heuvel, B.L.P., Raphael, M.F., Rijneveld, A.W., Appel, I.M., Vlot, A.J., Versluijs, A.B., van Gelder, M., Granzen, B., Janssen, M.C.H., Rennings, A.J.M., van de Veerdonk, F.L., Brons, P.P.T., Bakkeren, D.L., Nijziel, M.R., Vlasveld, L.T., Swinkels, D.W., 2016, American Journal of Hematology 91(12), E482-E490.
16. Elgsaeter, A., Stokke, B.T., Mikkelsen, A., Branton, D., 1986. Science 234(4781), 1217-1223.
17. Hejna, M., Jorapur, A., Song, J.S., Judson, R.L., 2017, Sci Rep. 7, 11943.
18. Hoppe, C.C., 2013, Int. J. Lab. Hematol. 35(3), 297-305.
19. Iolascon, A., Andolfo, I., Russo, R., 2019, Br J Haematol. 187(1), 13-24.
20. Iolascon, A., Andolfo, I., Russo, R., 2020, Blood 136(11), 1274-1283.
21. Iolascon, A., Rivella, S., Anagnou, N., Camaschella, C., Swinkels, D., Muckenthaler, M., Porto, G., Barcellini, W., Andolfo, I., Risitano, A., Kattamis, A., Cappellini, M.D., Taher, A., De Franceschi, L., Rees, D., Russo, R., Tamary, H., Stauder, R., Girelli, D., on behalf of the EHA Scientific Working Group on "Red Cells and Iron", 2021, HemaSphere, 5(7), e607.

22. Jaferzadeh, K. and Moon, I., 2016, J. Biomed Opt. 21(12), 126015.
23. Jaferzadeh, K., Sim, M., Kim, N., Moon, I., 2019, Sci. Rep. 9(14062).
24. Jo, Y., Cho, H., Lee, S.Y., Choi, G., Kim, G., Min, H., Park, Y.K., 2019, IEEE Journ. Sel. Topics Quant. Electron. 25(1).
25. Khan, J.R., Chowdhury, S., Islam, H., Raheem, E., 2019, Journal of Data Science 17(1), 195-218.
26. Kim, G., Jo, Y.J., Cho, H., Min, H.-S., Park, Y.K., 2019, Biosensors and Bioelectronics 123, 69-76.
27. King, M.J., Telfer, P., MacKinnon, H., Langabeer, L., McMahon, C., Darbyshire, P., Dhermy, D., 2008, Cytometry B Clin Cytom. 74, 244-250.
28. King, M.J., Garçon, L., Hoyer, J.D., Iolascon, A., Picard, V., Stewart, G., Bianchi, P., Lee, S.H., Zanella, A., 2015, Int Journ Lab Hem 37, 304-325.
29. Kleiber, A., Ramoji, A., Mayer, G., Neugebauer, U., Poppa, J., Henkel, T., 2020, Lab Chip.
30. Memmolo, P., Miccio, L., Merola, F., Gennari, O., Netti, P.A., Ferraro, P., 2014, Cytometry part A 85(12), 1030-1036.
31. Miccio, L., Cimmino, F., Kurelac, I., Villone, M.M., Bianco, V., Memmolo, P., Merola, F., Mugnano, M., Capasso, M., Iolascon, A., Maffettone, P.L., Ferraro, P., 2020, View 1(3).
32. Mir, M., Tangella, K., Popescu, G., 2011, Biomed. Opt. Exp. 2(12), 3259-3266.
33. Mitani, A., Huang, A., Venugopalan, S., Corrado, G.S., Peng, L., Webster, D.R., Hammel, N., Liu, Y., Varadarajan, A.V., 2020, Nature Biomedical Engineering 4, 18-27.
34. Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., Van Valen, D., 2019, Nature Methods 16, 1233-1246.
35. Moon, I., Javidi, B., Yi, F., Boss, D., Marquet, P., 2012, Opt. Express 20(9), 10295-10309.
36. Moon, I., Anand, A., Cruz, M., Javidi B., 2013, IEEE Photonics Journal 5(5), 6900207-6900207.
37. Muckenthaler, M.U., Rivella, S., Hentze, M.W., Galy, B., 2017, Cell 168(3), 344-361.
38. Mugnano, M., Memmolo, P., Miccio, L., Merola, F., Bianco, V., Bramanti, A., Gambale, A., Russo, R., Andolfo, I., Iolascon, A., Ferraro, P., 2018, Analytical chemistry 90(12), 7495-7501.
39. Natarajan, V., Sarkar, D., 2019, Ensemble Machine Learning Cookbook: Over 35 Practical Recipes to Explore Ensemble Machine Learning Techniques Using Python, Wiley (Ed), 2019. ISBN: 1789136601.
40. O'Connor, T., Doblas, A., Javidi, B., 2019, Opt. Lett. 44(9), 2326.
41. O'Connor, T., Shen, J., Liang, B.T., Javidi, B., 2021, Opt. Lett. 46(10), 2344-2347.
42. Park, Y., Best, C.A., Badizadegan, K., Dasari, R., Feld, M., Kuriabova, T., Henle, M., Levine, A., Popescu, G., 2010, PNAS 107(15), 6731-6736.
43. Patel, T., Trivedi, V., Mahajan, S., Chaniwal, V., Fournier, C., Lee, S., Javidi, B., Anand, A., 2018, Biomed. Opt. Express 9(6), 2779-2784.
44. Pham, H.V., Bhaduri, B., Tangella, K., Best-Popescu, C., Popescu, G., 2013, PloS one 8(2), e55676.
45. Qaseem, A., Humphrey, L.L., Fitterman, N., Starkey, M., Shekelle, P., 2013, Annals of internal medicine 159, 11, 770-779.

46. Rappaz, B., Barbul, A., Hoffmann, A., Boss, D., Korenstein, R., Depeursinge, C., Magistretti, P.J., Marquet, P., 2009, Blood Cells, Mol. Dis. 42(3), 228-232.
47. Russo, R., Gambale, A., Langella, C., Andolfo, I., Unal, S., Iolascon, A., 2014, The American Journal of Hematology 89(10), E169-E175.
48. Russo, R., Andolfo, I., Manna, F., Gambale, A., Marra, R., Rosato, B.E., Caforio, P., Pinto, V., Pignataro, P., Radhakrishnan, K., Unal, S., Tomaiuolo, G., Forni, G.L., Iolascon, A., 2018, Am J Hematol. 93(5), 672-682.
49. Russo, R., Marra, R., Rosato, B.E., Iolascon, A., Andolfo, I. 2020, Front Physiol 11, 613559.
50. Seo, S., Isikman, S.O., Sencan, I., Mudanyali, O., Su, T., Bishara, W., Erlinger, A., Ozcan, A., 2010, Anal Chem. 82(11), 4621-4627.
51. Shaked, N.T., Satterwhite, L.L., Truskey, G.A., Wax, A.P., Telen, M.J. 2011, Journ Biomed Opt 16(3), 030506.
52. Singh, A., 2018, A Comprehensive Guide to Ensemble Learning (with Python codes). https://www.analyticsvidhya.com
53. Singla, N., Srivastava, V., 2020, Optics & Laser Technology 130, 106335.
54. Smith, R.E., 2010, The American Journal of Managed Care 16, S59-S66.
55. Spence, R.K., 2010, Heart Fail Clin 6(3), 373-383.
56. Xu, M., Papageorgiou, D.P., Abidi, S.Z., Dao, M., Zhao, H., Karniadakis, G.E., 2017, PLoS Comput Biol 13(10), e1005746.
57. Yi, F., Moon, I., Lee, Y., 2015, J. Biomed. Opt. 20, 016005.
58. Zhang, Y., Koydemir, H.C., Shimogawa, M.M., Yalcin, S., Guziak, A., Liu, T., Oguz, I., Huang, Y., Bai, B., Luo, Y., Luo, Y., Wei, Z., Wang, H., Bianco, V., Zhang, B., Nadkarni, R., Hill, K., Ozcan, A., 2018, Light: Science & Applications 7(108), 1-18.
59. Zhu, H., Sencan, I., Wong, J., Dimitrov, S., Tseng, D., Nagashima, K., Ozcan, A. 2013, Lab Chip 13(7), 1282-1288.