Deep learning-assisted ultra-accurate smartphone testing of paper-based colorimetric ELISA assays

Sixuan Duan [a] [b], Tianyu Cai [a], Jia Zhu [a] [c], Xi Yang [a], Eng Gee Lim [a], Kaizhu Huang [d], Kai Hoettges [b], Quan Zhang [a], Hao Fu [e], Qiang Guo [f], Xinyu Liu [g], Zuming Yang [h], Pengfei Song [a] [b]
Show more

Abstract

Smartphone has long been considered as one excellent platform for disease screening and diagnosis, especially when combined with microfluidic paper-based analytical devices (μPADs) that feature low cost, ease of use, and pump-free operations. In this paper, we report a deep learning-assisted smartphone platform for ultra-accurate testing of paper-based microfluidic colorimetric enzyme-linked immunosorbent assay (c-ELISA). Different from existing smartphone-based μPAD platforms, whose sensing reliability is suffered from uncontrolled ambient lighting conditions, our platform is able to eliminate those random lighting influences for enhanced sensing accuracy. We first constructed a dataset that contains c-ELISA results (n = 2048) of rabbit IgG as the model target on μPADs under eight controlled lighting conditions. Those images are then used to train four different mainstream deep learning algorithms. By training with these images, the deep learning algorithms can well eliminate the influences of lighting conditions. Among them, the GoogLeNet algorithm gives the highest accuracy (>97%) in quantitative rabbit IgG concentration classification/prediction, which also provides 4% higher area under curve (AUC) value than that of the traditional curve fitting results analysis method. In addition, we fully automate the whole sensing process and achieve the "image in, answer out" to maximize the convenience of the smartphone. A simple and user-friendly smartphone application has been developed that controls the whole process. This newly developed platform further enhances the sensing performance of μPADs for use by laypersons in low-resource areas and can be facilely adapted to the real disease protein biomarkers detection by c-ELISA on μPADs.

1. Introduction

Microfluidic paper-based analytical devices (μPADs) have been demonstrated as a promising tool for detecting disease-relevant biomarkers at point-of-care testing (POCT), thanks to their low-cost, ease of operation, and self-driven capillary fluidic flow [[1], [2], [3], [4], [5]]. In the last decade, many biosensing mechanisms, such as electrochemistry, chemiluminescence, fluorescence immunoassay, and colorimetric enzyme-linked immunosorbent assay (c-ELISA) have been developed and realized on μPADs [[6], [7], [8], [9], [10]]. Among them, colorimetric μPADs are particularly suitable for being used in low resources settings, especially when using smartphones as colorimetric reading and analyzing tools, and no further complicated equipment is needed [11,12]. The smartphones feature the broadest user

community, a friendly operator interface, and excellent data transfer capability. It is a powerful computation platform capable of fully automating the colorimetric results collection, analysis, and displaying process, using its camera, processor, and screen, respectively [[13], [14], [15]]. Meanwhile, it is usually cheaper than the dedicated medical-grade colorimetric reader [16]. Hence, the smartphone-based colorimetric μPADs well meet all the criteria outlined by the World Health Organization (WHO) for POCT devices in low-resource settings, abbreviated as "ASSURED"; affordable, sensitive, specific, user-friendly, rapid and robust, equipment-free, and delivered [17]. And it can significantly improve healthcare worldwide by providing low-cost, timely, accurate disease screening.
Recently, smartphone-based colorimetric μPADs have been extensively studied and developed to detect various disease biomarkers such as glucose, cholesterol, and uric acid [15,[18], [19], [20]]. The c-ELISA is the most widely used analytical assay as it is the gold standard for detecting protein biomarkers in disease-related clinical samples [10,[21], [22], [23]]. The c-ELISA produces color signals that are quantitatively correlated with the different concentrations of the sensing target molecules, and those signals can be easily picked up by smartphones; however, smartphone cameras are subject to various camera settings and, more severely, to different environmental lighting conditions [[24], [25], [26], [27]]. Therefore, the accuracy (the ratio between the number of true positives and negatives to the total number of experimental results) and sensitivity (the ratio between the number of true positives to the number of true positives and false negatives) of the smartphone-based colorimetric μPADs are greatly jeopardized, and the smartphone has shown the lowest sensing performance when compared with the standard colorimetric μPADs analysis methods: the RGB sensor and RGB scanner, with or without the use of ambient lighting prevention tools [23,28]. Thus, effectively minimizing the influence of internal camera optics/settings and external ambient light conditions is still highly desired to improve the analytical performance of smartphone-based colorimetric μPADs, thereby facilitating the development of POCT in resource-limited areas.
Most recently, machine learning has been demonstrated as one of the most potent image data processing tools [29,30], and it can be readily implemented on smartphone platforms. However, the smartphone is not powerful enough to run the machine learning algorithms to process the data, which can be wirelessly uploaded to/downloaded from the cloud server before and after the processing [14,31,32]. Several works have applied machine learning algorithms to analyze the colorimetric signal for the lateral flow assays (LFA), and highly reliable qualitative and quantitative results were obtained [[33], [34], [35], [36], [37]]. However, LFA is much simpler than c-ELISA μPADs and usually cannot enable high sensitivity and high specificity biosensing [38]. Machine learning has also been used on smartphone-based colorimetric μPADs and provides higher accuracy [14,[39], [40], [41]]. But they all used conventional machine learning algorithm that requires human feature extraction, further increasing human intervention and hindering its usage by laypersons. Subsequently, two works have combined deep learning algorithms with smartphone paper devices for disease diagnosis, which can automatically extract features and achieve high accuracy, thanks to its higher computation power than conventional machine learning. But in these works, an additional homemade cassette is needed to prevent the ambient light [32,42], which increases the device complexity. Most recently, Ning et al. employed a deep learning algorithm for the colorimetric detection of C-reactive protein (CRP) and provided good accuracy of 96% [43]. However, this work relied on the manual image and data transfer between the smartphone and deep learning servers, which hinders the wide accessibility of the POCT devices. Therefore, a fully automated smartphone-based colorimetric μPAD that is free of the influences of human inventions and ambient lighting conditions is still highly desired and yet to develop.

In this work, we present a deep learning-assisted smartphone colorimetric μPADs platform for rapid, sensitive, and accurate detection of protein markers using c-ELISA. Unlike the previous work, there is no need to manually copy the resultant images extracted from the smartphone to the local computer, and the data transfer can be performed directly from the smartphone. In addition, a deep learning model is used to further avoid manual screening of features. Furthermore, an Android application with a friendly user interface is developed to interact with the whole process, and it enables the fully automated, streamlined manner of "image in, answer out" colorimetric sensing. Here, the complete automation is contributed by the deep learning algorithms. Specifically, we first constructed a comprehensive dataset (number = 2048) using Rabbit IgG as the c-ELISA sensing target on μPADs. Then, the three light sources and two different smartphones were used to enrich the complexity of the dataset for training the algorithm, with the objective of increasing the robustness and adaptability of the algorithm. By training all the images of experimental results under different lighting conditions, the deep learning algorithm can learn itself and eliminate the effect of illumination on the experimental results. In this work, we compared the performance of the four mainstream deep learning algorithms (AlexNet, GoogLeNet, ResNet34, and MobileNet_V2). Among them, GoogLeNet provides the highest accuracy (>97%), which also provides a 4% higher area under curve (AUC) value than that of the conventional smartphone sensing results analysis method that uses curve fitting. The overview of deep learning-assisted ultra-accurate smartphone testing of paper-based c-ELISA is shown in Fig. 1. In this work, we validated the feasibility of our platform by using rabbit IgG protein as a sensing target for c-ELISA, which is the most common sensing target protein and has been widely used for validation experiments in many works [10,23]. In future work, we will apply this approach to clinically relevant biomarkers. To sum up, the unique capabilities of our platform further enhance the advantages of μPAD for general use in resource-poor settings and hold great promise for real-world use.

2. Materials and methods

2.1. Reagents and materials

Potassium periodate (>99.5%, 80,106,916) was purchased from Sinopharm Chemical Reagent Co., Ltd (Shanghai, China). Rabbit IgG (I5006) was purchased from Sigma-Aldrich (Shanghai, China). Alkaline phosphatase (ALP) conjugated anti-rabbit IgG (A0239), 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT) alkaline phosphatase color development kit (C3206) was purchased from the Beyotime Institute of Biotechnology (Haimen, China). 10 × phosphate-buffered salines (PBS, G4207), Tween-20 (G5058), and bovine serum albumin (BSA, G5001) were purchased from Servicebio (Wuhan, China). The 10 × PBS was diluted using deionized (DI) water to 1 × PBS.

2.2. c-ELISA for detection of rabbit IgG in PBS on μPAD

We first used an aldehyde functionalization step to treat the pristine paper substrate before making it into the functional μPAD, in order to enhance the immobilize proteins' immobilization efficiency. Briefly, the 100 mm × 100 mm piece of Whatman No. 1 chromatography paper was placed in a 150 mm × 150 mm glass Petri dish, after which 100 ml of 0.031 M KIO_4 was added and put on a hot plate (CS-15956-31, Cole-Parmer) at 65 °C for 2 h with stirring [44]. Subsequently, the modified papers were stored in a humidity-proof cabinet (temperature: 24 °C, humidity: 30%) before use. We then designed the μPAD

with the multi-well format (diameter of the well = 5.6 mm) in AutoCAD software and printed solid wax patterns on the above pre-treated paper by a wax printer (8580DN, Xerox). It was then placed on a hot plate at 120 °C for 30 s to melt the printed wax and form a multi-well μPAD.

The schematic diagram of the protocol for a direct c-ELISA performed on our multi-well μPAD is shown in Fig. 2 [10]. The rabbit IgG was the most common model sensing target for the c-ELISA assay and was also used in this study. A polymethyl methacrylate (PMMA) holder was used to support the paper-based multi-well microplates and allow complete suspension and isolation of the test wells (avoiding the contamination caused by direct contact with the lab bench surface). Briefly, 3 μl of different concentrations (in 10-fold dilutions, 6.7 pM to 6.7 μM) of rabbit IgG antigen was first immobilized in the test wells. After 10 min of incubation under ambient conditions, the test well was blocked with 3 μl of blocking buffer (1% w/v BSA and 0.05% w/v Tween-20 in PBS) and allowed to dry for 10 min under ambient conditions. Then, 3 μl of ALP-conjugated anti-rabbit IgG antibody (1:100 dilutions of the stock antibody produced in goat in 0.05% w/v Tween-20 in PBS) was added to the test well and incubated for 1 min on the hot plate at 37 °C. Immediately, the test wells were washed twice with 10 μl of 1 × PBS. Finally, 3 μL of BCIP/NBT substrate solution (pH 9.5) was added to each well for color initiation (30 min).

2.3. Comprehensive dataset under different lighting conditions and smartphones

Different ambient lights and camera optics can significantly influence the accuracy of paper-based c-ELISA testing results when the smartphone is used as the reader [34]. Deep learning algorithms can mitigate those influences and achieve high accuracy for quantitative prediction of c-ELISA testing results by training with a comprehensive dataset that consists of colorimetric images (labeled with known sensing target concentration) under different lighting conditions [14]. To compensate for the influence posed by different smartphones, we also used two mainstream smartphones to take the colorimetric images. Firstly, we used a 3D-printed black box with a height of 130 mm to block the outside light and captured the images of the paper-based multi-well microplates under the automated flash of a smartphone. The smartphone flash was then replaced with three light sources: fluorescent light (F), table lamp (L), and natural light (N), which were used individually or in combination to form seven different lighting groups (F, L, N, FL, FN, LN, FLN) to mimic the different light conditions (as shown in Fig. 3). The fluorescent light (Philips 30 W) emits white colors (6500 K), the table lamp bulb (Midea 3 W) provides daylight colors (5700 K), and the experiments were conducted during summer in Suzhou (China), which provides around 2000 lx light intensity [45]. The image of the paper-based multi-well microplate was captured in a vertical position at a constant distance of 130 mm between the smartphone and the microplate.

As image quantification is affected by the smartphone camera resolution and optics [46]. Table S1 shows two mainstream smartphones (Android and iOS) used in this work. The smartphone was kept at the same height (130 mm) in capturing each concentration image for all light conditions. The smartphone was set to the default mode, further reducing user intervention. Eight concentrations (N = 16) for eight lighting conditions lead to having 1024 images per smartphone, resulting in 2048 images of 128 pixels × 128 pixels size in the entire dataset. The details of the comprehensive dataset under different lighting conditions and smartphones are shown in Table S2.

2.4. Image processing

We also applied the image processing open library (OpenCV) to automatically extract the microwell area (region of interest (ROI)) from the multi-well paper chip (as shown in Fig. 4a). The region of interest (ROI) consists of the circular detection well that produces the colorimetric signal and its outer rectangle. In the process of extracting ROI, the RGB values of the input image can be directly converted to the grayscale values according to the equation:

$$Gray = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \quad (1)$$

R, G, and B correspond to the pixel's color, respectively. The parameters of this equation are mainly assigned based on the ratio of the light-sensitive intensity of three different photoreceptor cells in the human eye, which is commonly used for this purpose [[47], [48], [49]]. Then, the image is Gaussian filtered to remove noise and facilitate the extraction of circular detection well contours. Otsu's method will be used to further convert the grayscale image into a binarized (black and white) image. Otsu's method will first assume the circular detection well and the background parts of the image, and then automatically fine-tune and determine a final optimal threshold to preferably distinguish the circular detection well from the background by a statistical method (maximum between-class variance). Subsequently, the circular detection well is set to 255 (white), and the background part is set to 0 (black). Then, we use the morphological operation function to remove the noise around the circular detection wells to make their boundaries smoother. It is convenient for canny edge detection and finding the contours of circular detection wells. The final ROI contours will be extracted and saved to the specified folder. The whole process of image processing is shown in Fig. 4. The final generated image containing an image of 128 × 128 pixel size will be transferred to a deep learning algorithm for further processing. The details of the image functions we used can be found in the supplementary materials.

2.5. Comparison of different deep learning algorithms

To identify the correlation relationship between different protein concentrations and the colors of the assay results, we randomly selected 80% of the images from the dataset (n = 2048) to train four mainstream deep learning algorithms (AlexNet [50], GoogLeNet [51], ResNet34 [52], MobileNet_V2 [53]) and evaluated the ability of these algorithms to predict rabbit IgG protein concentration accurately. The algorithm with the highest prediction accuracy was selected and integrated into the final smartphone-based platform among the tested algorithms.

The input layer of the general deep learning algorithms contains the image dataset, consisting of two basic parts: feature extraction and classification. Feature extraction consists of several convolutions and rectified linear unit (ReLU) followed by a pooling function. Moreover, the classification part usually involves the fully connected layers. The architecture of the generalmodel is shown in Fig. 5. In this work, the AlexNet algorithm consists of five convolutional layers and three fully connected layers, and it improves the speed of training by employing an activation function (ReLU) [50]. After ReLU, the local response normalization (LRN) layer is used to enhance the generalization ability of the model and further improve its prediction accuracy. The response-normalized activity ($b^i_{x,y}$) equation is

$$b_{x,y}^{i} = a_{x,y}^{i} / \left(k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} \left(a_{x,y}^{j} \right)^{2} \right)^{\beta}$$

(2)

Where $a_{x,y}^{i}$ is the output value after the ReLU. *N* is the total number of kernels. *K* =2, *n*=5, α 10−4 and β=0.75 are hyper-parameters [50]. And the GoogleNet algorithm uses night inception modular structures with 22 layers to reduce the computation and improve the convergence speed. The main feature of the model structure of the GoogLeNet algorithm is the enhanced utilization of computational resources inside the network. Average pooling is used instead of fully connected layers to minimize the model parameters and improve the accuracy of classification [51]. Similarly, the Resnet34 consists of 5 convolutional clusters and uses a residual network approach to deepen the network and better learn the features of the data. The residual mapping to be learned can be expressed as $\mathcal{F}(x, \{W_i\})$ so the relationship between input (x) and output (y) vectors of the layers can be represented using Eq (3) [52].

$$y = \mathcal{F}(x, \{W_i\}) + x \quad (3)$$

In the end, the Mobilenet_V2 is a lightweight neural network model with 17 layers. The core of its implementation of model lightweight is depth wise separable convolution, which can be expressed mathematically as:

$$\widehat{G}_{k,l,n} = \sum_{i,j} \widehat{K}_{i,j,m} \bullet F_{k+i-1,l+j-1,m} \quad (4)$$

Where $\widehat{K}$ is the depthwise convolutional kernel, $F_{k+i-1,l+j-1,m}$ is the input feature map, which will produce a feature map $\widehat{G}_{k,l,n}$. T is smaller and faster while ensuring the accuracy of the model [53].

Thanks to the advantage of the deep learning algorithms, it can automatically extract features from the original images for the training dataset [54], which will further reduce human intervention. Moreover, we use all the deep learning algorithms pre-trained on a large ImageNet dataset with 1000 categories. Therefore, four fully connected layers are added after the algorithm to accommodate the eight categories of our dataset to make the algorithm more suitable for the present dataset.

2.6. Smartphone application

Android system has the advantages of low development cost and high versatility [33]. Therefore, we developed an Android application with simple and user-friendly interfaces capable of automatic image processing and result analysis, further providing a more convenient data analysis platform. In addition, this application can satisfy the needs of different users in various scenarios. This work utilized Vue. js (JavaScript framework) [55] to create the user interface and Flask python to implement the application's primary functions [56]. Subsequently, this application can be applied to both smartphones and web pages.

Fig. S1 shows the screenshots of the application at different stages. Users can enter a new test by logging into the application (as shown in Fig. S1a) and clicking the “ADD test” button (as shown In Fig. S1b). Next, the user enters information about the test, such as name, test type, etc. The application will allow the user to capture a new image using the smartphone's camera or load the image from an album (as shown in Fig. S1c). Then, after tapping the “submit” button, the application will send the captured image to the cloud server. The ROI in the captured image will be automatically obtained and transferred to the deep learning algorithms, which will provide the final results. The result will be sent back to the application and displayed on the smartphone screen (as shown in Fig. S1d), which concludes the fully automated “image in, answer out” sensing process.

3. Results and discussion

3.1. Lighting conditions disturb c-ELISA results

In this work, a direct c-ELISA was used to detect the concentration of Rabbit IgG in PBS. We prepared seven different concentrations (N = 16 for each concentration) in 10-fold dilutions (6.7 pM–6.7 μM) ofrabbit IgG as the sensing target for c-ELISA, and PBS buffer (0 pM) was used as a negative control. After each test, we captured images of the same microwells detection zones under different lighting conditions with two different smartphones (iPhone 11 and Huawei P40 Pro), respectively. The surface of the microwell detection zone is purple, and the higher the rabbit IgG concentration, the darker the color of the detection zone. Fig. 6a shows the resulting photographs of different rabbit IgG concentrations captured by iPhone 11 under eight lighting conditions (No light, F, L, N, F + L, F + N, F + L + N, and L + N). We measured the average grayscale intensity of all test areas using ImageJ first. Then, the intensity data were fitted using the Hill equation (Fig. 6b) for the sigmoidal curve (s-curve), and the limit of detection (LOD) and coefficient of determination (COD, denoted as R^2) corresponding to each s-curve were calculated. Where the LOD is determined by using the 3σ of the lowest concentration sample [23], and the COD represents the strength of a curve fitting, and the better the curve fit, the closer the value of COD is to 1. We also verified the results of direct c-ELISA of rabbit IgG in PBS under No light conditions with two different smartphones (as shown in Fig. S2).
The example photos of the paper microwell shown in Fig. 6a confirmed that light conditions greatly affect the colorimetric signals, and the different calibration curves for the same c-ELSA assays are generated, as shown in Fig. 6b. From those calibration curves, different quantitative results can be obtained even for the same c-ELSA assays, which greatly disturbs the accuracy of the results reading. Moreover, the variations of LOD are as large as 50.24% (No light: LOD = 550 pM, COD = 0.945; under fluorescent light: LOD = 780 pM, COD = 0.978; under table lamp: LOD = 840 pM, COD = 0.989; under natural light: LOD = 418 pM, COD = 0.913.). These further demonstrated the urgency of removing the influences of light conditions.

3.2. Construction of the image dataset

In this work, we prepared eight experimental groups with different concentrations of Rabbit IgG. It contains a control group with a concentration of 0 (PBS buffer) and seven groups with different concentrations from 6.7 pM to 6.7 μM in 10-fold dilutions. Subsequently, each concentration will be repeated sixteen times of testing. Here, we will get $8\times16=128$ c-ELISA experimental results. In addition, we set eight different light conditions (No light, F, L, N,

F + L, F + N, F + L + N, and L + N) to simulate the illumination conditions that may be encountered in the experiment. Therefore, considering these eight different lighting conditions, a total of 128×8=1024 experimental results will be obtained. Then, two commonly used smartphones (iPhone 11 and Huawei P40 Pro) will be applied to capture all the experimental results, resulting in 1024×2=2048 images.

3.3. Different deep learning algorithms show varied performance

In recent works, deep learning has shown excellent signal predictive capability by efficiently removing the noise, especially in image recognitions and classifications [50,57]. We applied four current mainstream deep learning algorithms (AlexNet, GoogLeNet, ResNet34, MobileNet_V2) to quantify rabbit IgG concentrations on μPADs. First, we collected eight images of μPADs, and each μPADs image contained 16 replicate experiments of the same concentration. In addition, each μPAD image was collected by two different phones under eight different illumination conditions, resulting in a final dataset of 2048 images (2×8×8×16 = 2048 images). Before training the deep learning algorithm using this dataset, the collected images are automatically processed. Since the reaction area is light in color that shares good contrast with the dark surrounding area, the image processing algorithm can recognize the reaction area of μPADs easily by setting the appropriate threshold. Then, image processing is employed to obtain ROI, and after receiving the contours of the test well, the final data set of the same size is obtained by cropping along the contours. Next, we randomly selected 80% of the dataset images to train the model, and the remaining 20% was used as the test dataset. Finally, the deep learning algorithms will classify the test dataset to determine the concentration classes corresponding to the images.

First, the performance of a deep learning algorithm is evaluated in terms of accuracy and f1 score. Accuracy is defined as the sum up of true positive and negative rates in biosensor terminology (Eq. (5)). In biosensor communities and in this study, the true positives (TP) means that the sensors correctly give the positive classification of the samples (the class of non-zero IgG concentration in our study), and the false positives are defined that sensor incorrectly determine the negative samples to be the positive ones. The true negatives and false negatives follow the same. As shown in Eq. (6), the F1 score is the harmonic mean of precision (the precision for the positive class) and recall (known as the sensitivity, which is the ratio between true positive samples and samples classified positive). Also, a higher F1 score means that the deep learning algorithm has low false positives and false negatives [58]. As the accuracy and f1 score increase, it indicates better deep learning performance.

(5) $$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

(6) $$F1 = \frac{2TP}{2TP+FP+FN}$$

The performance comparison of AlexNet, GoogLeNet, ResNet34, and MobileNet_V2 algorithms is shown in Table 2. Then, to further evaluate the performance of the deep learning algorithm, we also provide the processing time and efficiency time in Table 2. Here, the processing time is the time required to compute one epoch, and the efficiency time is the training time required for the model to converge. Among them, GoogleNet provides the highest accuracy (>97%) and F1 score (>0.9713) due to its use of the inception module, which widens the network structure and thus increases the depth and width of the model,

allowing for a higher dimensionality of the learned features and more feature parameters to be extracted [59]. However, GoogLeNet takes the highest processing time and efficiency time. In this work, we pay more attention to the accuracy of the model. Because the accuracy reflects the sensitivity of the analytical method, and the high sensitivity is more important in immunoassays. In addition, once the deep learning model is trained, the subsequent experimental results can be tested. And this detection process can be completed in about 1 s. Therefore, the processing time and the efficiency time of the deep learning algorithm do not affect the detection time.

Table 2. Comparison of algorithms' performance.

Classification method	Accuracy	F1 score	Processing time	Efficiency time
MobileNet_V2	95.05%	0.95060	12s	5 m 49s
AlexNet	96.32%	0.96321	11s	4 m 31s
ResNet34	96.88%	0.96879	13s	18 m 54s
GoogLeNet	97.14%	0.97131	16s	49 m 31s

However, the accuracy of the different deep learning algorithms for predicting different Rabbit IgG concentrations varies. To show the algorithms' accuracy at different concentrations, the confusion matrix for each algorithm is given in Fig. 7, showing the correlation between the actual and predicted labels (concentration class) of the test dataset. And the confusion matrix is used in this paper to compare the predicted (calculated) and actual values of Rabbit IgG concentration for the four algorithms. Where, for the test dataset (around 20% × 2048, here we take 384 samples as test dataset), each concentration class has 48 (384/8) samples. Fig. 7 represents the confusion matrix of AlexNet (Fig. 7a), GoogLeNet (Fig. 7b), ResNet34 (Fig. 7c), and MobileNet_V2 (Fig. 7d) respectively. Here, the darker the color in the confusion matrix, the higher the number of correct predictions.

Then, we take the GoogLeNet algorithm as an example. The GoogLeNet algorithm provides ultra-high accuracy (100%) for classifying 6.7 μM, 670 nM, and 67 nM concentrations. It should be emphasized that although AlexNet, GoogLeNet, ResNet34, and MobileNet_V2 all show high performance (>95% accuracy) in image classification, the results are inconsistent for the same dataset due to their different model structures [60,61]. In addition, the number of correct predictions of the same deep learning algorithm for the test dataset categories is randomized, so the accuracy of GoogLeNet algorithm predictions for different concentrations is also different. Therefore, we provide the repeatability validation results of the GoogLeNet model. From Table 3, we can see that the GoogLeNet model provides an average accuracy of 97.582%, further demonstrating the reproducibility of the model.

Table 3. Repeatability validation results of the GoogLeNet model.

Number of times	Accuracy
1	97.13%
2	97.92%
3	97.40%

Number of times	Accuracy
4	97.66%
5	97.92%
6	97.40%
7	97.79%
8	97.66%
9	97.13%
10	97.79%
Average	97.58%
SD	0.299%

3.4. GoogLeNet outperformed the conventional c-ELISA analysis method

In addition, for a more intuitive comparison with traditional analysis methods, as shown in Fig. 8, we compared the ROC curves of the GoogLeNet algorithm model with that of the standard analysis method (OriginPro analysis software). The x-axis of the ROC curve represents 1-specificity, the y-axis represents sensitivity, and the area under curve (AUC) is defined as the area enclosed by the ROC curve and the coordinate axis. Further, the ROC and AUC are valid methods to assess the performance of diagnostic tests [62,63]. Moreover, it is worth noting that the sensitivity and specificity in ROC curves are not the parameters for evaluating the biosensor itself. Otherwise, they indicate the true positive and true negative rates of the sensing experiments, respectively. Here, for the standard analysis method, the PBS control experiment was used as a negative control, and other non-zero concentration samples were used as the positive samples. It can be seen that the ROC curve differs under different lighting conditions, further demonstrating the necessity of removing the influences of lighting. For a ROC curve, the closer the curve is to the upper left corner, the higher the sensitivity, specificity, and AUC value are. The closer the AUC is to 1, the better the diagnostic performance [63]. As shown in Fig. 8a, the GoogLeNet algorithm offers an AUC value of 0.99 (Micro-AVG). The standard analysis method provides the largest AUC value (0.95) under fluorescent lighting conditions, as shown in Fig. 8b. This further proves that the deep learning algorithm is one suitable and excellent analytical method for colorimetric detection. The ROC curve of AlexNet, ResNet34, and MobileNet_V2 algorithms are shown in Fig. S3.

After testing four deep learning algorithms, the GoogLeNet algorithm was integrated into our custom-designed Android application for c-ELISA concentration testing because of its good prediction results (the predicted accuracy >97%). Further, the AUC value of the GoogLeNet algorithm also provided 4% higher than standard analysis methods. Since the deep learning algorithm still provided ultra-high accuracy under different illumination, this indicates that the deep learning-assisted smartphone platform could accurately predict the results of the μPADs c-ELISA assay and eliminate the effect of light on the experimental results.

4. Conclusion

This paper reports a deep learning-assisted smartphone platform for ultra-accurate detection of protein markers by μPAD c-ELISA. The platform uses a deep learning algorithm trained

from images acquired using two different smartphones under eight lighting conditions to improve the platform's robustness to lighting variations and camera optics. Unlike existing platforms, our platform can fully automatically extract c-ELISA features from the original images regardless of smartphone brand, time, and location with ultra-high accuracy (>97%) for quantitative classification/prediction of c-ELISA results. In addition, a custom Android application allowing image processing has been developed. The image processing tools embedded in the application automatically find ROI and reduce human error, making the platform more user-friendly and accurate. Our platform's unique features further enhance the advantages of μPAD for general use in low-resource areas and offer great promise for real-world use. In this work, we select commonly used convolutional neural network (CNN) models (AlexNet, GoogLeNet, ResNet34, and MobileNet_V2), and these four algorithms have different model structures and reflect the development of CNN models. In our future work, we will continue to explore the new algorithms and apply them to our future experiments.

CRediT authorship contribution statement

Sixuan Duan: Conceptualization, Writing – original draft, Writing – review & editing. **Tianyu Cai:** Visualization. **Jia Zhu:** Investigation. **Xi Yang:** Methodology. **Eng Gee Lim:** Conceptualization. **Kaizhu Huang:** Conceptualization. **Kai Hoettges:** Conceptualization. **Quan Zhang:** Visualization. **Hao Fu:** Methodology. **Qiang Guo:** Methodology. **Xinyu Liu:** Supervision. **Zuming Yang:** Writing – review & editing. **Pengfei Song:** Conceptualization, Writing – review & editing.

Declaration of competing interest

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

Acknowledgment

The authors thank the financial support from the programs of Natural Science Foundation of the Jiangsu Higher Education (20KJB460024, 22KJB460033), Jiangsu Science and Technology Programme - Young Scholar (BK2020041995), Jiangsu Province High-level Innovation and Entrepreneurship Talent Plan (2020–30803), XJTLU Key Programme Special Fund – Exploratory Research Programme (KSF-E-39), and XJTLU Research Development Fund (RDF-18-02-20). The authors also acknowledge the financial support from Xi'an Jiaotong – Liverpool University to S. Duan (PGRS1912019).

Figures

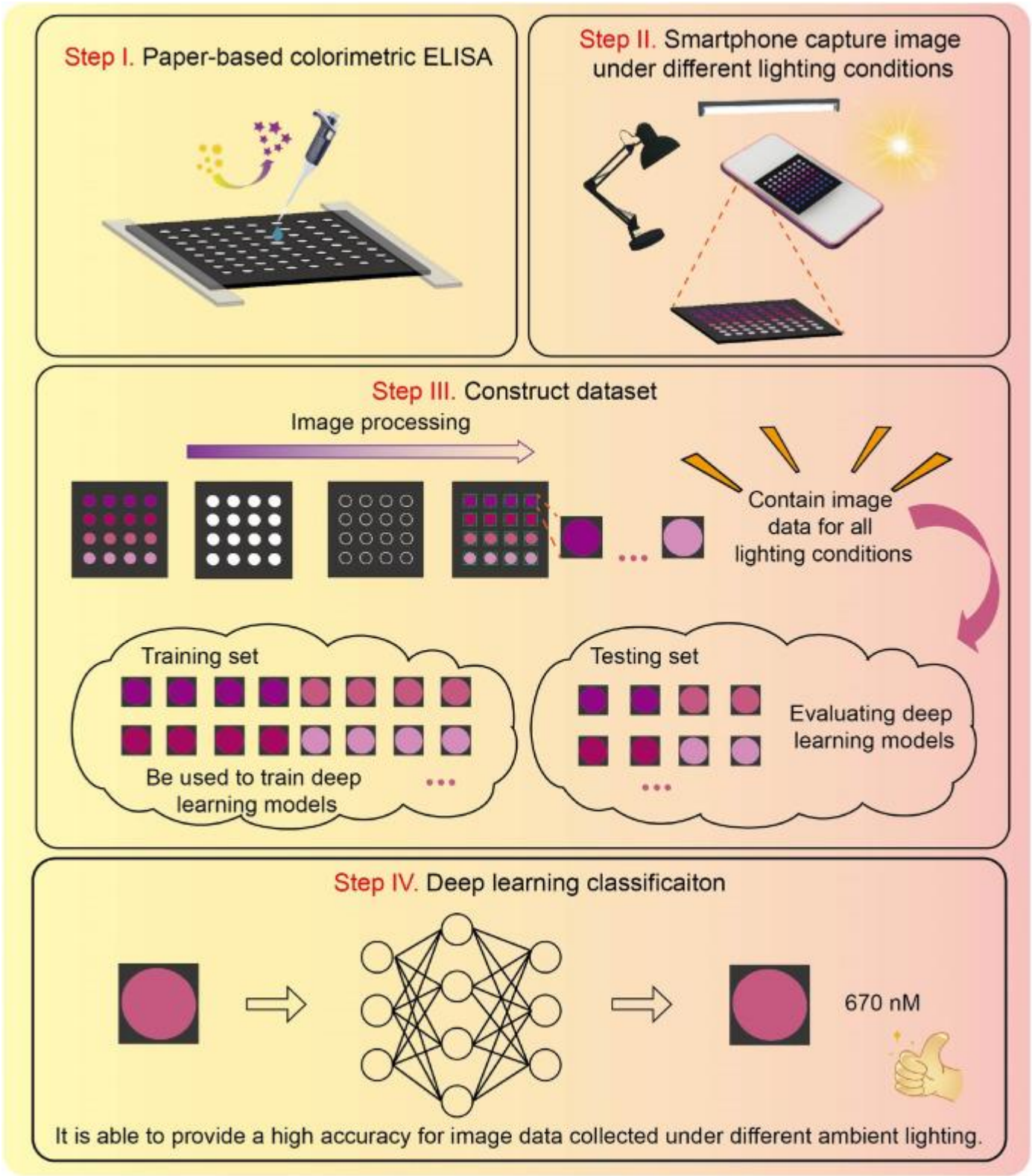

Fig. 1. Overview of deep learning assisted ultra-accurate smartphone testing of paper-based c-ELISA.

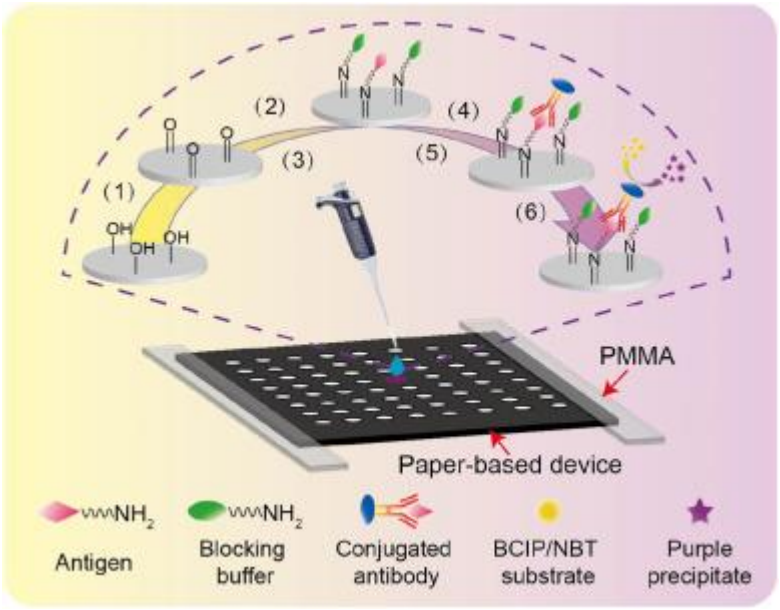

Fig. 2. Schematic diagram of a direct c-ELISA performed in a paper-based multi-well microplate. The protocol of a direct c-ELISA is carried out in six steps: (1) Modify the paper surface using periodate potassium; (2) Immobilize antigens; (3) Block each well with blocking buffer; (4) Immobilize enzyme-conjugated antibodies; (5) Add washing buffer; (6) Add enzyme substrate.

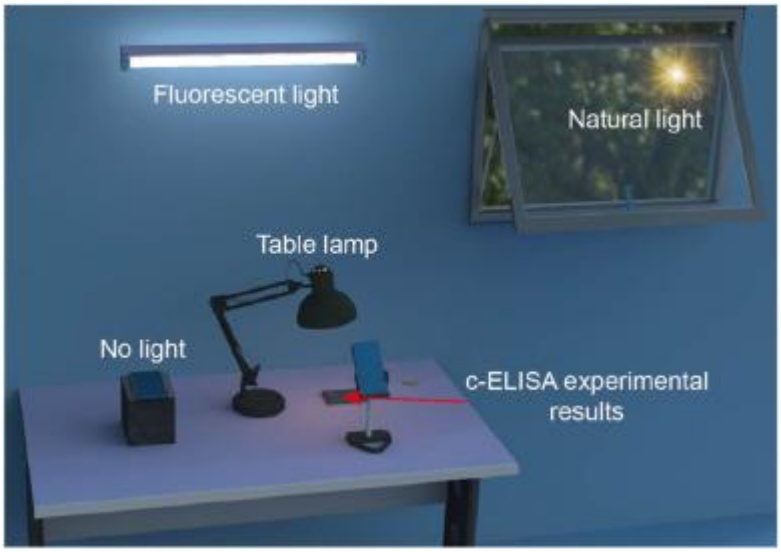

Fig. 3. Schematic illustration of an image captured with a smartphone camera under no-light conditions and different combinations of fluorescent light (F), table lamp (L), and natural light (N) conditions.

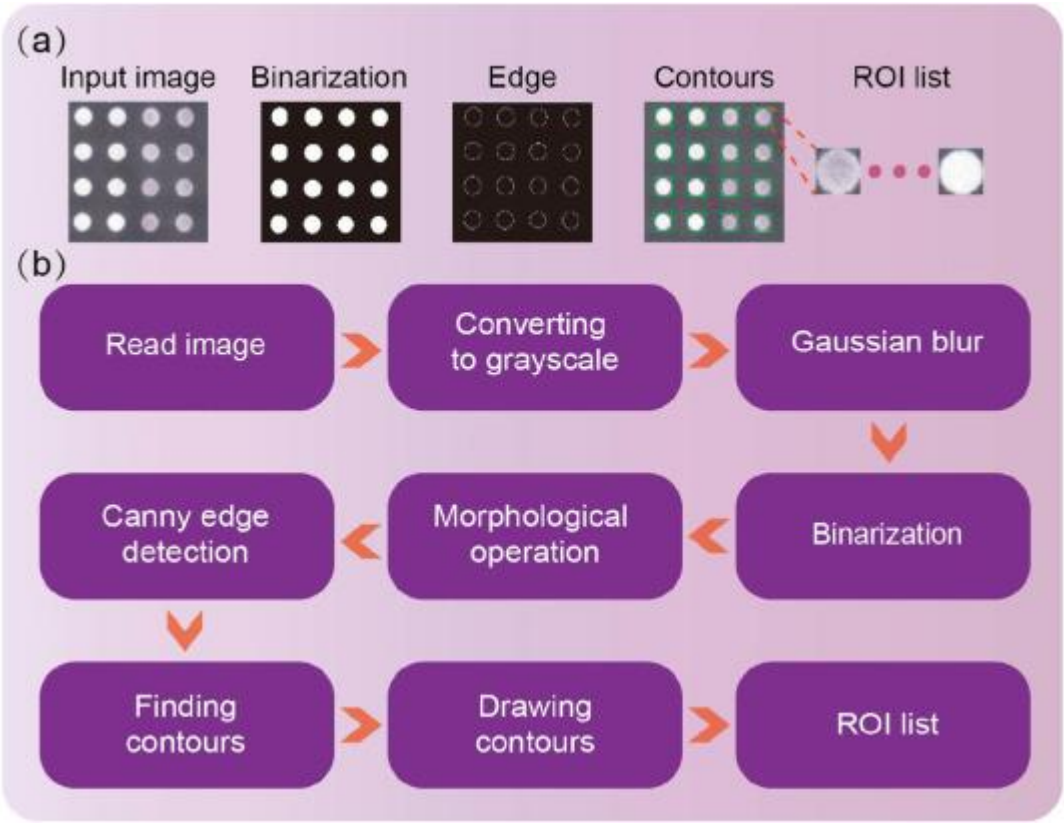

Fig. 4. Fully automated image processing. (a) An image with 10 nM IgG concentration and negative control is used as an example to illustrate the relevant image processing algorithm. (b) Flow chart of the image processing.

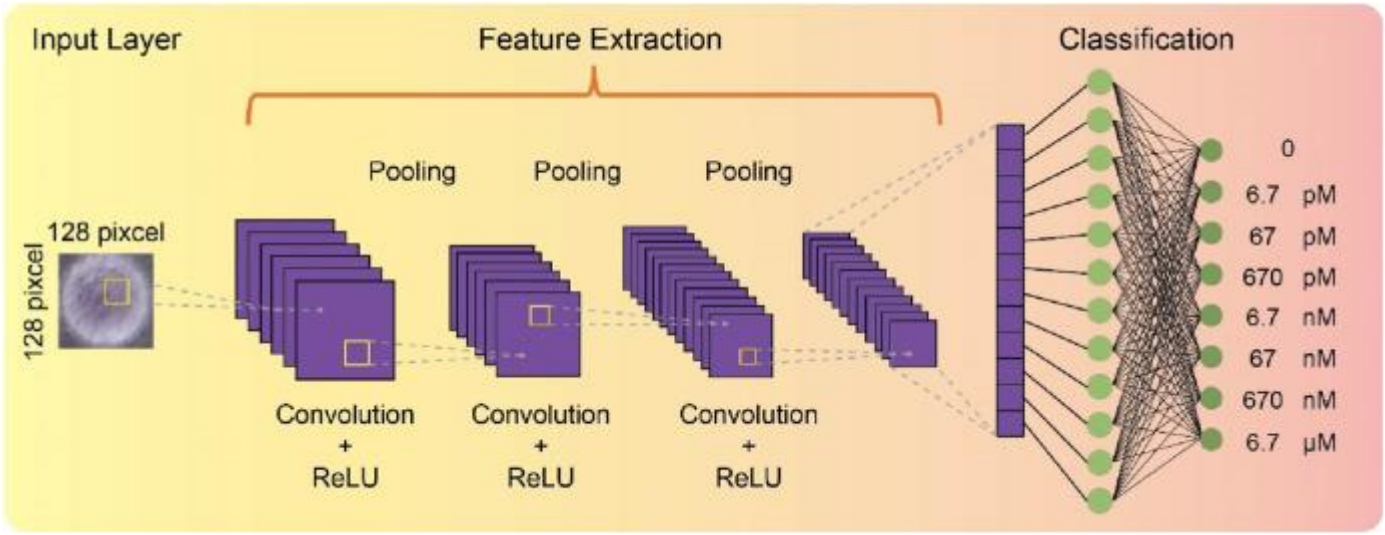

Fig. 5. The schematic of the deep learning architecture.

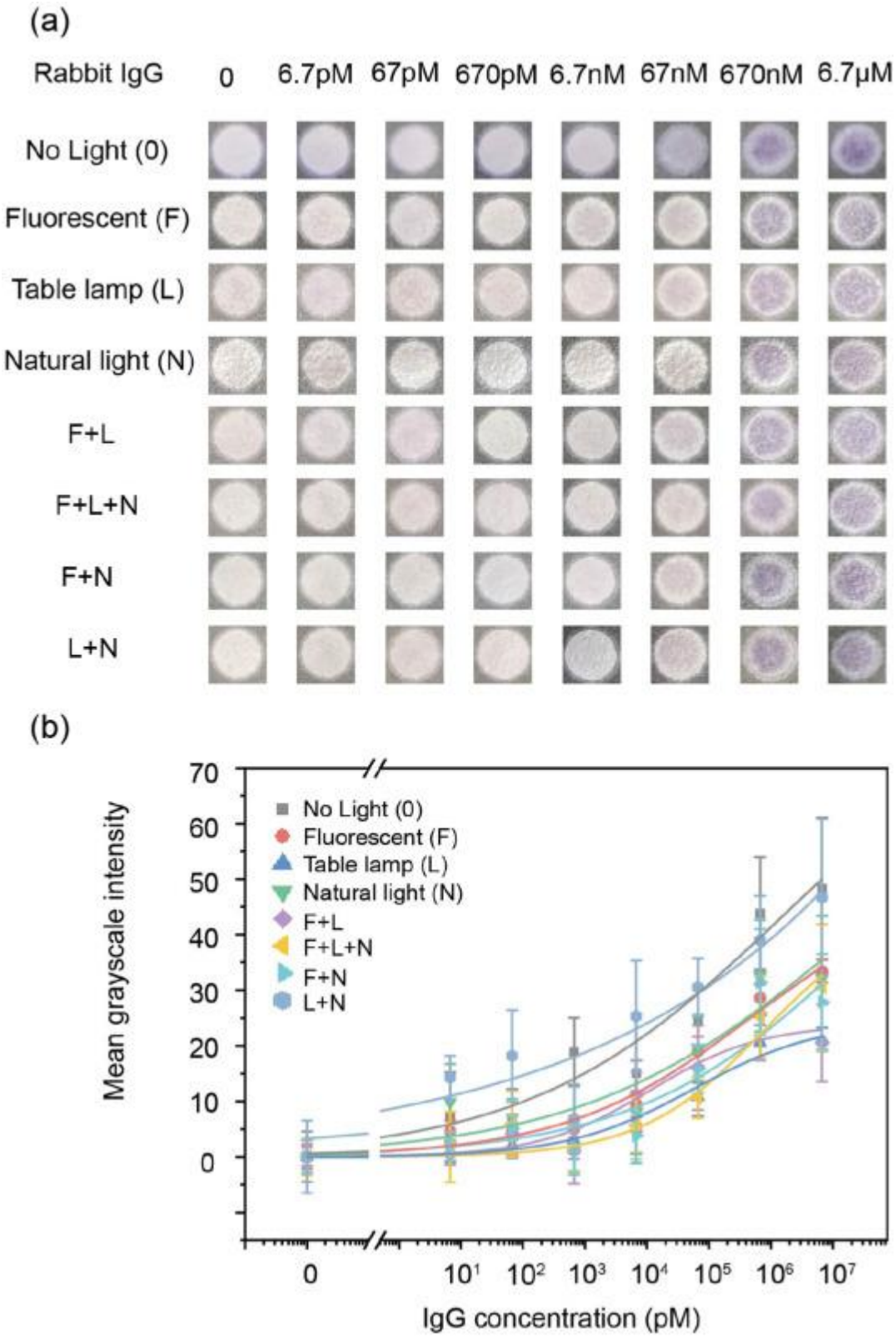

Fig. 6. Results of direct c-ELISA of rabbit IgG in PBS under different lighting conditions. (a) Photographs of the direct c-ELISA of rabbit IgG under four light conditions. (b) Calibration curve of the mean grayscale intensity versus the rabbit IgG concentration (N = 16).

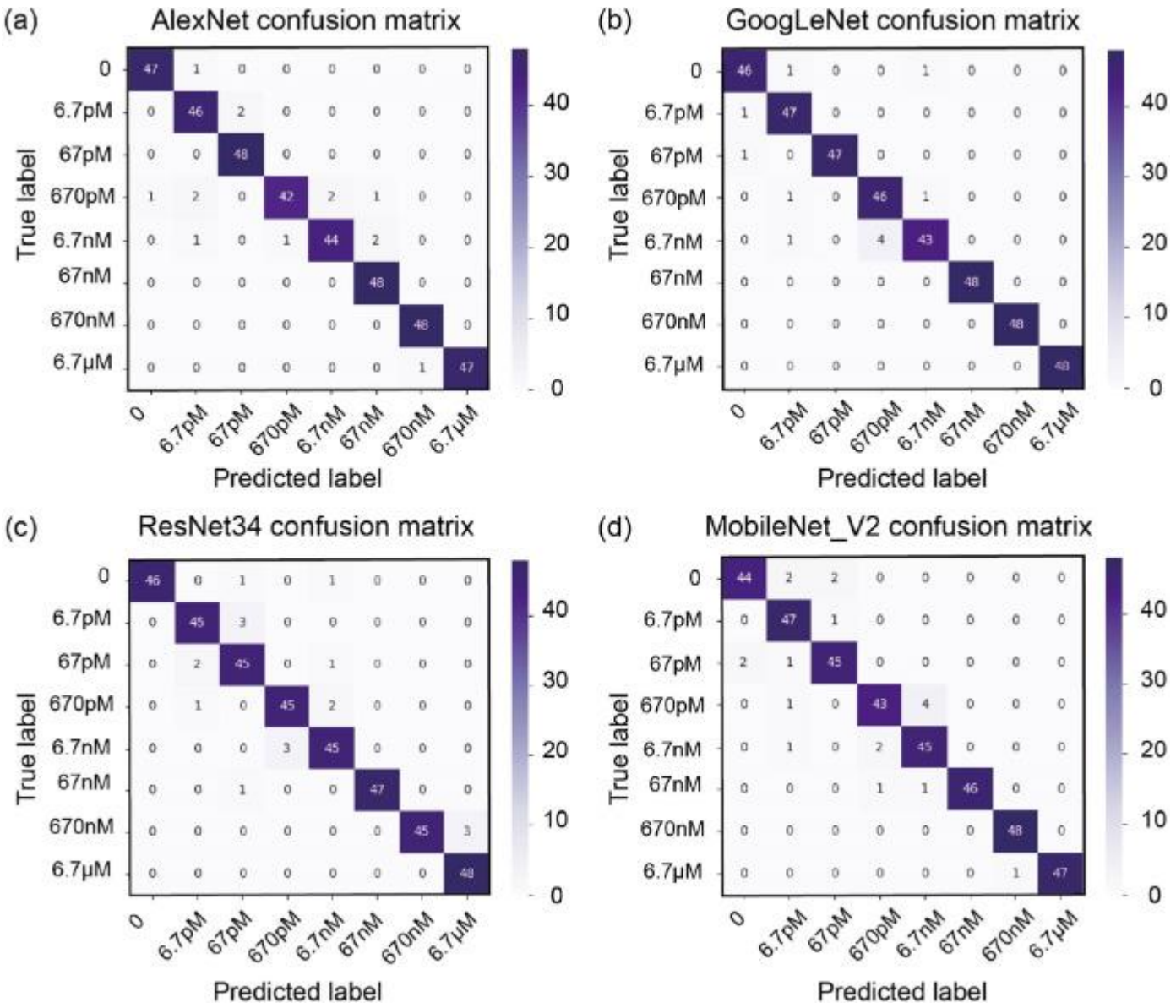

Fig. 7. Confusion matrices of (a) AlexNet, (b) GoogLeNet, (c) ResNet34, and (d) MobileNet_V2 algorithms. The horizontal coordinate of the confusion matrix represents the predicted Rabbit IgG concentration. The vertical coordinate represents the actual Rabbit IgG concentration. The darker the color in the confusion matrix, the higher the number of correct predictions (shown inside the square).

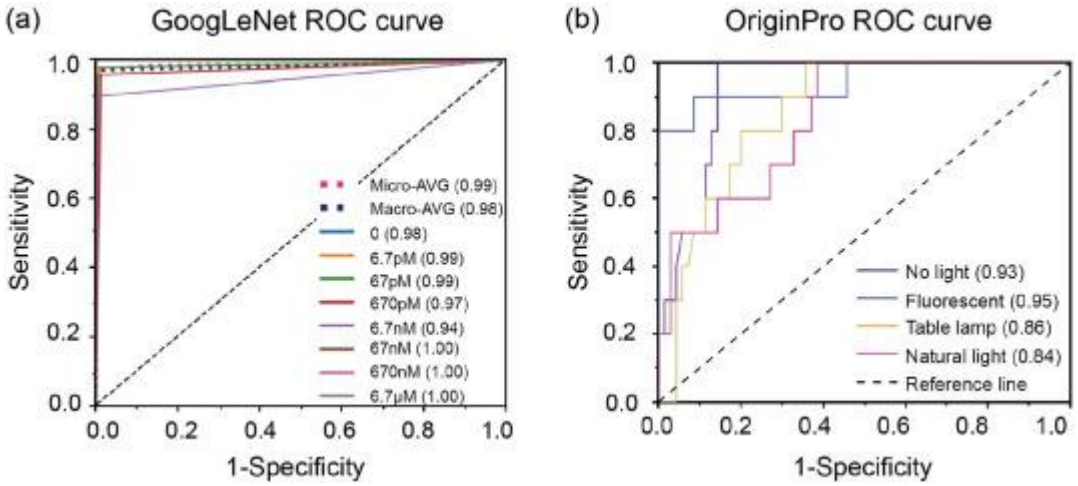

Fig. 8. Receiver operating characteristic (ROC) curve obtained by (a) GoogLeNet algorithm and (b) OriginPro.

References

[1] A.W. Martinez, S.T. Phillips, G.M. Whitesides, E. Carrilho, Diagnostics for the Developing World: Microfluidic Paper-Based Analytical Devices, ACS Publications, 2010.

[2] E. Noviana, T. Ozer, C.S. Carrell, J.S. Link, C. McMahon, I. Jang, C.S. Henry, Microfluidic paper-based analytical devices: from design to applications, Chem. Rev. 121 (2021) 11835–11885.

[3] Y. Xia, J. Si, Z. Li, Fabrication techniques for microfluidic paper-based analytical devices and their applications for biological testing: a review, Biosens. Bioelectron. 77 (2016) 774–789.

[4] T. Cai, S. Duan, H. Fu, J. Zhu, E.G. Lim, K. Huang, K. Hoettges, X. Liu, P. Song, A paper-based microfluidic analytical device with a highly integrated on-chip valve for autonomous ELISA, Proc. IEEE Int. Conf. Micro Electro Mech. Syst. (MEMS), IEEE (2022) 271–274.

[5] D.M. Cate, J.A. Adkins, J. Mettakoonpitak, C.S. Henry, Recent developments in paper-based microfluidic devices, Anal. Chem. 87 (2015) 19–41.

[6] Q. Shang, P. Zhang, H. Li, R. Liu, C. Zhang, A flow chemiluminescence paper-based microfluidic device for detection of chromium (III) in water, J. Innov. Opt. Health. Sci. 12 (2019), 1950016.

[7] J. Zhou, B. Li, A. Qi, Y. Shi, J. Qi, H. Xu, L. Chen, ZnSe quantum dot based ion imprinting technology for fluorescence detecting cadmium and lead ions on a three-dimensional rotary paper-based microfluidic chip, Sensor. Actuator. B Chem. 305 (2020), 127462.
[8] L. Cao, G.-C. Han, H. Xiao, Z. Chen, C. Fang, A novel 3D paper-based microfluidic electrochemical glucose biosensor based on rGO-TEPA/PB sensitive film, Anal. Chim. Acta 1096 (2020) 34–43.
[9] M. Taghizadeh-Behbahani, B. Hemmateenejad, M. Shamsipur, Colorimetric determination of acidity constant using a paper-based microfluidic analytical device, Chem. Pap. 72 (2018) 1239–1247.
[10] C.M. Cheng, A.W. Martinez, J. Gong, C.R. Mace, S.T. Phillips, E. Carrilho, K. A. Mirica, G.M. Whitesides, E.L.I.S.A. Paper-based, Angew. Chem. 122 (2010) 4881–4884.
[11] L.-M. Fu, Y.-N. Wang, Detection methods and applications of microfluidic paper-based analytical devices, Trends Anal. Chem. 107 (2018) 196–211.
[12] S. Balbach, N. Jiang, R. Moreddu, X. Dong, W. Kurz, C. Wang, J. Dong, Y. Yin, H. Butt, M. Brischwein, Smartphone-based colorimetric detection system for portable health tracking, Anal. Methods 13 (2021) 4361–4369.
[13] X. Huang, D. Xu, J. Chen, J. Liu, Y. Li, J. Song, X. Ma, J. Guo, Smartphone-based analytical biosensors, Analysis 143 (2018) 5339–5351.
[14] ¨O.B. Mercan, V. Kılıç, M. S¸ en, Machine learning-based colorimetric determination of glucose in artificial saliva with different reagents using a smartphone coupled µPAD, Sens. Actuator B-Chem. 329 (2021), 129037.
[15] M. Zhang, X. Cui, N. Li, Smartphone-based mobile biosensors for the point-of-care testing of human metabolites, Mater. Today Bio (2022), 100254.
[16] J.F. Bergua, R. Alvarez-Diduk, A. Idili, C. Parolo, M. Maym´o, L. Hu, A. Merkoci, Low-cost, user-friendly, all-integrated smartphone-based microplate reader for optical-based biological and chemical analyses, Anal. Chem. 94 (2022) 1271–1285.
[17] H. Kettler, K. White, S.J. Hawkes, Mapping the Landscape of Diagnostics for Sexually Transmitted Infections: Key Findings and Recommendations, World Health Organization, 2004.

[18] S.H. Baek, C. Park, J. Jeon, S. Park, Three-dimensional paper-based microfluidic analysis device for simultaneous detection of multiple biomarkers with a smartphone, Biosensors 10 (2020) 187.
[19] S. Cho, T. San Park, T.G. Nahapetian, J.-Y. Yoon, Smartphone-based, sensitive μPAD detection of urinary tract infection and gonorrhea, Biosens. Bioelectron. 74 (2015) 601–611.
[20]¨O. Kap, V. Kılıç, J.G. Hardy, N. Horzum, Smartphone-based colorimetric detection systems for glucose monitoring in the diagnosis and management of diabetes, Analysis 146 (2021) 2784–2806.
[21] A. Thiha, F. Ibrahim, A colorimetric enzyme-linked immunosorbent assay (ELISA) detection platform for a point-of-care dengue detection system on a lab-on-compact-disc, Sensors 15 (2015) 11431–11441.
[22] X. Liu, C. Cheng, A. Martinez, K. Mirica, X. Li, S. Phillips, M. Mascarenas, G. Whitesides, A portable microfluidic paper-based device for ELISA, in: Proc. IEEE Int. Conf. Micro Electro Mech. Syst. (MEMS), IEEE, 2011, pp. 75–78.
[23] R.C. Murdock, L. Shen, D.K. Griffin, N. Kelley-Loughnane, I. Papautsky, J.A. Hagen, Optimization of a paper-based ELISA for a human performance biomarker, Anal. Chem. 85 (2013) 11634–11642.
[24] A.W. Martinez, S.T. Phillips, E. Carrilho, S.W. Thomas III, H. Sindi, G. M. Whitesides, Simple telemedicine for developing regions: camera phones and paper-based microfluidic devices for real-time, off-site diagnosis, Anal. Chem. 80 (2008) 3699–3707.
[25] B. Coleman, C. Coarsey, M.A. Kabir, W. Asghar, Point-of-care colorimetric analysis through smartphone video, Sensor. Actuator. B Chem. 282 (2019) 225–231.
[26] B. Hunt, A.J. Ruiz, B.W. Pogue, Smartphone-based imaging systems for medical applications: a critical review, J. Biomed. Opt. 26 (2021), 040902.
[27] J.-H. Park, E.-K. Park, Y.K. Cho, I.-S. Shin, H. Lee, Normalizing the Optical Signal Enables Robust Assays with Lateral Flow Biosensors, ACS Omega, 2022.
[28] H. Fu, P. Song, Q. Wu, C. Zhao, P. Pan, X. Li, N.Y. Li-Jessen, X. Liu, A paper-based microfluidic platform with shape-memory-polymer-actuated fluid valves for automated multi-step immunoassays, Microsyst. Nanoeng. 5 (2019) 1–12.

[29] R. Liu, C. Ren, M. Fu, Z. Chu, J. Guo, Platelet detection based on improved YOLO_ v3, Cyborg and Bionic Systems 2022 (2022).
[30] G. Zhan, W. Wang, H. Sun, Y. Hou, L. Feng, Auto-CSC, A transfer learning based automatic cell segmentation and count framework, Cyborg and Bionic Systems 2022 (2022).
[31] Z.S. Ballard, H.-A. Joung, A. Goncharov, J. Liang, K. Nugroho, D. Di Carlo, O. B. Garner, A. Ozcan, Deep learning-enabled point-of-care sensing using multiplexed paper-based sensors, NPJ Digit. Med. 3 (2020) 1–8.
[32] X. Guo, M.A. Khalid, I. Domingos, A.L. Michala, M. Adriko, C. Rowel, D. Ajambo, A. Garrett, S. Kar, X. Yan, Smartphone-based DNA diagnostics for malaria detection using deep learning for local decision support and blockchain technology for security, Nat. Electron. 4 (2021) 615–624.
[33] A.Y. Mutlu, V. Kılıç, G.K.¨Ozdemir, A. Bayram, N. Horzum, M.E. Solmaz, Smartphone-based colorimetric detection via machine learning, Analysis 142 (2017) 2434–2441.
[34] M.H. Tania, K.T. Lwin, A.M. Shabut, M. Najlah, J. Chin, M.A. Hossain, Intelligent image-based colourimetric tests using machine learning framework for lateral flow assays, Expert Syst. Appl. 139 (2020), 112843.
[35] H.J. Min, H.A. Mina, A.J. Deering, E. Bae, Development of a smartphone-based lateral-flow imaging system using machine-learning classifiers for detection of Salmonella spp, J. Microbiol. Methods 188 (2021), 106288.
[36] M.E. Solmaz, A.Y. Mutlu, G. Alankus, V. Kılıç, A. Bayram, N. Horzum, Quantifying colorimetric tests using a smartphone app based on machine learning classifiers, Sensor. Actuator. B Chem. 255 (2018) 1967–1973.
[37] H. Kim, O. Awofeso, S. Choi, Y. Jung, E. Bae, Colorimetric analysis of saliva–alcohol test strips by smartphone-based instruments using machine-learning algorithms, J. Appl. Opt. 56 (2017) 84–92.
[38] Y. Liu, L. Zhan, Z. Qin, J. Sackrison, J.C. Bischof, Ultrasensitive and highly specific lateral flow assays for point-of-care diagnosis, ACS Nano 15 (2021) 3593–3611.
[39] V. Do˘gan, E. Yüzer, V. Kılıç, M. S¸ en, Non-enzymatic colorimetric detection of hydrogen peroxide using a μPAD coupled with a machine learning-based

smartphone app, Analysis 146 (2021) 7336–7344.
[40]¨O.B. Mercan, V. Kılıç, M. S¸ en, Machine learning-based colorimetric determination of glucose in artificial saliva with different reagents using a smartphone coupled µPAD, Sensor. Actuator. B Chem. 329 (2021), 129037.
[41] S.K. Biswas, S. Chatterjee, S. Bandyopadhyay, S. Kar, N.K. Som, S. Saha, S. Chakraborty, Smartphone-enabled paper-based hemoglobin sensor for extreme point-of-care diagnostics, ACS Sens. 6 (2021) 1077–1085.
[42] H.-A. Joung, Z.S. Ballard, J. Wu, D.K. Tseng, H. Teshome, L. Zhang, E.J. Horn, P. M. Arnaboldi, R.J. Dattwyler, O.B. Garner, Point-of-care serodiagnostic test for early-stage Lyme disease using a multiplexed paper-based immunoassay and machine learning, ACS Nano 14 (2019) 229–240.
[43] Q. Ning, W. Zheng, H. Xu, A. Zhu, T. Li, Y. Cheng, S. Feng, L. Wang, D. Cui, K. Wang, Rapid segmentation and sensitive analysis of CRP with paper-based microfluidic device using machine learning, Anal. Bioanal. Chem. 414 (2022) 3959–3970.
[44] H. Fu, X. Liu, Experimental comparison of surface chemistries for biomolecule immobilization on paper-based microfluidic devices, J. Micromech. Microeng. 29 (2019), 124003.
[45] E. Gündo˘gdu, A.C. Kunduraci, Effect of window glazings' visible transmittance to daylight factor and energy efficiency in an architecture studio, J. Emerg. Trends Eng. Appl. Sci. 10 (2019) 171–178.
[46] M. Hoque Tania, An Intelligent Image-Based Colourimetric Test Framework for Diagnosis, Anglia Ruskin University, 2018.
[47] S.G. Wu, F.S. Bao, E.Y. Xu, Y.-X. Wang, Y.-F. Chang, Q.-L. Xiang, A leaf recognition algorithm for plant classification using probabilistic neural network, in: 2007 IEEE International Symposium on Signal Processing and Information Technology, IEEE, 2007, pp. 11–16.
[48] I. Pratikakis, K. Zagoris, G. Barlas, B. Gatos, ICDAR2017 Competition on Document Image Binarization (DIBCO 2017), 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), IEEE, 2017, pp. 1395–1403.

Analytica Chimica Acta 1248 (2023) 340868