

Convolution Neural Networks for Point-of-Care Diagnostics of Bacterial Infections in Blood

Omkar Hegde^{1,†}, Ritika Chatterjee^{2,†}, Durbar Roy^{1,+}, Vivek Jaiswal^{1,+}, Dipshikha Chakravorty^{2,*}, Saptarshi Basu^{1,*}

¹Department of Mechanical Engineering, Indian Institute of Science, Bangalore

²Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore

[†]These authors contributed equally as first authors

⁺These authors contributed equally as second authors

*Corresponding Author¹ email – sbasu@iisc.ac.in

*Corresponding Author² email – dipa@iisc.ac.in

ABSTRACT:

A droplet of blood, when evaporated on a surface, leaves dried residue—the fractal patterns formed on the dried residues can act as markers for infection present in the blood. Exploiting the unique patterns found in the residues of a naturally dried droplet of blood, we propose a Point-of-Care (POC) diagnostic tool for detecting broad-spectrum of bacterial infections (such as *Enterobacter aerogenes*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Pseudomonas aeruginosa*, *Salmonella Typhi*) in blood. The diagnosis process we propose is straightforward and can be performed with the following steps: A droplet of blood (healthy or infected) of volume range 0.5 to 2 μl is allowed to dry on a clean glass surface and is imaged using a conventional optical microscope. A computer algorithm based on the framework of convolution neural network (CNN) is used to classify the captured images of dried blood droplets according to the bacterial infection. In total, our multiclass model reports an accuracy of 92% for detecting six bacterial species infections in the blood (with control being the uninfected or healthy blood). The high accuracy of detecting bacteria in the blood reported in this article is commensurate with the standard bacteriological tests. Thus, this article presents a proof-of-concept of a potential futuristic tool for a rapid and low-cost diagnosis of bacterial infection in the blood.

INTRODUCTION:

Antimicrobial resistance (AMR) has become a cause of global concern, aggravated by a lack of data on the different mechanisms and factors causing its emergence and spread. According to a WHO report, bacterial infections alone cause nearly 1 million deaths per year worldwide; this number is estimated to increase to 10 million per year by 2050, owing to AMR¹. Entero-pathogens (such as *Salmonella Typhi*) and ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter aerogenes*) pathogens are the deadliest, and every passing year they cause millions of infections and thousands of deaths worldwide²⁻⁶. These bacteria are clinically relevant strains causing ~42% bloodstream infection and have been associated with significant mortality and morbidity^{7,8}. The WHO report lists these pathogens as priority antibiotic-resistant pathogens as they are associated with a significant amount of mortality and morbidity due to the emergence of multiple drug resistance (MDR) and extensive drug resistance (XDR) strain⁶. Even developed economies like the United States⁹ and other European countries¹⁰ are finding it difficult to cope with the increasing AMR; an estimated cost of billions of dollars is attributed to the health care expenditure of hospital-acquired and community-acquired AMR infections¹¹. The first step to stop the spreading of infection would

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

47 be to diagnose it. Timely diagnosis of the causative organism and subsequent treatment of
48 bacterial infection can save millions of lives and inhibit the generation of various drug-
49 resistant strains. “Test, isolate (the person infected), and treat” is the primary formula used by
50 advanced countries to stop the spread of infection, as evidenced during the Global COVID-19
51 pandemic¹¹. Thus, it is crucial to test rapidly to prevent the further spread of infection and any
52 possible outbreak.

53 Since the abovementioned bacterial strains (Enteropathogens and ESKAPE pathogens)
54 majorly cause bloodstream infections^{12,13} and can mainspring changes in the composition and
55 fluid properties of the blood, it is prudent to test blood for diagnosis of these infections.
56 Moreover, blood profile indicates the overall health of the individuals; thus, it is a general
57 practice to test blood for several abnormalities, including bacterial infections.
58 Conventionally, various tests and investigations are undertaken to diagnose diseases in
59 blood—such as a complete blood count, chemistry profiling of urea and electrolytes, level of
60 C-Reactive Protein (CRP), glucose, and coagulation screen, to name a few. Several
61 biomarkers have been reported in the literature to diagnose sepsis¹⁴, such as pro-
62 inflammatory and anti-inflammatory cytokines and chemokines, altered cell surface marker
63 of immune cells. One of the main drawbacks of these markers is that they have low
64 specificity and high cost associated with high-end technique requirements for the detection¹⁵.
65 The gold standard for bacterial infection diagnosis is bacteriological culture and staining¹⁶.
66 Infectious agents are isolated and identified using various selective and differential culture
67 media. However, bacteriological culture diagnosis of the infection is often time-consuming
68 (require up to 48 hours). Alternatively, polymerase chain reaction (PCR) based diagnostics
69 are faster than standard bacteriological culture-based methods. Nonetheless, due to high
70 sensitivity, it often can give false-positive results. Another major disadvantage for PCR-based
71 detection involves too high cost and specialized instrumentation and setup requirement,
72 which may not be feasible in semi-urban or rural areas^{17–20}. Therefore, there is an urgent need
73 for innovative solutions to create accurate, low-cost, rapid diagnostic tools to detect bacterial
74 infections in the blood. This article addresses the above problems by providing a simple
75 method to rapidly detect a broad spectrum of bacterial infections in the blood (with a
76 turnaround time of less than 15 minutes) by leveraging the unique patterns formed on the
77 residue of the dried blood droplets.

78 With the advent of fundamental research on droplet evaporation and colloidal self-
79 assembly in droplets in the past few decades, the dried precipitates of biological fluids have
80 been deeply studied^{21–30}. A few of these studies have found applications in biomedical
81 diagnostics^{31–34}. The competitive interaction of multi-component elements in biological fluid
82 like blood leads to the formation of rich patterns on the dried droplet precipitate³⁵ such as
83 distinctive radial, spiral, and tangential cracks, color variations, varying thickness of the
84 coffee rings, finger formations, etc. (Refer to image S1 in the supplementary information for
85 the image of a dried droplet of blood taken from a healthy volunteer). These features depend
86 on the properties of the fluid (and thereby the components present in it), which in turn can be
87 related to the health of a person. Notably, in blood droplets, it has been shown previously that
88 the flow inside the droplet is driven by Marangoni, wettability, and natural evaporation leads
89 to the formation of distinctive patterns on the dried residues of the drop for anemic and
90 hyperlipidemic blood samples^{36–38}. Besides, the final dried blood droplet pattern is
91 susceptible to drying conditions such as temperature³⁹ and humidity⁴⁰, placement of blood
92 droplet⁴¹ (gentle or with impact), and depends on several other factors such as blood groups⁴²,
93 ethnicity⁴³, age⁴⁴, gender⁴³, size, shape, and mobility of cells⁴⁵. Further, blood infection leads
94 to cellular changes such as Leucocytosis⁶³ or leukopenia, and biochemical changes include
95 markers released upon red blood cell lysis or white blood cell lysis⁶³. Thus, the blood

96 carrying the infection would also exhibit different properties (compared to the healthy blood),
97 leading to a unique pattern formation on the dried blood drop⁴⁶; for example, it is possible to
98 detect malaria in dried blood droplets using a simple light microscope imaging⁴⁷.

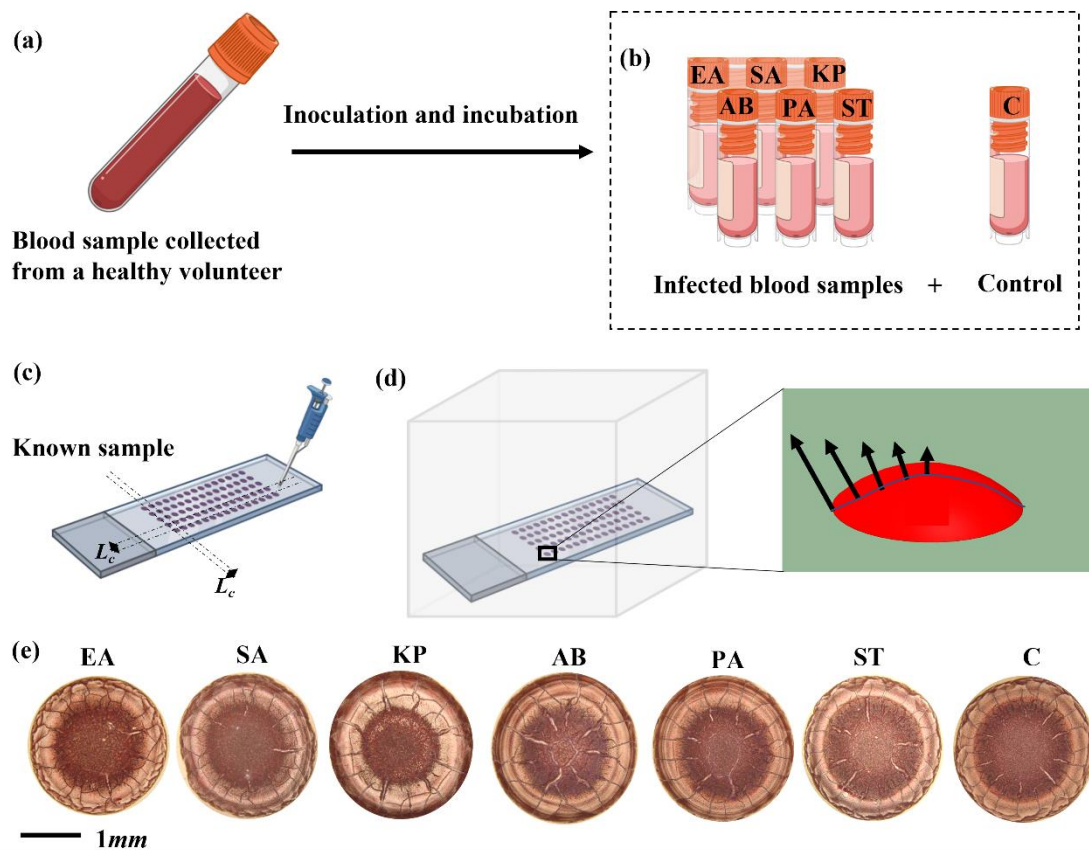
99 However, the differences in the final pattern on the blood droplet residue could be due to
100 several reasons, and markers for an infection in these patterns can be subtle. Thus, the human
101 interpretation of optical images of dried blood droplets for disease diagnosis can be prone to
102 errors; it is inevitable to use computer-assisted diagnosis. Also, the use of computers for the
103 diagnosis will lead to a rapid and more accurate diagnosis. Bio-Medical diagnostics have
104 extensively used Deep learning (DL) and Convolutional Neural networks (CNN). For
105 example, with the help of images generated by CT scans, breast cancer, lung nodule, and
106 segmentation can be detected using CNN and DL¹⁹⁻²⁷. These computer models reported 63-
107 85% accuracy (AlexNet 63.98%²⁸, DenseNet 80.7%²⁹, and Wang et al. 85%³⁰), but the
108 training data set had very few runs (less than five runs). Recently, Machine Learning (ML)
109 analysis was used to discriminate dried blood droplet patterns with varying physiological
110 conditions, i.e., changes in dried blood droplet patterns before or after physical exercise of the
111 volunteers were detected using ML with a 95% prediction accuracy⁴⁸. This demonstrates that
112 ML is a potent tool to classify patterns formed by blood droplets with even slight changes in
113 blood chemistry induced by the physical exercise of volunteers. However, it remains unclear
114 if the ML or DL can detect a wide variety of bacterial pathogens in the blood.

115 The current article presents a proof-of-concept of an alternative low-cost and rapid point-of-
116 care diagnostic tool by imaging a dried droplet of blood using simple optical microscopy. It
117 has been demonstrated that a wide range of deadly pathogens such as (*Enterobacter*
118 *aerogenes*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and
119 *Pseudomonas aeruginosa*, *Salmonella Typhi*) present in the blood can be detected using the
120 framework of CNN. The diagnostic procedure is described as follows: An image of a dried
121 blood droplet (that may contain the target infection) on a glass substrate is captured using an
122 optical microscope. The captured image is fed into an in-house developed computer
123 algorithm that predicts the possible infection based on CNN architecture. Our model reports
124 an accuracy of 92% for distinguishing the infection of the abovementioned bacterial species
125 in the blood. The computer algorithm for this application is developed by training and testing
126 the algorithm in a supervised learning environment with a large number of images of blood
127 droplets (~35,000 images), whose status of infection is known apriori. The model accuracy
128 remains invariant over the large dataset.

129 **RESULTS AND DISCUSSION:**

130 Blood collection from the volunteers is detailed in the experimental procedure section.
131 The blood is spiked with bacteria *in-vitro* (individual bacterial species are spiked in different
132 vials of blood (Fig.1 (a) and (b)) - *Enterobacter aerogenes* (EA), *Staphylococcus aureus*
133 (SA), *Klebsiella pneumoniae* (KP), *Acinetobacter baumannii* (AB), *Pseudomonas aeruginosa*
134 (PA)), *Salmonella Typhi* (ST)). The unspiked blood sample is used as a control (C) against
135 the abovementioned spiked samples, i.e., it does not contain any bacteria and is considered a
136 healthy blood sample. These samples are incubated for 24h at 37°C. Each blood sample in
137 the vials (including the control) is considered an individual class and labeled accordingly (see
138 Fig. 1(b)). Several hundred droplets (of volume 0.5 to 2 µl) of blood samples
139 (spiked/unspiked) are deposited on a clean glass slide on a given day and allowed to
140 evaporate in the controlled atmospheric condition (26 ±5 °C and 45 ±10% RH) (Fig.1 (c) and

141 (d)). The droplets are imaged using a conventional optical microscope—a total of 35000
142 images are acquired (number of images in each class (~5000 images of each class). Sample
143 images of the dried blood droplet of the glass slide spiked with different bacteria, and the
144 control (unspiked) is shown in Fig.1 (e). The blood droplet evaporates in constant contact
145 radius (CCR) mode as the droplet remains pinned throughout evaporation (irrespective of any
146 class). The droplets generally assume the shape of a spherical cap as the Bond number (Bo) is
147 $\ll 1$ and the dried residue is of a circular shape. However, few droplets show a shape
148 disparity due to errors in dispensing the drop from the micropipette. Thus, amongst the nearly
149 47,000 images taken, about 13,000 images are discarded to disregard the shape disparity of
150 the droplet. The analysis in this article is based on 34469 images considered after filtration. It
151 is evident from Fig.1 (e) that it is difficult to classify the blood droplet residues according to
152 the bacteria present in them from a visual inspection of the images. Hence, we have
153 developed an image classification computer algorithm using CNN to recognize different
154 classes.

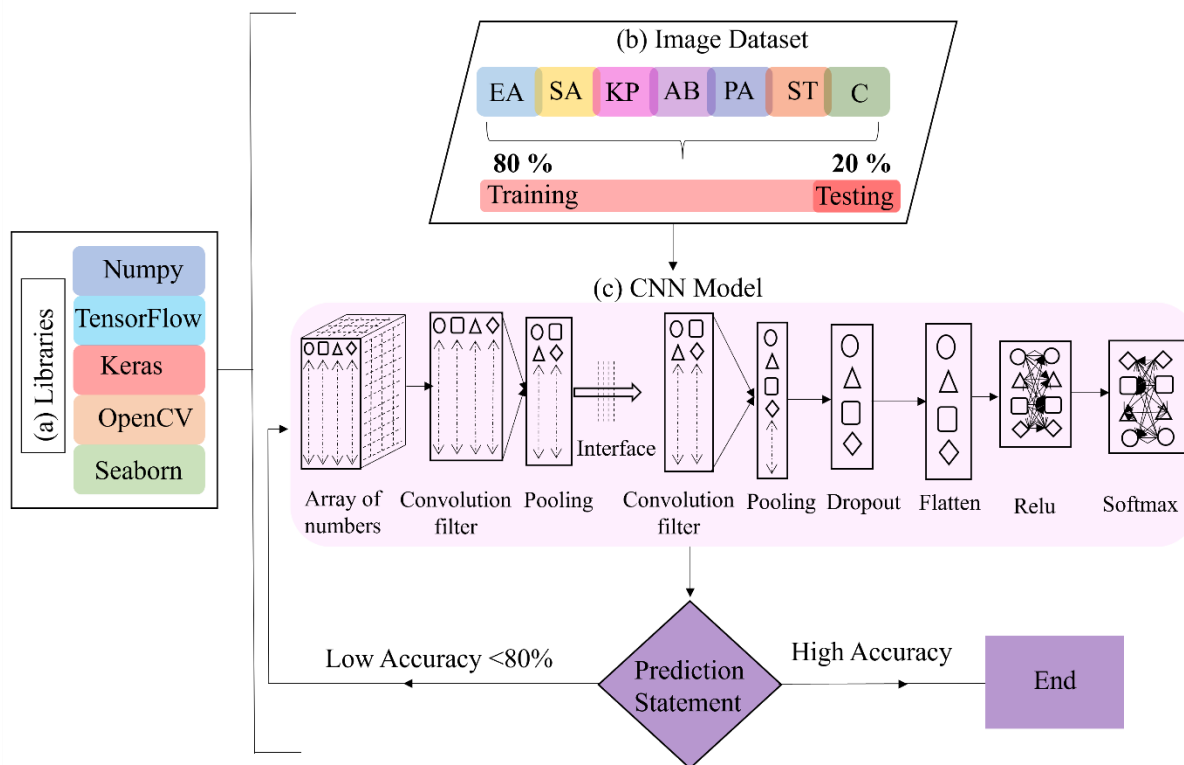


155

156 **Figure 1. Evaporation triggered pattern formation in infected and healthy blood**
157 **droplets. (a) 15 ml of blood is collected from a healthy volunteer in vacutainer tubes. (b)**
158 **the collected blood is infected with bacteria (in-vitro) and is labeled as follows:**
159 ***Enterobacter aerogenes* - EA, *Staphylococcus aureus* - SA, *Klebsiella pneumoniae* - KP,**
160 ***Acinetobacter baumannii* - AB, *Pseudomonas aeruginosa* - PA, *Salmonella Typhi* -ST.**
161 **The blood sample that does not contain any infection is considered as a control and is**
162 **labeled as - C . All the samples are incubated for 24h at 37° C. The individual labels are**
163 **referred to as “class”. (c) A few hundreds of droplets of 0.5-2 μ l of a known sample (as**
164 **labeled in (b)) are gently placed on clean glass slides using a micro-pipette. The**

165 individual droplets are placed far apart (L_c) such that there is no vapor-mediated
 166 interaction between them. (d) The deposited droplets are allowed to evaporate in an
 167 enclosed chamber. The evaporation profile is shown in the zoomed-in image of the
 168 droplet, with the evaporation flux being maximum at the edge and minimum at the
 169 center of the drop. (e) Sample images of dried blood droplets labeled according to their
 170 class (the respective label is on top of the image). The label at the top of the image
 171 indicates the species of bacterial infection in the blood. The last image in the row is not
 172 infected with bacteria and thus is the control.

173 The algorithm is developed in Python programming language and uses seven libraries to
 174 analyze the data set: NumPy⁴⁹, Keras⁵⁰, Matplotlib⁵¹, Seaborn⁵², Scikit⁵³, Tensorflow⁵⁴, and
 175 OpenCV⁵⁵ (Refer to Fig. 2(a), Matplotlib, Seaborn are for plotting the data). The data set
 176 (captured images of blood droplets) is imported and labeled according to the given class.
 177 Prior to analysis, the image data set is segregated in the ratio of 80:20 for training and testing,
 178 respectively (see Fig. 2(b)). This segregation is done randomly to avoid any bias. The images
 179 are further pre-processed using a series of operations using the ImageDataGenerator function
 180 in the Keras library to maintain uniformity for variables like contrast, color, and positioning
 181 of the drop in the image. The processed images are discretized and equalized into an array of
 182 numbers (vector). However, analyzing this data on traditional computing infrastructure (the
 183 hardware used is Intel(R) Core i7 2.90 GHz 64 GB RAM) is a challenge as the volume of
 184 data generated by the images is very large.

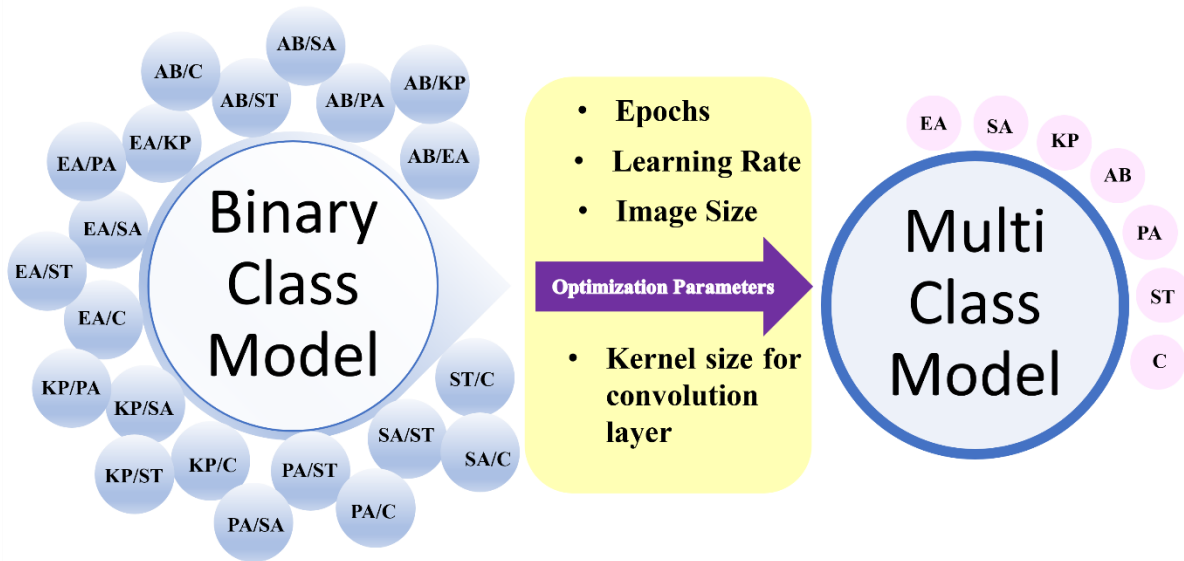


185
 186 **Figure 2. The overall design of the CNN algorithm used to train the computer to classify**
 187 **images of the residues of dried blood droplets according to their classes. (a) The**
 188 **algorithm imports standard libraries used for CNN. Importantly, NumPy⁴⁹, Keras⁵⁰,**
 189 **Tensorflow⁵⁴, OpenCV⁵⁵, and Seaborn⁵² are imported libraries, to name a few. (b) The**
 190 **imported image data set consisting of ~35000 images are labeled as per their respective**

191 **classes. The image dataset is divided into the training and testing sets in the ratio 80:20,**
192 **respectively. The model is trained using the training data set and is validated using the**
193 **test dataset. The test dataset is not used during training, thus ensuring no data leakage**
194 **from the test dataset. (c) The image data set is converted to a number array (vector) and**
195 **analyzed using multiple layers. Model parameters such as epochs, kernel size for**
196 **convolutional layers, image size, and learning rate are fed into the model, and the model**
197 **is trained using the training set. The model makes a prediction using the testing dataset.**
198 **Iterations are repeated with a new set of values for model parameters until the highest**
199 **possible accuracy is achieved.**

200 Hence the algorithm is developed in a two-step process. First, the model is optimized
201 for binary classes, i.e., the algorithm is trained to distinguish between any two classes with
202 high accuracy. Since there are seven classes, the model for 7C_2 binary classes is optimized
203 (see Fig. 3). Second, the parameters used for optimizing the binary classes (such as epochs -
204 the number of times the algorithm iterates the whole dataset for training, kernel size for
205 convolutional layers⁵¹, image size, and learning rate - the variation in the model with updated
206 weights) are used in the multiclass model with all the classes included, as shown in Fig. 3 (7
207 classes including the control case). While the basic structure of the algorithm remains the
208 same, the parameters optimized for binary classes are used as a template for the multiclass
209 model with all the classes.

210 To examine the algorithm for its ability to distinguish between the binary classes, the
211 vector output of the processed images from any two classes is input into the CNN Model. The
212 CNN model consists of the following layers: convolution layer⁵¹, pooling layer⁵⁶, dropout
213 layer⁵², flattening layer⁵³, and fully connected layer⁵⁴ (Refer to Figure 2(c)). The modeling
214 parameters such as filters used and kernel size in the convolution layer, the prevention of
215 overfitting the data in the dropout layer, and units in the dense layer⁵⁰ have a crucial role in
216 improving the overall accuracy of detection. The values of epochs, learning rate, kernel and
217 image size for which the highest accuracy of prediction is achieved are considered optimized
218 values. Using the trial and error method, the optimization of the model (for the given vector)
219 is achieved by setting the values of epochs, learning rate, kernel size, and image size - 50, 10⁻
220 ⁴, 9×9, and 256 pixels, respectively. The trained model is tested using the test image data set
221 (20 % of the segregated image data set for testing) and the results of accuracy, precision, and
222 recall for 21 binary classes are given in table 1.



223

224 **Figure 3. Schematic representation of the two-step development of the model used in the**
 225 **article. Optimal model parameters such as the epoch, learning rate, image size, and**
 226 **kernel size for the convolution layer obtained from the 7C_2 binary classes are adopted**
 227 **into the multiclass model consisting of seven classes.**

228

Table 1: Accuracy, precision, and recall values for 7C_2 Binary classes

Binary class	Accuracy	Precision	Recall
AB/ EA	0.97	0.97	0.96
AB/KP	0.84	0.96	0.92
AB/PA	0.89	0.96	0.95
AB/SA	0.91	0.95	0.93
AB/ST	0.92	0.99	0.98
AB/C	0.85	0.91	0.92
EA/ KP	0.98	0.98	0.98
EA/PA	0.88	0.95	0.96
EA/SA	0.92	0.92	0.92
EA/ST	0.94	0.98	0.97
EA/C	0.98	0.98	0.98
KP/PA	0.92	0.93	0.92
KP/SA	0.91	0.9	0.92
KP/ST	0.93	0.94	0.94
KP/C	0.80	0.85	0.81
PA/SA	0.96	0.96	0.96
PA/ST	0.92	0.98	0.97
PA/C	0.92	0.94	0.93
SA/ST	0.98	0.98	0.98
SA/C	0.88	0.9	0.88
ST/C	0.92	0.93	0.92

229 The performance of the model is gauged importantly by the following output parameters:
 230 accuracy, loss, precision, and recall. The accuracy of the model is the ratio of correctly

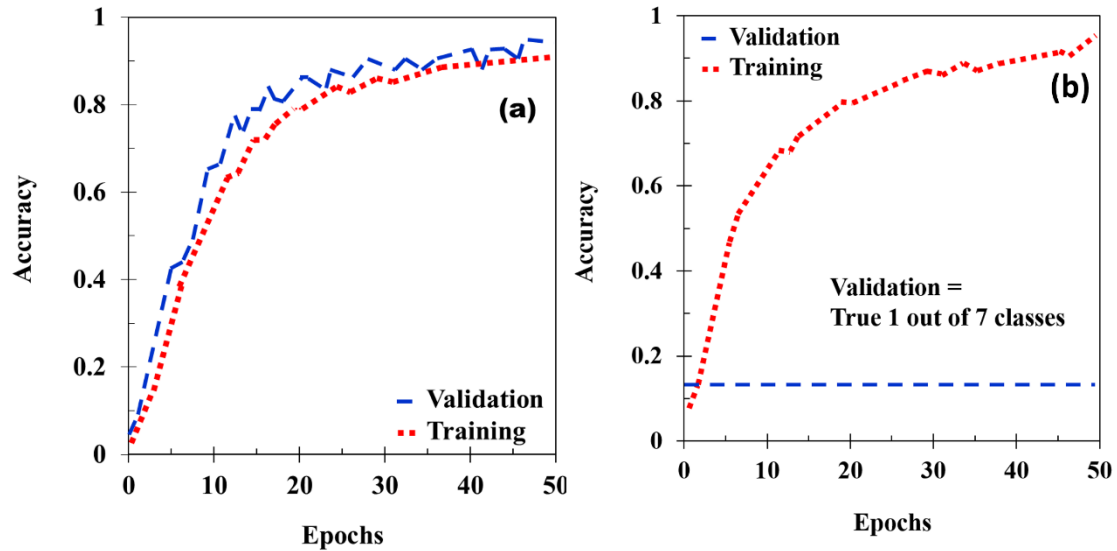
231 predicted to the total samples. The loss is the summation of errors made in prediction by the
232 model on the training and test image datasets (see Fig. 5 (b)). The precision is the percentage
233 of actual positive samples to the predicted positive samples (which may be correct or
234 incorrect). Recall (sensitivity) is the ratio of correctly predicted positive samples out of total
235 positive samples. As shown in Table 1 (for binary classes), the accuracy, precision, and recall
236 range from 80- 98% to 85-98% and 81-98%, respectively. The model parameters used for the
237 binary classes are considered to be optimized as the accuracy achieved for binary classes is
238 maximum for the given image data set. The optimized parameters used by the model for
239 binary classes are introduced into the multiclass model containing all the classes. The
240 algorithm is iterated for the multiclass model with the optimized parameters.

241

242 **Table 2: Accuracy, precision, and recall values of a multiclass model consisting of 7**
243 **classes**

Class	Accuracy	Precision	Recall
EA	0.92	0.97	0.96
SA		0.94	0.99
KP		0.96	0.9
AB		0.91	0.88
PA		0.91	0.91
ST		0.95	0.99
C		0.81	0.98

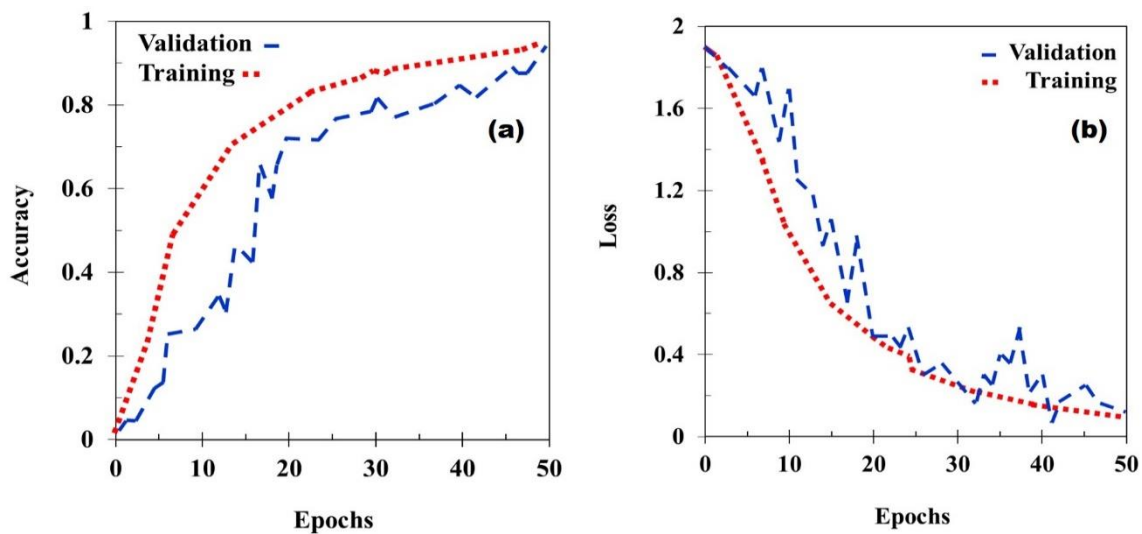
244 Since the model is trained with the training image dataset, it (the training image
245 dataset) should fit the model accurately upon validation. Fig. 4 (a) shows that this model has
246 97% accuracy when validated with the training image dataset. Subsequently, if the training
247 image dataset of only one of the classes fed to the algorithm is correct, the accuracy of the
248 model is stunted to 1/7 (refer to Fig. 4(b)). Thus, the model predicts appropriately for the
249 trained image dataset. The test image data set is unknown to the trained model. Thus, the
250 model is validated using the test image data set to mimic a real-life scenario. Upon validation,
251 the precision and recall values range from 88 to 99% and 81 to 97%, respectively, for
252 different classes (see table 2). The overall accuracy of the multiclass model achieved is 92 %
253 (see table 2). The overall accuracy of the model is calculated by averaging the values of
254 precision of the individual classes through their equal shared weights. The loss curve in Fig.
255 5(b) signifies the reduction of signal noises as the training proceeds.



256

257 **Figure 4. Plots of validation curve for the multiclass model using training image dataset.**
 258 **(a) accuracy of the model validated using a training image set inclusive of all seven**
 259 **classes. (b) The accuracy of the model with only one dataset out of seven classes is**
 260 **correct. Thus the accuracy of the model is $\sim 1/7$ (blue line in the graph).**

261



262

263 **Figure 5. Plots of (a) Accuracy curve and (b) loss curve of the multiclass model**

264 The precision of the multiclass model varies widely and can be classified based on its range.
 265 (1) The precision of class C is low $< 90\%$. (2) The precision of classes AB, PA, SA, and ST is
 266 between 90% to $\leq 95\%$. (3) EA and KP have precision $> 95\%$. The values of precision
 267 obtained in binary classes have a similar range for the given bacteria. For, e.g., binary classes
 268 of - EA and KP, PA and SA, and SA and ST have has 98% , 96% , and 98% precision,
 269 respectively. Therefore, it can be concluded that accuracy values are dependent on the
 270 individual bacteria and the extent of change they cause in the blood properties by their
 271 presence. Our investigation found that *Enterococcus faecium* (EF) (one of the ESKAPE
 272 pathogens) did not cause sufficient change in the blood properties, and the algorithm failed to
 273 distinguish the blood droplet residues that were infected with EF. Hence, the bacterial species
 274 EF has not been considered for investigation in the article.

275 In conclusion, we have demonstrated a technique to detect six bacterial species in the blood
276 by recognizing the patterns found in the dried residue of blood droplets using CNN. The
277 detection process is cost-efficient as it does not require any biomarkers/chemicals or high-end
278 equipment. Since the detection procedure involves only capturing an optical image, this can
279 be performed by a non-professional with minimal expertise in the healthcare sector. The
280 trained algorithm developed into a mobile application with a user interface has the potential
281 to revolutionize POC diagnostics. Moreover, the promising results with high accuracy
282 obtained in this article are equivalent to the accuracy of standard bacteriological tests⁵⁷⁻⁵⁹.
283 Besides, the technique has added benefits such as rapid turnaround time (less than 15
284 minutes), low cost, and is suitable for POC.

285 **EXPERIMENTAL PROCEDURE:**

286 All bacterial cultures, namely, *Enterobacter aerogenes*, *Staphylococcus aureus*, *Klebsiella*
287 *pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Salmonella Typhi*,
288 were grown in Luria Bertani medium. Briefly, a single colony of bacterial culture from a
289 freshly streaked LB agar plate was inoculated in LB broth overnight at 37°C at 170 rpm
290 shaker incubator. Next, the bacteria are normalized with OD_{600nm}. The blood was collected in
291 a lavender vacutainer (K3-EDTA, Quantum Biomedicals) from five healthy volunteers, age
292 24-30 years, with diverse blood groups (O +ve, O -ve, AB +ve, B +ve, B -ve), and it was
293 stored at 4°C and used for experiments up to a week. These storage conditions do not
294 influence the blood properties (for a week), and thus the drying process is unaffected due to
295 storage³⁶. Bloodstream infection diagnosis ranges from 1-10 CFU/mL^{60,61} to 10³-10⁴
296 CFU/mL⁶²; blood collected from the donor was thus spiked with bacterial culture to a final
297 concentration of 10³ CFU/mL to mimic tangible infection in real-life patients, and un-spiked
298 blood is used as a control (C). Since equal volumes of phosphate-buffered saline (PBS- 137
299 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, and 2 mM KH₂PO₄) is added to normalize all the
300 bacterial numbers to a constant of 1000 CFU/mL, an equal volume of PBS without any
301 bacteria was added as a control. To observe bacterial-induced changes in the blood (as per the
302 model of bloodstream infection), we incubated the blood at static conditions 37°C for 24
303 hours⁶⁴ before drop-casting it. Subsequently, the control sample (C) is also placed in the same
304 static conditions for 24 hours before drop-casting the same.

305 **Experimental setup for droplet evaporation and imaging:**

306 Optically flat plain glass slides procured from Blue Star© are cleaned using the following
307 procedure: the glass slides are immersed in an ultrasound bath sonicator (from Rivotek™)
308 containing isopropyl alcohol (commonly known as isopropanol) and are sonicated for 15
309 minutes. This is followed by a de-ionized (DI) water rinse and blow-dry. The average surface
310 roughness (R_a) is ~40 nm (measured using the optical profilometer tool by The TalySurf CCI
311). As shown in Figures 1 (b) and (c), the incubated samples from the vials are drop cast gently
312 onto the clean glass slides using a micropipette (Finnpipette®). The volume of the droplets
313 ranges from 0.5 to 2 µl. Hundreds of droplets are placed on the glass slide in linear
314 arrangements separated at least by the distance of 2 times the droplet diameter (L_C), ensuring
315 no vapor-mediated interactions between the individual droplets⁶⁵ (refer to Fig.1 (c)). The
316 glass slides are labeled as the sample deposited on a particular slide is known. The droplets
317 are allowed to evaporate under controlled atmospheric conditions (26 ±5 °C and 45 ±10%
318 relative humidity measured by TSP-01, Thorlabs) in an acrylic enclosure to minimize the

319 external convection disturbances. All experiments are conducted inside a bio-safety hood
320 with HEPA filters following the biosafety laboratory -II protocols. The bacterial samples and
321 other plastic wares were discarded with 10% bleach, followed by double autoclaving all the
322 waste.

323 The blood drops generally evaporate in 10-12 minutes; they are imaged using an optical
324 microscope (Olympus microscope). The droplet is illuminated by an LED light source in-line
325 with the objective lens of the microscope—a digital camera (Nikon D7200) attached to the
326 microscope captures the reflected light and images of the droplet. The captured images were
327 manually filtered for uniformity and shape (as the $Bo \ll 1$, the droplet assumes a spherical
328 shape). The background of the images (other than the droplet region) is removed to reduce
329 the redundant information. The captured images constitute the image data set to be processed
330 through the CNN algorithm.

331

332 **Credit statement**

333 Conceptualization: SB, DC; Methodology: OH, RC, DR; Investigation: OH, RC, VJ,
334 Visualization: OH, RC, DR, VJ; Funding acquisition: SB, DC; Project administration: OH,
335 VJ; Supervision: SB, DC; Writing original draft: OH; Editing and revision: OH, RC, VJ, DC,
336 SB.

337 **Acknowledgments**

338 The authors acknowledge Suraksha Sunil, Amey Agharkar, Ankur Chattopadhyay, Srinivas
339 Rao S, Rovin Pinto, and Gannena K.S. Raghuram for assistance in drop-casting and imaging
340 of blood droplets. The authors acknowledge all the volunteers who participated in blood
341 donation - Amey, Anmol, Vivek, Omkar, Durbar. SB: DRDO Chair Professorship,
342 DC: Infrastructure support from ICMR (Center for Advanced Study in Molecular Medicine),
343 DST (FIST), UGC-CAS (special assistance), DAE-SRC fellowship, ASTRA-Chair
344 fellowship, TATA Innovation grant, DBT-IOE partnership grant. RC duly acknowledges
345 CSIR-SRF for financial assistance.

346 **Ethics declarations and approval**

347 Human Ethical Review Committee of the Indian Institute of Science, Bangalore, approved
348 the experimental protocols.

349 **Informed consent**

350 Informed consent was obtained from all participants/volunteers.

351 **Data availability statement**

352 All relevant data are within the paper and its Supporting Information files. All materials and
353 additional data are available from the corresponding author upon request.

354 **References**

- 355 1. W. H. O. New report calls for urgent action to avert antimicrobial resistance crisis. *Jt.*
356 *News Release* **29**, 2019–2022 (2019).
- 357 2. Bloom, D. E. & Cadarette, D. Infectious disease threats in the twenty-first century:
358 Strengthening the global response. *Frontiers in Immunology* vol. 10 549 (2019).
- 359 3. Biggest Threats and Data | Antibiotic/Antimicrobial Resistance | CDC.
360 <https://www.cdc.gov/drugresistance/biggest-threats.html>.

- 361 4. El-Mahallawy, H. A., Hassan, S. S., El-Wakil, M. & Moneer, M. M. Bacteremia due to
362 ESKAPE pathogens: An emerging problem in cancer patients. *J. Egypt. Natl. Canc.*
363 *Inst.* **28**, 157–162 (2016).
- 364 5. Gandra, S. *et al.* The Mortality Burden of Multidrug-resistant Pathogens in India: A
365 Retrospective, Observational Study. *Clin. Infect. Dis.* **69**, 563–570 (2019).
- 366 6. Chatterjee, A. & Duerkop, B. A. Beyond bacteria: Bacteriophage-eukaryotic host
367 interactions reveal emerging paradigms of health and disease. *Front. Microbiol.* **9**,
368 1394 (2018).
- 369 7. Collignon, P. J. & McEwen, S. A. One Health—Its Importance in Helping to Better
370 Control Antimicrobial Resistance. *Trop. Med. Infect. Dis. 2019, Vol. 4, Page 22* **4**, 22
371 (2019).
- 372 8. Marturano, J. E., Ave, H., Marturano, J. E. & Lowery, T. J. ESKAPE Pathogens in
373 Bloodstream Infections Are Associated With Higher Cost and Mortality but Can Be
374 Predicted Using Diagnoses Upon Admission. *Open Forum Infect. Dis.* **6**, (2019).
- 375 9. Centers for Disease Control, U. Antibiotic Resistance Threats in the United States,
376 2019. doi:10.15620/cdc:82532.
- 377 10. Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. & Pardesi, K. R. Emerging
378 strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A
379 review. *Front. Microbiol.* **10**, 539 (2019).
- 380 11. Mercer, T. R. & Salit, M. Testing at scale during the COVID-19 pandemic. *Nat. Rev.*
381 *Genet.* **2021 227** **22**, 415–426 (2021).
- 382 12. Birru, M. *et al.* Bacterial profile, antimicrobial susceptibility patterns, and associated
383 factors among bloodstream infection suspected patients attending Arba Minch General
384 Hospital, Ethiopia. *Sci. Reports 2021 111* **11**, 1–13 (2021).
- 385 13. Infectious disease: Battling bacterial blood infection. *Nat.* **2010 4677316** **467**, 637–637
386 (2010).
- 387 14. O’Sullivan, S. *et al.* Developments in transduction, connectivity and AI/machine
388 learning for point-of-care testing. *Sensors (Switzerland)* **19**, (2019).
- 389 15. Faix, J. D. Biomarkers of sepsis. <https://doi.org/10.3109/10408363.2013.764490> **50**,
390 23–36 (2013).
- 391 16. Giuliano, C., Patel, C. R. & Kale-Pradhan, P. B. A Guide to Bacterial Culture
392 Identification And Results Interpretation. *Pharm. Ther.* **44**, 192 (2019).
- 393 17. Freeman, W. M., Walker, S. J. & Vrana, K. E. Quantitative RT-PCR: Pitfalls and
394 potential. *BioTechniques* vol. 26 112–125 (1999).
- 395 18. Gibson, U. E. M., Heid, C. A. & Williams, P. M. A novel method for real time
396 quantitative RT-PCR. *Genome Res.* **6**, 995–1001 (1996).
- 397 19. Louie, M., Louie, L. & Simor, A. E. The role of DNA amplification technology in the
398 diagnosis of infectious diseases. *CMAJ* vol. 163 301–309 (2000).
- 399 20. Yang, S. & Rothman, R. E. PCR-based diagnostics for infectious diseases: Uses,
400 limitations, and future applications in acute-care settings. *Lancet Infectious Diseases*
401 vol. 4 337–348 (2004).

- 402 21. Tarafdar, S., Tarasevich, Y. Y., Dutta Choudhury, M., Dutta, T. & Zang, D. Droplet
403 Drying Patterns on Solid Substrates: From Hydrophilic to Superhydrophobic Contact
404 to Levitating Drops. *Advances in Condensed Matter Physics* vol. 2018 (2018).
- 405 22. Gorr, H. M., Zueger, J. M., McAdams, D. R. & Barnard, J. A. Salt-induced pattern
406 formation in evaporating droplets of lysozyme solutions. *Colloids Surfaces B*
407 *Biointerfaces* **103**, 59–66 (2013).
- 408 23. Carreón, Y. J. P. *et al.* Effects of substrate temperature on patterns produced by dried
409 droplets of proteins. *Colloids Surfaces B Biointerfaces* **203**, 111763 (2021).
- 410 24. Sett, A., Ayushman, M., Desgupta, S. & Dasgupta, S. Analysis of the Distinct Pattern
411 Formation of Globular Proteins in the Presence of Micro- and Nanoparticles. *J. Phys.*
412 *Chem. B* **122**, 8972–8984 (2018).
- 413 25. Pal, A., Gope, A., Obayemi, J. D. & Iannacchione, G. S. Concentration-driven phase
414 transition and self-assembly in drying droplets of diluting whole blood. *Sci. Reports*
415 *2020 101* **10**, 1–12 (2020).
- 416 26. Davidson, Z. S. *et al.* Deposition and drying dynamics of liquid crystal droplets. *Nat.*
417 *Commun.* **8**, (2017).
- 418 27. Song, Y. *et al.* Budding-like division of all-aqueous emulsion droplets modulated by
419 networks of protein nanofibrils. *Nat. Commun.* *2018 91* **9**, 1–7 (2018).
- 420 28. Gerber, J., Lendenmann, T., Eghlidi, H., Schutzius, T. M. & Poulikakos, D. Wetting
421 transitions in droplet drying on soft materials. *Nat. Commun.* *2019 101* **10**, 1–10
422 (2019).
- 423 29. Tavana, H. *et al.* Nanolitre liquid patterning in aqueous environments for spatially
424 defined reagent delivery to mammalian cells. *Nat. Mater.* *2009 89* **8**, 736–741 (2009).
- 425 30. Hegde, O., Chatterjee, R., Rasheed, A., Chakravorty, D. & Basu, S. Multiscale vapor-
426 mediated dendritic pattern formation and bacterial aggregation in complex respiratory
427 biofluid droplets. *J. Colloid Interface Sci.* **606**, 2011–2023 (2022).
- 428 31. Sefiane, K. On the Formation of Regular Patterns from Drying Droplets and Their
429 Potential Use for Bio-Medical Applications. *J. Bionic Eng.* **7**, S82–S93 (2010).
- 430 32. Cameron, J. M., Butler, H. J., Palmer, D. S. & Baker, M. J. Biofluid spectroscopic
431 disease diagnostics: A review on the processes and spectral impact of drying. *J.*
432 *Biophotonics* **11**, e201700299 (2018).
- 433 33. Gökçe, O., Castonguay, S., Temiz, Y., Gervais, T. & Delamarche, E. Self-coalescing
434 flows in microfluidics for pulse-shaped delivery of reagents. *Nat.* *2019 5747777* **574**,
435 228–232 (2019).
- 436 34. Mehlem, A., Hagberg, C. E., Muhl, L., Eriksson, U. & Falkevall, A. Imaging of neutral
437 lipids by oil red O for analyzing the metabolic status in health and disease. *Nat.*
438 *Protoc.* *2013 86* **8**, 1149–1154 (2013).
- 439 35. Chen, R., Zhang, L., Zang, D. & Shen, W. Blood drop patterns: Formation and
440 applications. *Adv. Colloid Interface Sci.* **231**, 1–14 (2016).
- 441 36. Brutin, D., Sobac, B., Loquet, B. & Sampol, J. Pattern formation in drying drops of
442 blood. *J. Fluid Mech.* **667**, 85–95 (2011).

- 443 37. Iqbal, R., Shen, A. Q. & Sen, A. K. Understanding of the role of dilution on
444 evaporative deposition patterns of blood droplets over hydrophilic and hydrophobic
445 substrates. *J. Colloid Interface Sci.* **579**, 541–550 (2020).
- 446 38. Smith, F. R. & Brutin, D. Wetting and spreading of human blood: Recent advances
447 and applications. *Curr. Opin. Colloid Interface Sci.* **36**, 78–83 (2018).
- 448 39. Pal, A., Gope, A. & Iannacchione, G. Temperature and Concentration Dependence of
449 Human Whole Blood and Protein Drying Droplets. *Biomol. 2021, Vol. 11, Page 231*
450 **11**, 231 (2021).
- 451 40. Bou-Zeid, W. & Brutin, D. Effect of relative humidity on the spreading dynamics of
452 sessile drops of blood. *Colloids Surfaces A Physicochem. Eng. Asp.* **456**, 273–285
453 (2014).
- 454 41. Laan, N. *et al.* Bloodstain Pattern Analysis: implementation of a fluid dynamic model
455 for position determination of victims. *Sci. Reports 2015 51* **5**, 1–8 (2015).
- 456 42. Ewald, D. R. & Sumner, S. C. J. Blood type biochemistry and human disease. *Wiley*
457 *Interdiscip. Rev. Syst. Biol. Med.* **8**, 517–535 (2016).
- 458 43. Kaniyas, T. *et al.* Ethnicity, sex, and age are determinants of red blood cell storage and
459 stress hemolysis: results of the REDS-III RBC-Omics study. *Blood Adv.* **1**, 1132–1141
460 (2017).
- 461 44. Simmonds, M. J., Meiselman, H. J. & Baskurt, O. K. Blood rheology and aging. *J.*
462 *Geriatr. Cardiol.* **10**, 291 (2013).
- 463 45. Ahmed, G., Arjmandi Tash, O., Cook, J., Trybala, A. & Starov, V. Biological
464 applications of kinetics of wetting and spreading. *Adv. Colloid Interface Sci.* **249**, 17–
465 36 (2017).
- 466 46. Sikarwar, B. S., Roy, M., Ranjan, P. & Goyal, A. Automatic Pattern Recognition for
467 Detection of Disease from Blood Drop Stain Obtained with Microfluidic Device. *Adv.*
468 *Intell. Syst. Comput.* **425**, 655–667 (2016).
- 469 47. Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S. & Thoma, G. Image analysis and
470 machine learning for detecting malaria. *Transl. Res.* **194**, 36–55 (2018).
- 471 48. Hamadeh, L. *et al.* Machine Learning Analysis for Quantitative Discrimination of
472 Dried Blood Droplets. *Sci. Reports 2020 101* **10**, 1–13 (2020).
- 473 49. Harris, C. R. *et al.* Array programming with NumPy. *Nat. 2020 5857825* **585**, 357–362
474 (2020).
- 475 50. Gulli, A. & Pal, S. *Deep learning with Keras.* (2017).
- 476 51. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95
477 (2007).
- 478 52. Waskom, M. L. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021
479 (2021).
- 480 53. Pedregosa Fabianpedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J.*
481 *Mach. Learn. Res.* **12**, 2825–2830 (2011).
- 482 54. Abadi, M. *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous

- 483 Distributed Systems. (2016).
- 484 55. Bradski, G. (2000). The openCV library. *Journal of Software Tool*.
- 485 56. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep
486 Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **25**, (2012).
- 487 57. Opota, O., Croxatto, A., Prod'hom, G. & Greub, G. Blood culture-based diagnosis of
488 bacteraemia: state of the art. *Clin. Microbiol. Infect.* **21**, 313–322 (2015).
- 489 58. Giuliano, C., Patel, C. R. & Kale-Pradhan, P. B. A Guide to Bacterial Culture
490 Identification And Results Interpretation. *Pharm. Ther.* **44**, 192 (2019).
- 491 59. Bergey, D. H. (David H. & Holt, J. G. *Bergey's manual of determinative bacteriology*.
492 *Bergey's manual of determinative bacteriology*. (Williams & Wilkins, 1994).
- 493 60. Wain, J. *et al.* Quantitation of bacteria in blood of typhoid fever patients and
494 relationship between counts and clinical features, transmissibility, and antibiotic
495 resistance. *J. Clin. Microbiol.* **36**, 1683–1687 (1998).
- 496 61. Henry, N. K. *et al.* Microbiological and clinical evaluation of the isolator lysis-
497 centrifugation blood culture tube. *J. Clin. Microbiol.* **17**, 864 (1983).
- 498 62. Werner, A. S., Cobbs, C. G., Kaye, D. & Hook, E. W. Studies on the bacteremia of
499 bacterial endocarditis. *JAMA* **202**, 199–203 (1967).
- 500 63. Wester, A. L., Dunlop, O., Melby, K. K., Dahle, U. R. & Wyller, T. B. Age-related
501 differences in symptoms, diagnosis and prognosis of bacteremia. *BMC Infect. Dis.* **13**,
502 (2013).
- 503 64. Dong, M. *et al.* Standardized methods to generate mock (spiked) clinical specimens by
504 spiking blood or plasma with cultured pathogens. *J. Appl. Microbiol.* **120**, 1119–1129
505 (2016).
- 506 65. Shaikeea, A., Basu, S., Hatte, S. & Bansal, L. Insights into Vapor-Mediated
507 Interactions in a Nanocolloidal Droplet System: Evaporation Dynamics and Affects on
508 Self-Assembly Topologies on Macro- to Microscales. *Langmuir* **32**, 10334–10343
509 (2016).
- 510

