Convolution Neural Networks for Point-of-Care Diagnostics of 1 **Bacterial Infections in Blood** 2 Omkar Hegde^{1,†}, Ritika Chatterjee^{2,†}, Durbar Roy^{1,+}, Vivek Jaiswal^{1,+}, Dipshikha 3 Chakravortty^{2*}, Saptarshi Basu^{1*} 4 ¹Department of Mechanical Engineering, Indian Institute of Science, Bangalore 5 ²Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 6 7 [†] These authors contributed equally as first authors ⁺These authors contributed equally as second authors 8 9 *Corresponding Author¹ email – sbasu@iisc.ac.in *Corresponding Author² email – dipa@iisc.ac.in 10 11 12

13 **ABSTRACT:**

A droplet of blood, when evaporated on a surface, leaves dried residue—the fractal patterns 14 formed on the dried residues can act as markers for infection present in the blood. Exploiting 15 the unique patterns found in the residues of a naturally dried droplet of blood, we propose a 16 Point-of-Care (POC) diagnostic tool for detecting broad-spectrum of bacterial infections 17 (such as Enterobacter aerogenes, Staphylococcus aureus, Klebsiella pneumoniae, 18 Acinetobacter baumannii, and Pseudomonas aeruginosa, Salmonella Typhi) in blood. The 19 diagnosis process we propose is straightforward and can be performed with the following 20 steps: A droplet of blood (healthy or infected) of volume range 0.5 to 2 μl is allowed to dry 21 on a clean glass surface and is imaged using a conventional optical microscope. A computer 22 algorithm based on the framework of convolution neural network (CNN) is used to classify 23 the captured images of dried blood droplets according to the bacterial infection. In total, our 24 25 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 26 detecting bacteria in the blood reported in this article is commensurate with the standard 27 bacteriological tests. Thus, this article presents a proof-of-concept of a potential futuristic tool 28 for a rapid and low-cost diagnosis of bacterial infection in the blood. 29

30 INTRODUCTION:

Antimicrobial resistance (AMR) has become a cause of global concern, aggravated by a 31 lack of data on the different mechanisms and factors causing its emergence and spread. 32 According to a WHO report, bacterial infections alone cause nearly 1 million deaths per year 33 34 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, 35 36 Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter aerogenes) pathogens are the deadliest, and every passing 37 year they cause millions of infections and thousands of deaths worldwide^{2–6}. These bacteria 38 are clinically relevant strains causing ~42% bloodstream infection and have been associated 39 with significant mortality and morbidity^{7,8}. The WHO report lists these pathogens as priority 40 antibiotic-resistant pathogens as they are associated with a significant amount of mortality 41 42 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 43 European countries¹⁰ are finding it difficult to cope with the increasing AMR; an estimated 44 45 cost of billions of dollars is attributed to the health care expenditure of hospital-acquired and community-acquired AMR infections⁹. The first step to stop the spreading of infection would 46

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

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

With the advent of fundamental research on droplet evaporation and colloidal self-78 assembly in droplets in the past few decades, the dried precipitates of biological fluids have 79 been deeply studied²¹⁻³⁰. A few of these studies have found applications in biomedical 80 diagnostics^{31–34}. The competitive interaction of multi-component elements in biological fluid 81 like blood leads to the formation of rich patterns on the dried droplet precipitate³⁵ such as 82 distinctive radial, spiral, and tangential cracks, color variations, varying thickness of the 83 84 coffee rings, finger formations, etc. (Refer to image S1 in the supplementary information for the image of a dried droplet of blood taken from a healthy volunteer). These features depend 85 on the properties of the fluid (and thereby the components present in it), which in turn can be 86 related to the health of a person. Notably, in blood droplets, it has been shown previously that 87 the flow inside the droplet is driven by Marangoni, wettability, and natural evaporation leads 88 to the formation of distinctive patterns on the dried residues of the drop for anemic and 89 hyperlipidemic blood samples $\frac{1}{36-38}$. Besides, the final dried blood droplet pattern is 90 susceptible to drying conditions such as temperature³⁹ and humidity⁴⁰, placement of blood 91 droplet⁴¹ (gentle or with impact), and depends on several other factors such as blood groups⁴², 92 ethnicity⁴³, age⁴⁴, gender⁴³, size, shape, and mobility of cells⁴⁵. Further, blood infection leads 93 to cellular changes such as Leucocytosis⁶³ or leukopenia, and biochemical changes include 94 markers released upon red blood cell lysis or white blood cell lysis⁶³. Thus, the blood 95

- 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 $drop^{46}$; for example, it is possible to
- 98 detect malaria in dried blood droplets using a simple light microscope imageing 47 .

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

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

129 **RESULTS AND DISCUSSION:**

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

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





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

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

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





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

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

200 Hence the algorithm is developed in a two-step process. First, the model is optimized for binary classes, i.e., the algorithm is trained to distinguish between any two classes with 201 high accuracy. Since there are seven classes, the model for ${}^{7}C_{2}$ binary classes is optimized 202 (see Fig. 3). Second, the parameters used for optimizing the binary classes (such as epochs -203 the number of times the algorithm iterates the whole dataset for training, kernel size for 204 convolutional layers⁵¹, image size, and learning rate - the variation in the model with updated 205 weights) are used in the multiclass model with all the classes included, as shown in Fig. 3 (7 206 classes including the control case). While the basic structure of the algorithm remains the 207 same, the parameters optimized for binary classes are used as a template for the multiclass 208 model with all the classes. 209

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



223

Figure 3. Schematic representation of the two-step development of the model used in the article. Optimal model parameters such as the epoch, learning rate, image size, and kernel size for the convolution layer obtained from the $^{7}C_{2}$ binary classes are adopted into the multiclass model consisting of seven classes.

228

Table 1: Accuracy, precision, and recall values for ⁷C₂ 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 |

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

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

241

Table 2: Accuracy, precision, and recall values of a multiclass model consisting of 7 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 |
| С | | 0.81 | 0.98 |

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



256

261

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





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

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

In conclusion, we have demonstrated a technique to detect six bacterial species in the blood 275 by recognizing the patterns found in the dried residue of blood droplets using CNN. The 276 detection process is cost-efficient as it does not require any biomarkers/chemicals or high-end 277 278 equipment. Since the detection procedure involves only capturing an optical image, this can be performed by a non-professional with minimal expertise in the healthcare sector. The 279 trained algorithm developed into a mobile application with a user interface has the potential 280 to revolutionize POC diagnostics. Moreover, the promising results with high accuracy 281 obtained in this article are equivalent to the accuracy of standard bacteriological tests^{57–59}. 282 Besides, the technique has added benefits such as rapid turnaround time (less than 15 283 minutes), low cost, and is suitable for POC. 284

285 **EXPERIMENTAL PROCEDURE:**

All bacterial cultures, namely, Enterobacter aerogenes, Staphylococcus aureus, Klebsiella 286 pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Salmonella Typhi, 287 288 were grown in Luria Bertani medium. Briefly, a single colony of bacterial culture from a freshly streaked LB agar plate was inoculated in LB broth overnight at 37°C at 170 rpm 289 shaker incubator. Next, the bacteria are normalized with OD_{600nm}. The blood was collected in 290 a lavender vacutainer (K3-EDTA, Quantum Biomedicals) from five healthy volunteers, age 291 24-30 years, with diverse blood groups (O +ve, O -ve, AB +ve, B +ve, B -ve), and it was 292 stored at 4°C and used for experiments up to a week. These storage conditions do not 293 influence the blood properties (for a week), and thus the drying process is unaffected due to 294 storge³⁶. Bloodstream infection diagnosis ranges from 1-10 CFU/mL^{60,61} to 10³-10⁴ 295 CFU/mL⁶²; blood collected from the donor was thus spiked with bacterial culture to a final 296 concentration of 10³ CFU/mL to mimic tangible infection in real-life patients, and un-spiked 297 blood is used as a control (C). Since equal volumes of phosphate-buffered saline (PBS-137 298 299 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, and 2 mM KH₂PO₄) is added to normalize all the bacterial numbers to a constant of 1000 CFU/mL, an equal volume of PBS without any 300 bacteria was added as a control. To observe bacterial-induced changes in the blood (as per the 301 model of bloodstream infection), we incubated the blood at static conditions 37°C for 24 302 hours⁶⁴ before drop-casting it. Subsequently, the control sample (C) is also placed in the same 303 static conditions for 24 hours before drop-casting the same. 304

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

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

external convection disturbances. All experiments are conducted inside a bio-safety hood
with HEPA filters following the biosafety laboratory -II protocols. The bacterial samples and
other plastic wares were discarded with 10% bleach, followed by double autoclaving all the
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 with the objective lens of the microscope—a digital camera (Nikon D7200) attached to the 325 microscope captures the reflected light and images of the droplet. The captured images were 326 327 manually filtered for uniformity and shape (as the Bo<<1, the droplet assumes a spherical shape). The background of the images (other than the droplet region) is removed to reduce 328 the redundant information. The captured images constitute the image data set to be processed 329 through the CNN algorithm. 330

331

332 Credit statement

Conceptualization: SB, DC; Methodology: OH, RC, DR; Investigation: OH, RC, VJ,
Visualization: OH, RC, DR, VJ; Funding acquisition: SB, DC; Project administration: OH,

VJ; Supervision: SB, DC; Writing original draft: OH; Editing and revision: OH, RC, VJ, DC,
SB.

337 Acknowledgments

The authors acknowledge Suraksha Sunil, Amey Agharkar, Ankur Chattopadhyay, Srinivas 338 339 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, DC: Infrastructure support from ICMR (Center for Advanced Study in Molecular Medicine), 342 343 DST (FIST), UGC-CAS (special assistance), DAE-SRC fellowship, ASTRA-Chair 344 fellowship, TATA Innovation grant, DBT-IOE partnership grant. RC duly acknowledges CSIR-SRF for financial assistance. 345

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

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

354 **References**

- W. H. O. New report calls for urgent action to avert antimicrobial resistance crisis. *Jt. News Release* 29, 2019–2022 (2019).
- Bloom, D. E. & Cadarette, D. Infectious disease threats in the twenty-first century:
 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.

- 4. El-Mahallawy, H. A., Hassan, S. S., El-Wakil, M. & Moneer, M. M. Bacteremia due to
 ESKAPE pathogens: An emerging problem in cancer patients. *J. Egypt. Natl. Canc. Inst.* 28, 157–162 (2016).
- Gandra, S. *et al.* The Mortality Burden of Multidrug-resistant Pathogens in India: A
 Retrospective, Observational Study. *Clin. Infect. Dis.* 69, 563–570 (2019).
- 366 6. Chatterjee, A. & Duerkop, B. A. Beyond bacteria: Bacteriophage-eukaryotic host interactions reveal emerging paradigms of health and disease. *Front. Microbiol.* 9, 1394 (2018).
- Collignon, P. J. & McEwen, S. A. One Health—Its Importance in Helping to Better
 Control Antimicrobial Resistance. *Trop. Med. Infect. Dis. 2019, Vol. 4, Page 22* 4, 22
 (2019).
- Marturano, J. E., Ave, H., Marturano, J. E. & Lowery, T. J. ESKAPE Pathogens in Bloodstream Infections Are Associated With Higher Cost and Mortality but Can Be 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.
- Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. & Pardesi, K. R. Emerging
 strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A
 review. *Front. Microbiol.* 10, 539 (2019).
- Mercer, T. R. & Salit, M. Testing at scale during the COVID-19 pandemic. *Nat. Rev. Genet.* 2021 227 22, 415–426 (2021).
- Birru, M. *et al.* Bacterial profile, antimicrobial susceptibility patterns, and associated factors among bloodstream infection suspected patients attending Arba Minch General Hospital, Ethiopia. *Sci. Reports 2021 111* **11**, 1–13 (2021).
- Infectious disease: Battling bacterial blood infection. *Nat. 2010* 4677316 467, 637–637 (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, 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).
- Freeman, W. M., Walker, S. J. & Vrana, K. E. Quantitative RT-PCR: Pitfalls and 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 quantitative RT-PCR. *Genome Res.* 6, 995–1001 (1996).
- Louie, M., Louie, L. & Simor, A. E. The role of DNA amplification technology in the diagnosis of infectious diseases. *CMAJ* vol. 163 301–309 (2000).
- Yang, S. & Rothman, R. E. PCR-based diagnostics for infectious diseases: Uses,
 limitations, and future applications in acute-care settings. *Lancet Infectious Diseases*vol. 4 337–348 (2004).

- Tarafdar, S., Tarasevich, Y. Y., Dutta Choudhury, M., Dutta, T. & Zang, D. Droplet
 Drying Patterns on Solid Substrates: From Hydrophilic to Superhydrophobic Contact
 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).
- Pal, A., Gope, A., Obayemi, J. D. & Iannacchione, G. S. Concentration-driven phase
 transition and self-assembly in drying droplets of diluting whole blood. *Sci. Reports*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 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., Chakravortty, D. & Basu, S. Multiscale vapor426 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).
- 33. Gökçe, O., Castonguay, S., Temiz, Y., Gervais, T. & Delamarche, E. Self-coalescing
 flows in microfluidics for pulse-shaped delivery of reagents. *Nat. 2019* 5747777 574,
 228–232 (2019).
- 436 34. Mehlem, A., Hagberg, C. E., Muhl, L., Eriksson, U. & Falkevall, A. Imaging of neutral 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 applications. *Adv. Colloid Interface Sci.* 231, 1–14 (2016).
- 36. Brutin, D., Sobac, B., Loquet, B. & Sampol, J. Pattern formation in drying drops of
 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).
- 38. Smith, F. R. & Brutin, D. Wetting and spreading of human blood: Recent advances 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).
- 40. Bou-Zeid, W. & Brutin, D. Effect of relative humidity on the spreading dynamics of sessile drops of blood. *Colloids Surfaces A Physicochem. Eng. Asp.* **456**, 273–285 (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).
- 43. Kanias, T. *et al.* Ethnicity, sex, and age are determinants of red blood cell storage and stress hemolysis: results of the REDS-III RBC-Omics study. *Blood Adv.* 1, 1132–1141 (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 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 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).
- Wain, J. *et al.* Quantitation of bacteria in blood of typhoid fever patients and relationship between counts and clinical features, transmissibility, and antibiotic resistance. J. Clin. Microbiol. 36, 1683–1687 (1998).
- 496 61. Henry, N. K. *et al.* Microbiological and clinical evaluation of the isolator lysis497 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 bacterial endocarditis. *JAMA* 202, 199–203 (1967).
- 63. Wester, A. L., Dunlop, O., Melby, K. K., Dahle, U. R. & Wyller, T. B. Age-related differences in symptoms, diagnosis and prognosis of bacteremia. *BMC Infect. Dis.* 13, (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

















