nature medicine

LETTERS

https://doi.org/10.1038/s41591-021-01384-9

Check for updates

Deep learning of HIV field-based rapid tests

Valérian Turbé[1]✉, Carina Herbst[2], Thobeka Mngomezulu[2], Sepehr Meshkinfamfard[1], Nondumiso Dlamini[2], Thembani Mhlongo[2], Theresa Smit[2], Valeriia Cherepanova[3], Koki Shimada[3], Jobie Budd[1,4], Nestor Arsenov[1], Steven Gray[5], Deenan Pillay[2,6], Kobus Herbst[2,7]✉, Maryam Shahmanesh[2,8]✉ and Rachel A. McKendry[1,4]✉

Although deep learning algorithms show increasing promise for disease diagnosis, their use with rapid diagnostic tests performed in the field has not been extensively tested. Here we use deep learning to classify images of rapid human immunodeficiency virus (HIV) tests acquired in rural South Africa. Using newly developed image capture protocols with the Samsung SM-P585 tablet, 60 fieldworkers routinely collected images of HIV lateral flow tests. From a library of 11,374 images, deep learning algorithms were trained to classify tests as positive or negative. A pilot field study of the algorithms deployed as a mobile application demonstrated high levels of sensitivity (97.8%) and specificity (100%) compared with traditional visual interpretation by humans—experienced nurses and newly trained community health worker staff—and reduced the number of false positives and false negatives. Our findings lay the foundations for a new paradigm of deep learning–enabled diagnostics in low- and middle-income countries, termed REASSURED diagnostics[1], an acronym for real-time connectivity, ease of specimen collection, affordable, sensitive, specific, user-friendly, rapid, equipment-free and deliverable. Such diagnostics have the potential to provide a platform for workforce training, quality assurance, decision support and mobile connectivity to inform disease control strategies, strengthen healthcare system efficiency and improve patient outcomes and outbreak management in emerging infections.

Rapid diagnostic tests (RDTs) save lives by informing case management, treatment, screening, disease control and elimination programs[1]. Lateral flow tests are among the most common RDTs, and hundreds of millions of these tests are performed worldwide each year. They have the potential to support near-person testing and decentralized management of a range of clinically important diseases (including malaria, HIV, syphilis, tuberculosis, influenza and noncommunicable diseases[2]), making it convenient for the end user and more affordable for health systems[3]. However, RDTs also present some issues, namely: errors in performing the test and interpreting the result[4,5], quality control and lack of electronic data capture records of the test and results within health systems and surveillance. Many of these would be overcome with the real-time connectivity associated with REASSURED—the new criterion for an ideal test to reflect the importance of digital connectivity, coined by Peeling and coworkers[1]. Real-time connectivity involves the use of mobile-phone-connected RDTs. To date there have been few peer-reviewed studies or evaluations of the effectiveness of connected lateral flow tests at scale in populations in low- and middle-income countries.

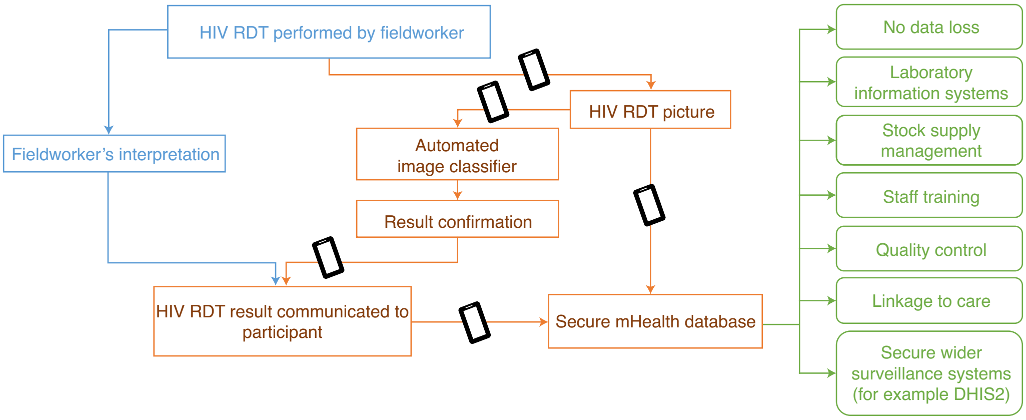

HIV RDT performed by fieldworker
Fieldworker's interpretation
HIV RDT picture
Automated image classifier
Result confirmation
HIV RDT result communicated to participant
Secure mHealth database
No data loss
Laboratory information systems
Stock supply management
Staff training
Quality control
Linkage to care
Secure wider surveillance systems (for example DHIS2)

Fig. 1 | Infographic illustrating the benefits of data capture in supporting field decisions. Current workflow used by fieldworkers (blue); our proposed mHealth system of automated RDT classifier plus data capture and transmission to a secure mHealth database (orange); and the benefits arising from deploying the proposed system (green). Black rectangles represent tablets or smartphones.

[1]London Centre for Nanotechnology, University College London, London, UK. [2]Africa Health Research Institute, Nelson R. Mandela Medical School, Durban, South Africa. [3]Department of Computer Science, University College London, London, UK. [4]Division of Medicine, University College London, London, UK. [5]UCL Centre for Advanced Spatial Analysis, London, UK. [6]Division of Infection and Immunity, University College London, London, UK. [7]DSI-MRC South African Population Research Infrastructure Network, Durban, South Africa. [8]Institute for Global Health, University College London, London, UK. ✉e-mail: v.turbe@ucl.ac.uk; Kobus.Herbst@ahri.org; m.shahmanesh@ucl.ac.uk; r.a.mckendry@ucl.ac.uk

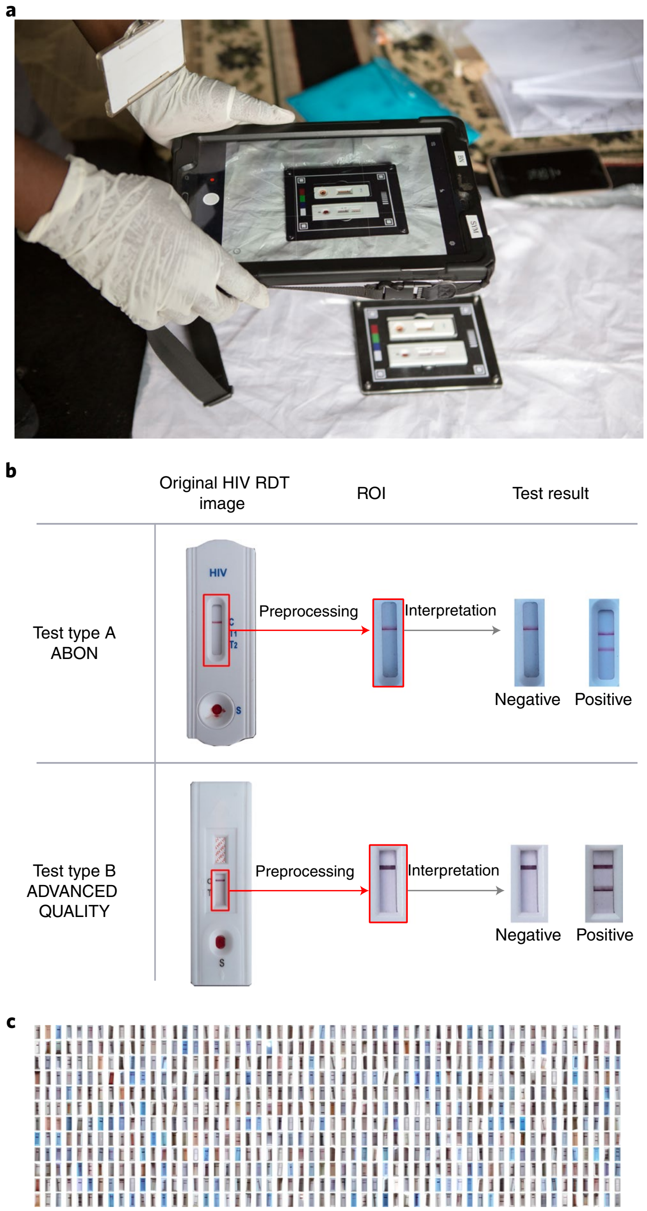


Fig. 2 | Standardization of image capture, image preprocessing and training library. a, Fieldworker capturing a photograph of two HIV RDTs at the time of interpretation, in the field in rural South Africa (image credit: Africa Health Research Institute). The two HIV RDTs are fitted in a plastic tray designed to standardize image capture and facilitate image preprocessing. **b**, Interpretation process, starting from the original picture of HIV RDTs used during the study, preprocessing to select the ROI then interpretation of the test result. If two lines (control + test) are present on the paper strip at the time of interpretation, the test result is positive. Note: for the ABON HIV RDT, one or two different test lines can appear (T1 and T2) depending on the type of HIV infection (HIV-1 and HIV-2, respectively). The test result is positive regardless of which test line is present, or if both test lines are present on the paper strip at the time of interpretation. If only the top line (control) is present, the test is negative; if no control line can be seen, the test is deemed invalid. **c**, Snapshot of the image library of HIV RDTs collected in the field in rural South Africa (162 randomly selected images out of 11,374), illustrating the diversity of color, background and brightness.

Recent studies comparing the human interpretation of a HIV RDT to various gold standards, such as immunoblot[6–9], enzyme immunoassay[7,9–11], standardized test panels[12] or different HIV RDTs[13–15], have highlighted the common issue of subjective interpretation of the test result, which can lead to incorrect diagnosis. User error (especially in the case of weak reactive lines) and inadequate supervision of testers were identified as prime factors for misinterpretation[16]. In a study of differently experienced users interpreting results of HIV RDTs by looking at pictures of tests[17], the accuracy of interpretation varied between 80 and 97%. This highlights the importance of experience in reading the test, as well as the subjectivity involved in reading a weak test line. Evidence also suggests that some fieldworkers struggle to interpret RDTs because of color blindness or short-sightedness[18]. Another study used photographs of HIV RDTs to quantify the subtle difference in tests with faint lines declared as true positive (TP) or false positive (FP) by a panel of human users[19]. While these were small-scale studies ($n = 148$ and 8, respectively), both highlighted the potential for photographs to improve quality control and decision making.

Deep learning algorithms, harnessing advances in large datasets and processing power, have recently shown the ability to exceed human performance in a plethora of visual tasks, including cell-based diagnostics[20], interpretation of dermatologic[21], ophthalmologic[22] and radiographic images[23], playing strategic games[24] and in clinical medicine when used alongside appropriate guidelines[25,26]. While some emerging studies are looking at the application of deep learning to the interpretation of RDTs[27,28], little is known about the ability of machine learning models to analyze field-acquired diagnostic test data, with concerns about the potential uniformity of images (for example, focus and tilt), harsh environmental factors such as lighting (for example, brightness and shadowing), and the variety of test types. In addition, there is a general lack of large real-world datasets available to successfully train deep learning classifiers, particularly from low- and middle-income countries. Recent advances in consumer electronic devices and deep learning have the potential to improve RDT quality assurance, staff training and connectivity, eventually supporting self-testing such as for HIV, which has been shown to be cost effective[29], to appeal to young people[30] and help reduce anxiety[31].

Mobile health (mHealth) approaches, which marry RDTs with widely available mobile phones, take advantage of inbuilt sensors (for example, cameras) found in the phones, battery life, processing power, screens to display results and connectivity to send results to health databases. A recent field study has shown high levels of acceptability for a device sending HIV RDT results to online databases in real time[32]. An array of approaches have been piloted at small scales ($n \leq 283$) and have shown good performance. However, most require a physical attachment such as a dongle (92–100% sensitivity, 97–100% specificity)[33], a cradle[34] or a portable reader (97–98% sensitivity)[35], which increases cost and complexity, and these are typically reliant on simple image analysis software.

We explore the potential of deep learning algorithms to classify field-based RDT images as either positive or negative, focusing on HIV as an exemplar and piloting at scale in population 'test beds' in KwaZulu-Natal, typical of semi-rural settings in subSaharan Africa. Figure 1 shows the concept of our deep learning–enabled REASSURED diagnostic system to capture and interpret RDT results. Our approach first involved building a large image library of field-acquired test images as a training dataset, optimizing algorithms for high sensitivity and specificity and then deploying our classifier in a pilot study to assess its performance compared to traditional visual interpretation with a range of end users having varying levels of training.

Our standard image collection protocol (Fig. 2a) and library are described in Methods. In brief, 11,374 photographs of HIV RDT were captured by >60 fieldworkers using Samsung tablets (SM-P585, 8-megapixel camera, f1/9 with autofocus capability). Embedding of routine image collection into staff workflows was acceptable and feasible, and participant consent rate was 96%.

a

Test type	A	B	Total
	ABON HIV 1/2/O Tri-Line Human Immunodeficiency Virus Rapid Test Device (whole blood/serum/plasma)	ADVANCED QUALITY ONE STEP Anti-HIV (1&2) Test	
Positive	282	530	812
Negative	4,161	6,401	10,562
Total	4,443	6,931	11,374

b

	Type A, ABON		Type B, ADVANCED QUALITY	
Fold	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
1	100.0	98.1	96.5	100.0
2	96.2	99.0	100.0	99.4
3	100.0	99.0	95.7	99.8
4	100.0	99.5	96.6	99.8
5	100.0	99.3	98.0	100.0
6	83.3	100.0	100.0	100.0
7	96.9	98.5	100.0	99.5
8	96.8	99.0	100.0	100.0
9	96.2	98.1	100.0	99.8
10	90.0	99.8	100.0	99.7
Average	95.9 ± 5.1	99.0 ± 0.6	98.7 ± 1.7	99.8 ± 0.2

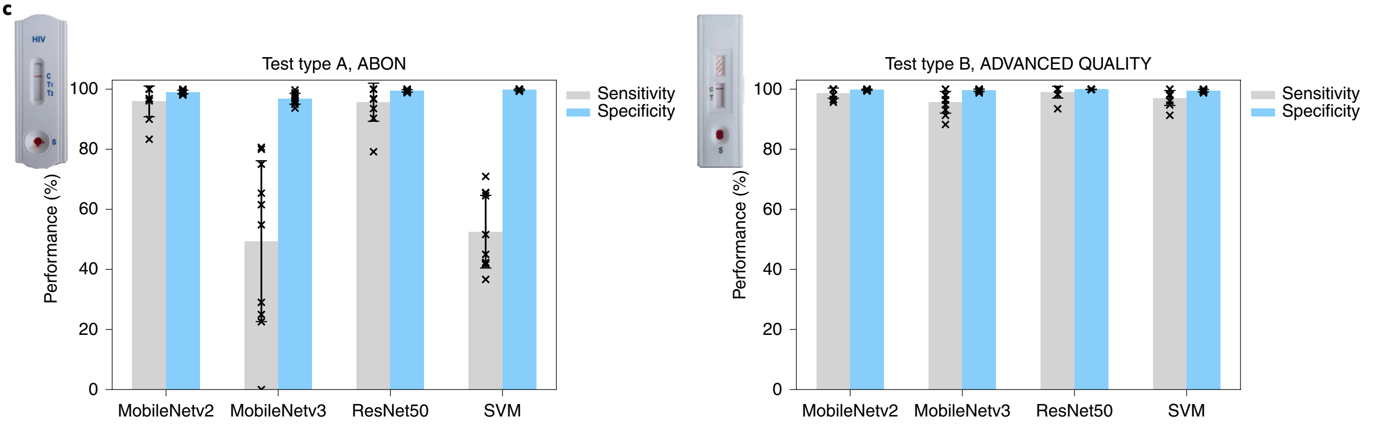


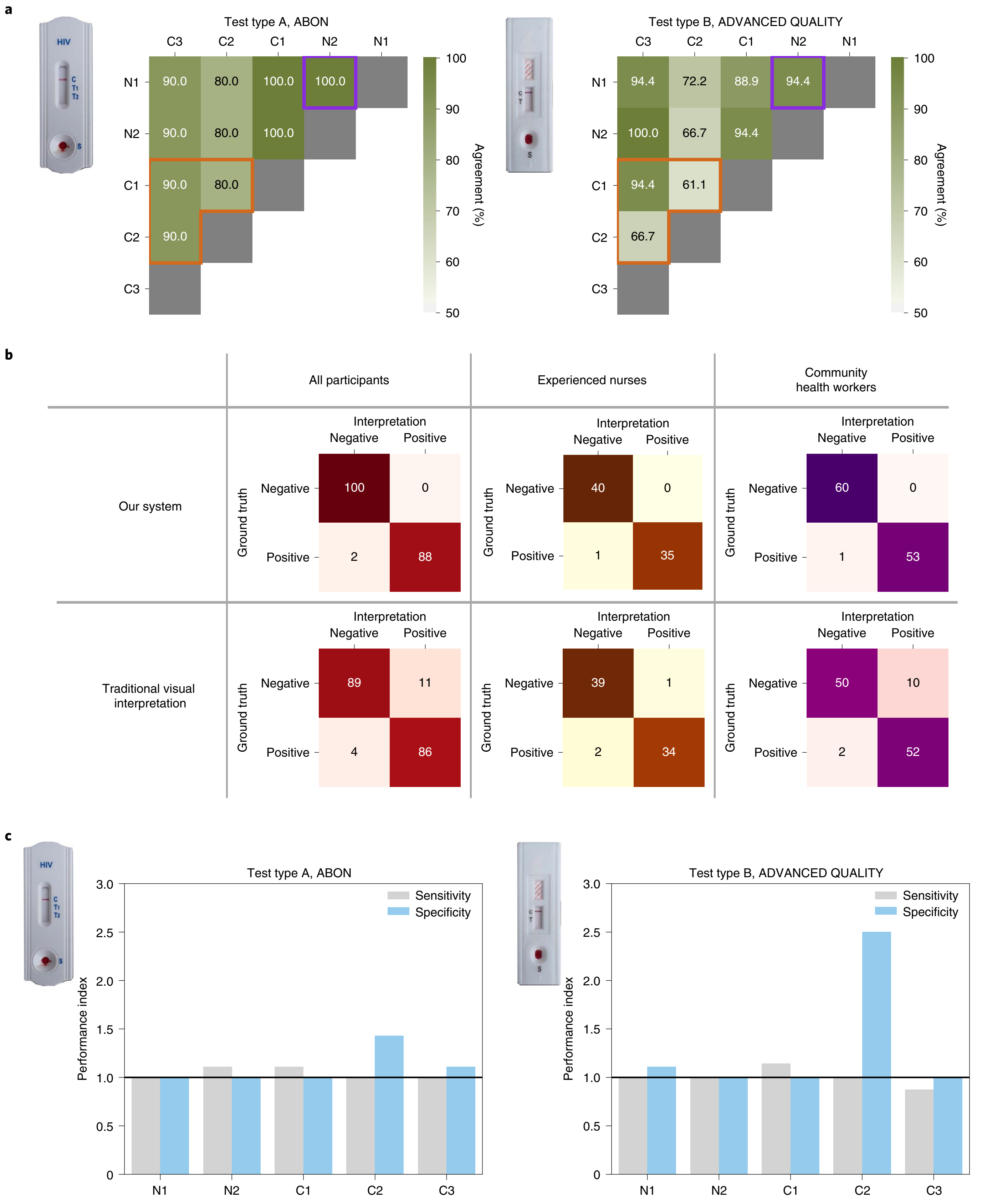
Fig. 3 | Algorithm training and performance. a, Table showing the number of images in the training library, divided into two label categories (positive and negative), as well as two subcategories corresponding to the test type. **b**, Table summarizing the training process using cross-validation, with a training set of $n = 3{,}998$ (type A) and $n = 6{,}221$ (type B). Sensitivity and specificity were obtained using a hold-out testing dataset of $n = 445$ (type A) and $n = 693$ (type B). **c**, Barplots showing the average performance (sensitivity and specificity) of four classification methods trained on our dataset, using cross-validation (error bars represent s.d. from the mean). The three CNNs pretrained on the ImageNet dataset (ResNet50, MobileNetV2 and MobileNetV3) were retrained and tested using our dataset. The SVM was trained using features extracted by the histogram of oriented gradients. All four classifiers were trained using the training set described in **b**. Sensitivity and specificity were obtained using the hold-out testing dataset described in **b**.

We optimized our mHealth system for the two different HIV RDTs used in the study as part of routine household population surveillance. At first glance these RDTs appear similar, but have different features and numbers of test lines. To reduce the number of variables, we cropped the images around the region of interest (ROI) (Fig. 2b). Figure 2c shows a snapshot of the very diverse real-world field conditions where the images were captured (indoors, outdoors, in the shade and in direct sunlight).

Each image was labeled (Methods) according to the test result. Figure 3a details the number of images used to train classifiers to automatically read the result of HIV RDT images. The training process is described in Methods. To test the reproducibility of the process, we performed a tenfold cross-validation. As can be seen in Fig. 3b, the average sensitivity (95.9 ± 5.1% for type A, 98.7 ± 1.7% for type B) and specificity (99.0 ± 0.6% for type A, 99.8 ± 0.2% for type B) achieved across the ten folds was high and consistent for both types of HIV RDT. We therefore used all available data to train a final classifier for each type of test, which were then used in our field study. We investigated different common classification methods in use for clinical diagnostics (support vector machine[36] (SVM) and convolutional neural networks (CNNs)), including three different CNN architectures (ResNet50 (ref. [37]), MobileNetV2 (refs. [38,39]) and MobileNetV3 (ref. [40]), and found MobileNetV2 the most appropriate for our task, as can be seen in Fig. 3c.

We then conducted a field pilot study in rural South Africa to assess the performance of our mHealth system compared to visual interpretation, with a range of end users having varying levels of training (Methods). Five participants (two nurses and three newly trained community health workers) were each asked to give their interpretation of 40 HIV RDTs and to acquire a photograph of the RDT via the application. The plastic trays used to collect the image library were not used in this pilot study. All five participants (100%) were able to use our mHealth system without training, demonstrating its feasibility and acceptability. The photographs were then evaluated by an expert RDT interpreter, followed by our deep learning algorithms on a secure server. The results were not fed back to the study participants, to avoid confirmation bias. The performance results can be seen in Fig. 4.

When comparing the traditional visual interpretation of RDTs we observed varied levels of agreement between participants (61–100%) as can be seen in Fig. 4a. As expected, agreement between nurses (N1 and N2: 100 and 94.4% agreement for test types A and B, respectively) was greater than that between newly trained community health workers (C1, C2 and C3: 80–90 and 61.1–94.4% for test types A and B, respectively). Test type B showed the lower level of agreement. The low level of agreement between participants, and variability due to the type of HIV RDT, were of concern and highlighted the need for a more objective and consistent method to interpret HIV RDTs in the field. The confusion matrices in Fig. 4b demonstrate that our mHealth system reduced the number of errors in reading RDTs. The number of FP results from our mHealth system was found to be lower than that for traditional visual interpretation (0 compared to 11—the largest variation being observed for community health workers, 10), which translates as an improvement in specificity from 89 to 100% and an improvement in positive predictive value from 88.7 to 100%. Similarly, the number of false-negative (FN) results was just two in our mHealth system compared to four for traditional visual interpretation, which translates as an improvement in sensitivity from 95.6 to 97.8% and an improvement in negative predictive value from 95.7 to 98%.



We plotted the ratio of our mHealth system performance to participant performance, for both sensitivity and specificity (Fig. 4c). All participants had a sensitivity index ≥1 for test type A; four out of five participants (N1, N2, C1 and C2) also had the same index for test type B, demonstrating that our mHealth system was more effective than those participants at reading positive test results. Our system was also more reliable at reading negative tests, because all participants had a specificity index ≥1 for both types of HIV RDT.

Fig. 4 | Performance evaluation of our mHealth system compared to traditional visual interpretation: field pilot study. a, Graphics showing the agreement (%) between pairs of study participants when asked to interpret HIV RDT results using traditional visual interpretation. Participants were two experienced nurses (N1, N2) and three community health workers (C1, C2, C3). For each pair of participants there were $n=38$ HIV RDTs. Observations are separated according to the two types of HIV RDT used in the study. Purple-bordered area on both graphics highlights agreement between the two experienced nurses, while the orange-bordered area highlights agreement between the three pairs of community health workers. **b**, Confusion matrices showing the number of TN, FP, FN and TP results when comparing the interpretation of our mHealth system (top row) and traditional visual interpretation (bottom row) to the ground truth. Red matrices on the left include the results for all study participants, which are broken down into experienced nurses (orange matrices) and community health workers (purple matrices). **c**, Barplots showing the performance index for individual participants. Participants are divided between experienced nurses and community health workers. The performance index is the ratio of the performance of our mHealth system to that of traditional visual interpretation: performance index ≥1 indicates that our mHealth system performed better than (or as well as) traditional visual interpretation. The observations are separated according to the two types of HIV RDT used in the study.

We acknowledge the following limitations of our study. First, our pilot study involved a relatively small number of participants (five) although we note this is comparable to other similar pilot studies reported in the field. In future, larger evaluation studies and clinical trials will be needed to assess the performance of the system, involving participants with a broader range of demographics including age, gender and different levels of digital literacy, as well as more expert readers. In addition, future studies would benefit from the inclusion of an invalid test classifier and different mobile phone types with varying camera specifications. Although images were analyzed on a secure server, future analysis could be on-device and thus overcome the need to upload images. We are also currently investigating an image segmentation approach using deep learning for the next iteration of the smartphone application.

To conclude, we have demonstrated the potential of deep learning for accurate classification of RDT images, with an overall performance of 98.9% accuracy, notably higher than traditional visual interpretation of study partipants (92.1%), comparable to reports of 80–97% accuracy[17]. Given that >100 million HIV tests are performed annually, even a small improvement in quality assurance could impact the lives of millions of people by reducing the risk of FP and FN. We believe our real-world image library is the first of its kind at this scale and we demonstrate that deep learning models can be deployed with mobile devices in the field, without the need for cradles, dongles or other attachments. It lays the foundation for deep learning–enabled REASSURED diagnostics, demonstrating that RDTs linked to a mobile device could standardize the capture and interpretation of test results for decision makers, reducing interpretation and transcription errors and workforce training. Our findings are based on HIV testing decision support for fieldworkers, nurses and community health workers, but in future could be applicable to decision support for self-testing. We focused on HIV as an exemplar, but the capacity of the classifier for adaptation to two different test types suggests that it is amenable to a large range of RDTs spanning both communicable and noncommunicable diseases. This platform could be utilized for workforce training, quality assurance, decision support and mobile connectivity to inform disease control strategies, strengthen healthcare system efficiency and improve patient outcomes and outbreak management. The ideal connected system would link connected RDTs to laboratory systems, whereby remote monitoring of RDT functionality and utilization could also allow health programs to optimize testing deployment and supply management to deliver sustainable development goals and ensure that no one is left behind. The real-time alerting capability of connected RDTs could also support public health outbreak management by mapping ‘hotspots’ for epidemics, including COVID-19, to protect populations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-021-01384-9.

Received: 8 October 2019; Accepted: 6 May 2021;
Published online: 17 June 2021

References

1. Land, K. J., Boeras, D. I., Chen, X.-S., Ramsay, A. R. & Peeling, R. W. REASSURED diagnostics to inform disease control strategies, strengthen health systems and improve patient outcomes. *Nat. Microbiol.* **4**, 46–54 (2019).
2. *Second WHO Model List of Essential In Vitro Diagnostics* (WHO, 2019).
3. Peeling, R. W. Diagnostics in a digital age: an opportunity to strengthen health systems and improve health outcomes. *Int. Health* **7**, 384–389 (2015).
4. Ghani, A. C., Burgess, D. H., Reynolds, A. & Rousseau, C. Expanding the role of diagnostic and prognostic tools for infectious diseases in resource-poor settings. *Nature* **528**, S50–S52 (2015).
5. Figueroa, C. et al. Reliability of HIV rapid diagnostic tests for self-testing compared with testing by health-care workers: a systematic review and meta-analysis. *Lancet HIV* **5**, e277–e290 (2018).
6. Klarkowski, D. B. et al. The evaluation of a rapid in situ HIV confirmation test in a programme with a high failure rate of the WHO HIV two-test diagnostic algorithm. *PLoS ONE* **4**, e4351 (2009).
7. Gray, R. H. et al. Limitations of rapid HIV-1 tests during screening for trials in Uganda: diagnostic test accuracy study. *Brit. Med. J.* **335**, 188 (2007).
8. Martin, E. G., Salaru, G., Paul, S. M. & Cadoff, E. M. Use of a rapid HIV testing algorithm to improve linkage to care. *J. Clin. Virol.* **52**, S11–S15 (2011).
9. Cham, F. et al. The World Health Organization African region external quality assessment scheme for anti-HIV serology. *Afr. J. Lab. Med.* **1**, 39 (2012).
10. Galiwango, R. M. et al. Evaluation of current rapid HIV test algorithms in Rakai, Uganda. *J. Virol. Methods* **192**, 25–27 (2013).
11. Louis, F. J. et al. Evaluation of an external quality assessment program for HIV testing in Haiti, 2006–2011. *Am. J. Clin. Pathol.* **140**, 867–871 (2013).
12. Peck, R. B. et al. What should the ideal HIV self-test look like? A usability study of test prototypes in unsupervised HIV self-testing in Kenya, Malawi, and South Africa. *AIDS Behav.* **18**, 422–432 (2014).
13. Baveewo, S. et al. Potential for false positive HIV test results with the serial rapid HIV testing algorithm. *BMC Res. Notes* **5**, 154 (2012).
14. Crucitti, T., Taylor, D., Beelaert, G., Fransen, K. & Van Damme, L. Performance of a rapid and simple HIV testing algorithm in a multicenter phase III microbicide clinical trial. *Clin. Vaccine Immunol.* **18**, 1480–1485 (2011).
15. Tegbaru, B. et al. Assessment of the implementation of HIV-rapid test kits at different levels of health institutions in Ethiopia. *Ethiop. Med. J.* **45**, 293–299 (2007).
16. Johnson, C. C. et al. To err is human, to correct is public health: a systematic review examining poor quality testing and misdiagnosis of HIV status. *J. Int. AIDS Soc.* **20**, 21755 (2017).
17. Learmonth, K. M. et al. Assessing proficiency of interpretation of rapid human immunodeficiency virus assays in nonlaboratory settings: ensuring quality of testing. *J. Clin. Microbiol.* **46**, 1692–1697 (2008).
18. García, P. J. et al. Rapid syphilis tests as catalysts for health systems strengthening: a case study from Peru. *PLoS ONE* **8**, e66905 (2013).
19. Sacks, R., Omodele-Lucien, A., Whitbread, N., Muir, D. & Smith, A. Rapid HIV testing using DetermineTM HIV 1/2 antibody tests: is there a difference between the visual appearance of true- and false-positive tests? *Int. J. STD AIDS* **23**, 644–646 (2012).
20. Doan, M. & Carpenter, A. E. Leveraging machine vision in cell-based diagnostics to do more with less. *Nat. Mater.* **18**, 414–418 (2019).

21. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).
22. De Fauw, J. et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* **24**, 1342–1350 (2018).
23. Xu, Y. et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin. Cancer Res.* **25**, 3266–3275 (2019).
24. Silver, D. et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362**, 1140–1144 (2018).
25. Ascent of machine learning in medicine. *Nat. Mater*. **18**, 407 (2019).
26. Ching, T. et al. Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* **15**, 20170387 (2018).
27. Zeng, N., Wang, Z., Zhang, H., Liu, W. & Alsaadi, F. E. Deep belief networks for quantitative analysis of a gold immunochromatographic strip. *Cogn. Comput.* **8**, 684–692 (2016).
28. Carrio, A., Sampedro, C., Sanchez-Lopez, J. L., Pimienta, M. & Campoy, P. Automated low-cost smartphone-based lateral flow saliva test reader for drugs-of-abuse detection. *Sensors (Basel)* **15**, 29569–29593 (2015).
29. Neuman, M. et al. The effectiveness and cost-effectiveness of community-based lay distribution of HIV self-tests in increasing uptake of HIV testing among adults in rural Malawi and rural and peri-urban Zambia: protocol for STAR (self-testing for Africa) cluster randomized evaluations. *BMC Public Health* **18**, 1234 (2018).
30. Aicken, C. R. H. et al. Young people's perceptions of smartphone-enabled self-testing and online care for sexually transmitted infections: qualitative interview study. *BMC Public Health* **16**, 974 (2016).
31. Witzel, T. C., Weatherburn, P., Rodger, A. J., Bourne, A. H. & Burns, F. M. Risk, reassurance and routine: a qualitative study of narrative understandings of the potential for HIV self-testing among men who have sex with men in England. *BMC Public Health* **17**, 491 (2017).
32. Nsabimana, A. P. et al. Bringing real-time geospatial precision to HIV surveillance through smartphones: feasibility study. *JMIR Public Health Surveill.* **4**, e11203 (2018).
33. Laksanasopin, T. et al. A smartphone dongle for diagnosis of infectious diseases at the point of care. *Sci. Transl. Med.* **7**, 273re1 (2015).
34. Mudanyali, O. et al. Integrated rapid-diagnostic-test reader platform on a cellphone. *Lab Chip* **12**, 2678–2686 (2012).
35. Allan-Blitz, L.-T. et al. Field evaluation of a smartphone-based electronic reader of rapid dual HIV and syphilis point-of-care immunoassays. *Sex. Transm. Infect.* **94**, 589–593 (2018).
36. Feng, S. et al. Immunochromatographic diagnostic test analysis using Google Glass. *ACS Nano* **8**, 3069–3079 (2014).
37. Guan, Q. et al. Diagnose like a radiologist: attention guided convolutional neural network for thorax disease classification. Preprint at https://arxiv.org/abs/1801.09927 (2018).
38. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. & Chen, L.-C. MobileNetV2: inverted residuals and linear bottlenecks. Preprint at https://arxiv.org/abs/1801.04381 (2018).
39. Chaturvedi, S. S., Gupta, K. & Prasad, P. S. Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet. In *International Conference on Advanced Machine Learning Technologies and Applications* 165–176 (Springer, 2021).
40. Howard, A. et al. Searching for MobileNetV3. Preprint at https://arxiv.org/abs/1905.02244 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Ethics. Ethical approval for the demographic surveillance study was granted by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal, South Africa (no. BE435/17). Separate informed consent was required for the main household survey, the HIV sero-survey, the HIV point of care test and photographs of the HIV test.

Ethical approval for the collection of human blood samples used in the pilot study was granted by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal, South Africa (no. BFCJ 11/18).

Recruitment of participants to the Africa Health Research Institute Population Implementation Platform for the image library. Eligible participants were all individuals aged 15 years and older and resident within the geographic boundaries of the Africa Health Research Institute (AHRI) population intervention program surveillance area (see ref. [41] for the cohort profile). Individuals who had died or outmigrated before the surveillance visit were no longer eligible. There were three contact attempts by the fieldworker team and a further three contact attempts by a tracking team before an individual was considered uncontactable. All individuals in the study gave informed consent. Specifically, all contacted eligible individuals who gave informed consent for this study were offered a rapid HIV test if they were not currently being administered antiretroviral therapy. For children under the age of 18 years, written consent for rapid HIV testing was obtained from the parent or guardian and assent from the participant.

HIV RDT image library collection. The original RDT images library was collected in rural South Africa by a team of 60 fieldworkers between 2017 and 2019. AHRI fieldworkers survey a population of 170,000 people in rural KwaZulu-Natal. Participants were visited at their home, those giving informed consent were tested for HIV using a combination of two HIV RDTs and, following further consent, a photograph of their two HIV RDTs was captured by the fieldworker on a tablet at the time of interpretation. Both HIV RDTs were used as part of routine demographic surveillance in AHRI. The test type continued to change during this study following recommendations by the South African government, exemplifying the need for robust systems in reading multiple test formats.

While the two HIV RDTs used in this study have their own instructions for use (see manufacturer's instructions), they all generally follow the same principle of collecting a drop of blood from the participant's fingertip, delivering that drop of blood to the sample pad and using a drop of chase buffer to facilitate sample flow through the length of the paper strip. The result (a combination of one or two lines appearing on the paper strip) is then read out after a period of 10–40 min, depending on the type of HIV RDT used.

For minimal disturbance of workflow, a plastic tray designed to hold both HIV RDTs was given to each fieldworker (Fig. 2a). This ensured that fieldworkers were required to capture only one image per participant. The tasks of separating the two HIV RDTs and isolating the ROI used to train the classifier were conducted further down the line as part of data preprocessing.

A standard operating procedure (SOP) on how to capture the image was cocreated and optimized with the team of fieldworkers; a copy of the SOP can be found in Extended Data Fig. 1. The SOP was designed to minimize the impact of environmental factors, as well as to ensure a standard means of capturing images. All fieldworkers attended a 2-day initial training program during which the objectives of data collection and design of the plastic tray were clearly explained, and each fieldworker was personally trained and given feedback on how to capture valid photographs. A training protocol was also established to ensure that newly enrolled fieldworkers who did not attend the initial training session could also be trained to capture images for the project. Finally, picture quality assessment sessions were conducted to give the fieldworkers team feedback, and to ensure that most images were of sufficient quality for use in training the classifier.

All images were captured using Samsung tablets (SM-P585, 8-megapixel camera, f1/9 with autofocus capability) using the native Android camera application and stored on the device until the end of the day, when they were transferred to a secure database at AHRI. Our mHealth system allows the saving of only one picture per test and per participant to the tablet and uploading to the AHRI database. After anonymization (including stripping of geocoordinates from the image EXIF data), batches of 2,000–3,000 images were securely transferred to University College London team members on a quarterly basis and stored securely in a 'data-safe haven' managed by the university.

Levels of both feasibility (93%) and acceptability (98%) of the system used to capture HIV RDT images were high, according to a survey taken by fieldworkers involved in the study.

For the purposes of this study, an initial batch of 11,374 images were used. Because only very few invalid results were obtained from the field, it was decided, for the purposes of this proof-of-concept study, to focus on training the classifier to distinguish between positive and negative results. To optimize this task, the ROI around each HIV RDT was isolated and used to train the classifier.

Image labeling. All preprocessed images were labeled by a group of three RDT experts (99.2% agreement with fieldworkers' labeling). Labeling is the process of sorting images into categories, which are then used to train the classifier. The categories chosen here correspond to the possibilities for the HIV RDT result—that is, positive and negative. We recognize that a third outcome, 'invalid', is also possible and needs to be considered when using the system to provide a confident diagnosis. However, the absence of invalid test results in our library of images collected by fieldworkers did not allow us to train the classifier on this third category in the present study. We therefore focused training on the two main categories (positive and negative), and are exploring other ways to incorporate the invalid outcome in our mHealth system. This could mean either using data augmentation techniques on the low numbers of invalid test results images, or adding a preprocessing step to detect the presence of a control line on the image before deciding to feed it (or not, in the case where the control line is absent) to the classifier.

Training library. The labeled images were divided into two subcategories corresponding to the HIV RDT type. The two types of test in our library are:

- Type A: ABON HIV 1/2/O Tri-Line Human Immunodeficiency Virus Rapid Test Device (whole blood/serum/plasma) (ABON Biopharm (Hangzhou) Co., Ltd)
- Type B: ADVANCED QUALITY ONE STEP Anti-HIV (1&2) Test (InTec PRODUCTS, INC.).

While two tests were administered per patient, in this study we treat each test individually since the tests are from different manufacturers and therefore could respond differently to the same blood sample. The collection system design also guaranteed that there was never more than one image of a given test per participant.

Image normalization. Before being used for training, each image was resized to the dimensions of the input layer then standardized. Standardization of the data was performed using equation (1) below, where x_s is the standardized pixel value, x_o the original pixel value and μ and σ are the mean and s.d. of all pixels in the image, respectively.

$$x_s = \frac{x_o - \mu}{\sigma} \tag{1}$$

Cross-validation. Each dataset (one for each type of HIV RDT) was randomly divided into ten equal folds. Using the leave-one-out method, ten classifiers were trained using nine folds as the training set (further randomly divided into 80% training and 20% validation). To account for imbalanced datasets (roughly 13:1 negative:positive ratio), we forced every batch during training to contain 50% positive images and 50% negative images using random sampling. Each model was then optimized by creating a receiver operating characteristic curve using the validation set. This yielded an optimal threshold which was used to evaluate the model performance on the testing set (the remaining tenth fold). The deployment models were obtained by retraining using all the available data, for each type of HIV RDT. All training and evaluation were conducted using the scikit-learn and Tensorflow libraries in Python.

Comparison with established classification methods. The SVM was trained using preprocessed features extracted using the histogram of oriented gradients, with principal component analysis used to filter out less important features. The three CNNs (ResNet50, MobileNetV2 and MobileNetV3) were pretrained using the ImageNet dataset then retrained using our dataset. For all four methods, training and evaluation were conducted using the scikit-learn and Tensorflow libraries in Python.

Android application. We developed a smartphone/tablet Android application designed for end users to capture a picture of their HIV RDT at the time of reading of the test result. Together with end users, we optimized the design to maximize the simplicity of the process to make our mHealth system accessible to end users with a broad range of digital literacy. All that is required from the end user is to roughly align a semitransparent template of the HIV RDT with their HIV RDT and press a button to capture an image. Cropping around the ROI is then performed automatically in the background (using the pixel coordinates of the template overlay), as is the process of sending the ROI to our classifier and receiving our mHealth system result. For the purpose of this pilot study, participants were not made aware of our mHealth system's interpretation of the test results, to avoid bias for their own interpretation. Screenshots of the application can be found in Extended Data Fig. 2.

Field pilot study protocol. The Android application was deployed in a field pilot study in KwaZulu-Natal, South Africa. Five participants were randomly selected from the staff at AHRI—two experienced nurses and three community health workers. Forty HIV RDTs (20 type A, 20 type B) were performed following the manufacter's guidelines using discarded, anonymized human blood samples (ten positive, ten negative according to enzyme-linked immunosorbent assay). For each of the 40 HIV RDTs, every participant was asked to record their visual interpretation of the test result, then to use our mHealth system on a tablet to capture a photograph of the HIV RDT. The system consisted of our Android application (described above) installed on a single Samsung SM-P585 tablet, identical to those used by fieldworkers for data collection. Participants were not shown the automated interpretation of the test result provided by our mHealth system, to avoid confirmation bias. The field pilot study took place at the AHRI rural site in the heart of the community (Mtubatuba, KwaZulu-Natal) under lighting conditions identical to those under which the mHealth system is intended to be used. A short (10-min) demonstration on how to use the smartphone application was given to all participants, who were then left on their own to proceed with the task of reading the HIV RDTs and capturing images.

Field pilot study data analysis. The data analysis consisted of the comparison of three datasets:

1. Traditional visual interpretation by study participants
2. Independent expert interpretation of the images captured by study participants
3. Automated machine learning interpretation by our classifier.

Traditional visual interpretaiton was recorded on the tablet by each study participant immediately after being shown the HIV RDTs. Only two of the 40 HIV RDTs (corresponding to ten images out of 200) had to be discarded from the analysis, because one participant took a photograph of the wrong HIV RDTs and it was therefore not possible to compare interpretation results across all five participants.

An independent RDT expert subsequently visually interpreted all 190 HIV RDT images; this expert had substantial experience conducting performance evaluations of lateral flow rapid tests for ocular and genital *Chlamydia trachomatis* in the Phillippines, the Gambia and Senegal. Visual interpretation was performed 1–5 h after sample addition. The independent expert certified that none of the HIV RDT results had changed during this time frame.

The automated machine learning interpretation by our classifiers was processed on our secured server. The results were compared to traditional visual interpretation (shown in the confusion matrices in Fig. 4) while the independent expert then analyzed the results using the performance indicators described below.

Performance indicators. The four indicators of performance investigated were sensitivity, specificity, positive predictive value (PPV) and negative predicitve value (NPV). For each image, the classifier produces an outcome that belongs to one of the four categories TP, true negative (TN), FP or FN. Whether the outcome is true or false depends on comparison with the gold standard chosen.

Sensitivity is the ability of the classifier to correctly detect a positive result by measuring the ratio $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$, while the specificity is the ratio $\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$ and translates the ability of the classifier to correctly detect a negative result. PPV is the ratio $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$ and NPV is the ratio $\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}$. These indicate the proportions of positive and negative results, as determined by a diagnostic test, that are true positves and true negatives, respectively.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated during and/or analyzed during the current study are available from the AHRI data repository https://doi.org/10.23664/AHRI.M-AFRICA.2019.V1.

Code availability

Custom code used in this study is available at the public repository https://xip.uclb.com/product/classify_ai.

References

41. Gareta, D. et al. Cohort profile update: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *Int. J. Epidemiol.* **50**, 33 (2021).

Acknowledgements

We thank the community of the uMkhanyakude district and the study participants, as well as the AHRI team of fieldworkers and their supervisors. We thank A. Koza, Z. Thabethe, T. Madini, N. Okesola and S. Msane for their help with the pilot study; D. Gareta and J. Dreyer for IT support; V. Lampos and I. J. Cox for useful discussions; and E. Manning and J. McHugh for their help with editing and project management. This research was funded by the m-Africa Medical Research Council GCRF Global Infections Foundation Award (no. MR/P024378/1, to C.H., D.P., K.H., M.S., R.A.M. and V.T.) and is part of the EDCTP2 program supported by the European Union, i-sense Engineering and Physical Sciences Research Council Interdisciplinary Research Collaboration (EPSRC IRC) in Early Warning Sensing Systems for Infectious Disease (no. EP/K031953/1, to R.A.M., V.T., D.P., S.M., S.G., N.A. and M.S.), the i-sense: EPSRC IRC in Agile Early Warning Sensing Systems for Infectious Diseases and Antimicrobial Resistance (no. EP/R00529X/1, to R.A.M., V.T., D.P., S.G., N.A. and S.M.) and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre (R.A.M. and S.M.). We thank the m-Africa and i-sense investigators and advisory boards. The AHRI is supported by core funding from the Wellcome Trust (core grant no. 082384/Z/07/Z, to T.S., D.P. and K.H.).

Author contributions

V.T. and R.A.M. wrote the manuscript with input from coauthors. V.T., C.H., T. Mngomezulu, N.D. and T. Mhlongo collected field data. V.T. and S.M. developed the machine learning models with contributions from V.C., K.S., S.G. and R.A.M. V.T., N.A. and J.B. were involved in manual data preprocessing. K.H. oversaw data collection and management. T.S. and M.S. provided access to anonymized blood samples used in the pilot study. R.A.M., V.T., M.S., K.H. and D.P. conceived the overall project, designed the study and secured funding. R.A.M. was the principal investigator with overall responsibility for the i-sense EPSRC IRC and m-Africa programs, and was supervisor of the research associates (V.T., S.M. and N.A.) and students (V.C., K.S. and J.B.) involved in this study.

Competing interests

The authors declare no competing interests.

Additional information

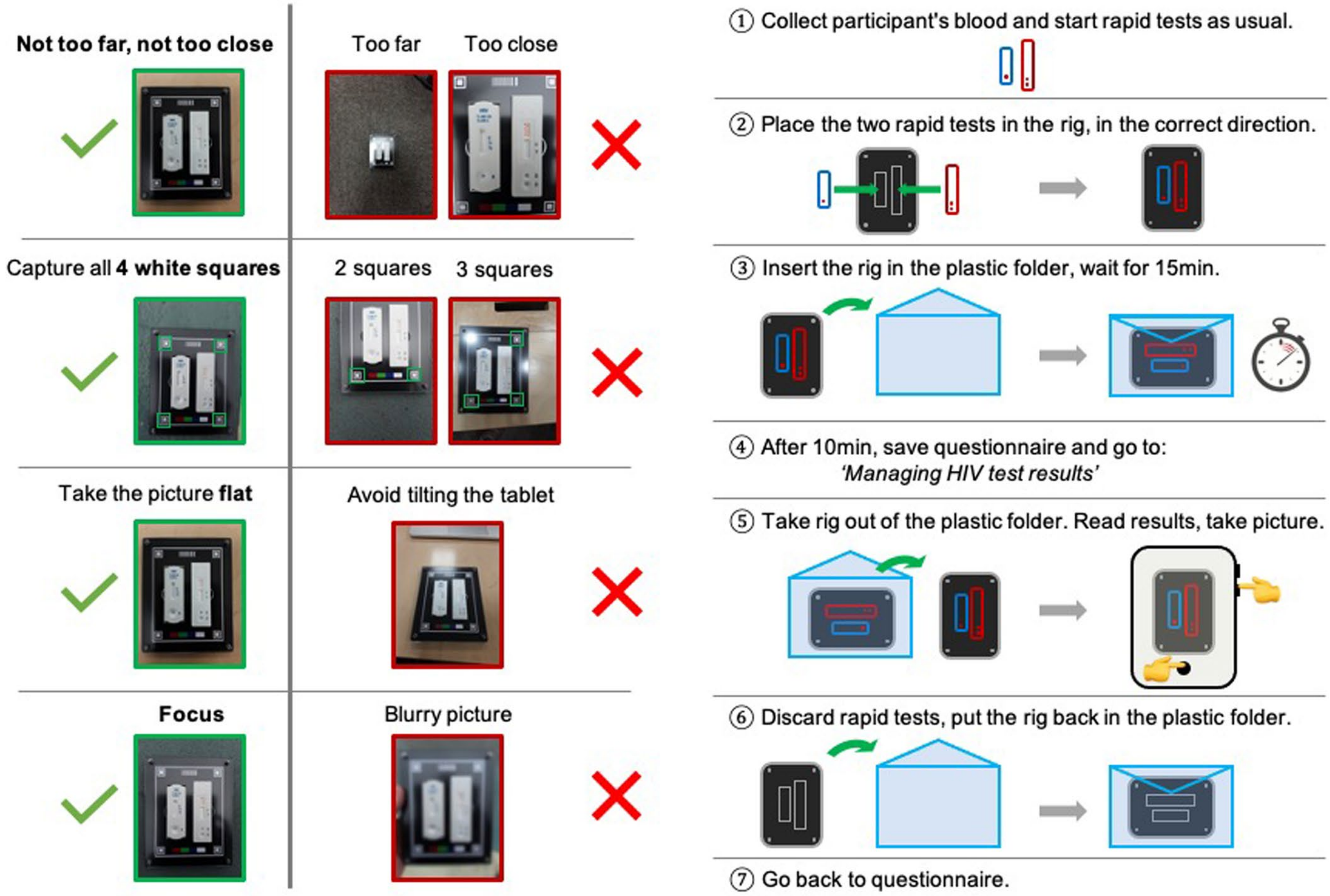
Extended data is available for this paper at https://doi.org/10.1038/s41591-021-01384-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-021-01384-9.

Correspondence and requests for materials should be addressed to V.T., K.H., M.S. or R.A.M.

Peer review information *Nature Medicine* thanks Nicholas Durr and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Michael Basson was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.



① Collect participant's blood and start rapid tests as usual.

② Place the two rapid tests in the rig, in the correct direction.

③ Insert the rig in the plastic folder, wait for 15min.

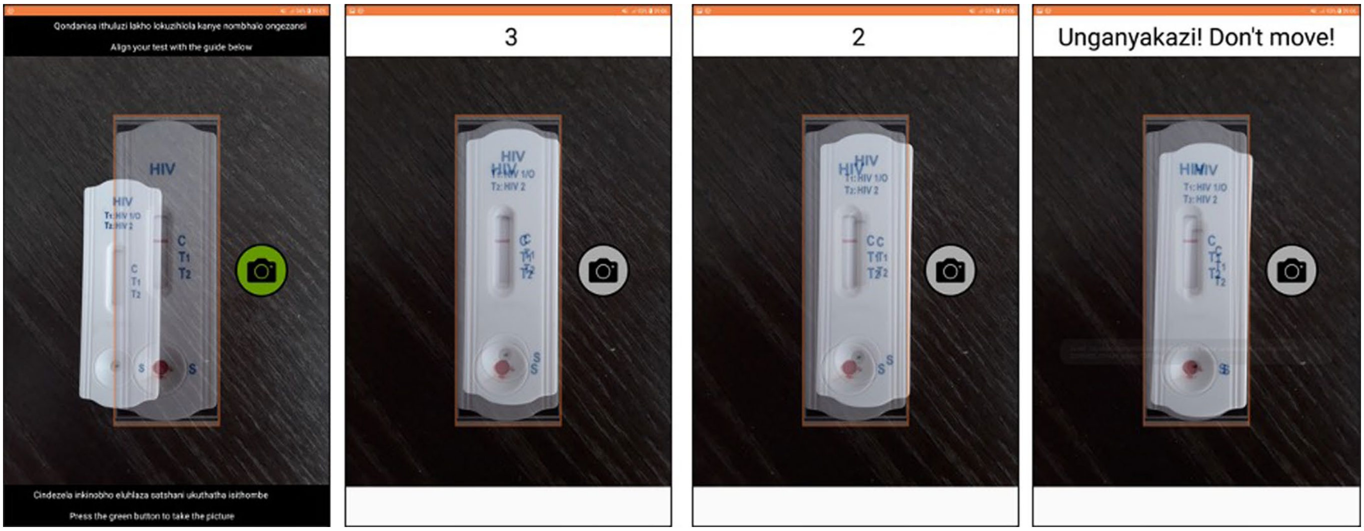
④ After 10min, save questionnaire and go to: *'Managing HIV test results'*

⑤ Take rig out of the plastic folder. Read results, take picture.

⑥ Discard rapid tests, put the rig back in the plastic folder.

⑦ Go back to questionnaire.

Extended Data Fig. 1 | Standard Operating Procedure for HIV RDT image collection. Document used for training and distributed to all AHRI fieldworkers involved in data collection. Left-hand side: example of valid and invalid photographs. Right-hand side: step-by-step guidelines for capturing pictures of HIV RDTs.



Qondanisa ithuluzi lakho lokuzihlola kanye nombhalo ongezansi

Align your test with the guide below

Cindezela inkinobho eluhlaza satshani ukuthatha isithombe

Press the green button to take the picture

3

2

Unganyakazi! Don't move!

Extended Data Fig. 2 | Screenshots of the Android application, to illustrate the capture of the HIV RDT image at the time of reading the test result. Images were captured sequentially from left to right. The end user is asked to align the test with the overlay on the screen, then continuously press the capture button for 3 seconds, after which the image is automatically captured and processed to extract the ROI. The 3 seconds press feature was implemented as a result of consultation with end users in the optimisation phase of the app development.

nature research

Corresponding author(s): McKendry, Rachel

Last updated by author(s): Mar 9, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
□	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	□	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	□	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	□	A description of all covariates tested
☒	□	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
□	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	□	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	□	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	□	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	□	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Pictures of HIV RDT were collected in the field using the native camera application on Samsung tablets (SM-P585, 8MPixels camera, f1/9, with autofocus capability). Pictures and data for the pilot study were collected using a custom made Android application available on request.
Data analysis	The code used for data analysis is available on the public repository: https://github.com/i-sense-epsrc/mAfrica-deep-learning-HIV'. Analysis was written using python 3.6.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data availability

The datasets generated during and/or analysed during the current study are available from the AHRI data repository:
Herbst, K., & McKendry, R. (2019). m-Africa: Building mobile phone-connected diagnostics and online care pathways for optimal delivery of population HIV testing,

prevention and care in decentralised settings (Version 1) [Data set]. Africa Health Research Institute (AHRI). https://doi.org/10.23664/AHRI.M-AFRICA.2019.V1

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

□ Life sciences ☒ Behavioural & social sciences □ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

The pilot study presented in this manuscript is a qualitative study, looking at a direct comparison between traditional visual interpretation of HIV RDTs and computer vision system performing the same task of interpreting the result of HIV RDTs. A comparison of traditional visual interpretation between study participants is also presented.

Research sample

AHRI fieldworkers survey a population of 140,000 people in rural KwaZulu-Natal. Eligible participants are all individuals age 15 years and older resident within the geographic boundaries of the AHRI population intervention programme surveillance area (Ref Gareta, D., Baisley, K., Mngomezulu, T., Smit, T., Khoza, T., Nxumalo, S., ... & Herbst, K. (2021). Cohort profile update: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. International journal of epidemiology, 50(1), 33.).

The survey population is not a sample and consists of members of all households resident within the boundary of the health and demographic surveillance area. This survey population was selected because of the large scale HIV-serosurveilance that have been conducted in this population over the last 15 years, including the use of POC rapid HIV testing by lay fieldworkers since 2017. Individuals who have died or outmigrated prior to the surveillance visit are no longer eligible.

There are three contact attempts by the fieldworker team and a further three contact attempts by a tracking team before the individual is considered to be uncontactable. All contacted eligible individuals are consented and offered a rapid HIV test if they are not currently on anti-retroviral therapy.

Sampling strategy

When separating the datasets used in training the classifiers into 10 folds (following a common approach in the field of machine learning), the dataset was divided into 2 groups depending on the type of HIV RDT used, then each group was divided in two categories depending on the label of the image (Negative or Positive). The folds were then populated randomly.

Data collection

Training library data: Participants were visited at their home, those giving informed consent were tested for HIV using a combination of two HIV RDT, and upon further consent, a picture of their two HIV RDTs was captured by the fieldworker on a tablet at the time of interpretation. Researchers were not present for the capture of the training library data.
Pilot study data: Participants were left to perform the test readings and use the tablet application on their own. Researchers were not blind to HIV RDTs results as they ran them beforehand, but they but did not intervene beyond providing the HIV RDTs used for the study.

Timing

The original RDT images library was collected in rural South Africa by a team of 60 fieldworkers (between 2017 and 2019).

Data exclusions

As only very few invalid results were obtained from the field, it was decided, for the purpose of this proof of concept study, to focus on training the classifier to distinguish between positive and negative results.

Non-participation

Training library: Over the period of the study, the consent rate for being tested for HIV was approximately 27%. Of those who consented to take the RDT, approximately 95% consented for a picture of their RDT to be captured and used for research.
Pilot study: all participants accepted and fully consented to taking part in the pilot study. No participant dropped out or declined.

Randomization

All 5 participants performed the same task, in the same experimental conditions. In the analysis of Figure 4b, the data analysis looks at the results as whole, but also depending on the study participants professional group (either professional nurse or community health worker).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics

Eligible participants are all individuals age 15 years and older resident within the geographic boundaries of the AHRI population intervention programme surveillance area (Cohort profile: Africa Centre demographic information system (ACDIS) and population-based HIV survey. International journal of epidemiology. 2007 Nov 12;37(5):956-62.). Individuals who have died or outmigrated prior to the surveillance visit are no longer eligible. The population pyramid can be found in Figure 2 of the above publication.

Recruitment

Training library: There are three contact attempts by the fieldworker team and a further three contact attempts by a tracking team before the individual is considered to be uncontactable. All contacted eligible individuals are consented and offered a rapid HIV test if they are not currently on anti-retroviral therapy. The following study conducted in the same region suggests that HIV-infected individuals are less likely to participate than HIV-uninfected persons to participate in HIV surveillance because they fear the negative consequences of others learning about their HIV infection. It also concludes that the increased knowledge of HIV status that accompanies improved ART access can reduce surveillance participation of HIV-infected persons, but that this effect decreases after ART initiation, in particular in successfully treated patients. (Bärnighausen, T., F. Tanser, A. Malaza, K. Herbst, and M-L. Newell. "HIV status and participation in HIV surveillance in the era of antiretroviral treatment: a study of linked population-based and clinical data in rural South Africa." Tropical Medicine & International Health 17, no. 8 (2012): e103-e110.")

Pilot study: Five participants were randomly selected from the staff at AHRI – two experienced nurses and three community healthworkers.

Ethics oversight

Ethical approval for the demographic surveillance study was granted by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal, South Africa, Reference Number BE435/17. Separate informed consent is required for the main household survey, for the HIV sero-survey, the HIV point of care test and the photographs of the HIV test. For children under the age of 18, written consent for Rapid HIV testing is obtain for the parent or guardian and assent from the participant.

Ethical approval for the collection of human blood samples was granted by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal, South Africa, Reference Number BFCJ 11/18.

Note that full information on the approval of the study protocol must also be provided in the manuscript.