Supplementary information: Machine learning to support visual auditing of home-based lateral flow immunoassay self-test results for SARS-CoV-2 antibodies

Supplementary Information 2 - Datasets and Machine Learning Details

Table S2-1: The 1D CNN architecture. "conv" is 1D convolution (no. of inputs, no. of outputs, filter length), "FC" (No. of inputs, no. of outputs), "batchnorm" are 1D batch normalisations and "MaxPool" are 1D max pooling functions with kernel size of 2. Note that "Input(10)" refers to 10 projection signatures of length 100, please look at Figure S1-3.

Model architecture
Input (10)
Conv(10,30,5)
Batchnorm + MaxPool
Conv(30,30,5)
Batchnorm + MaxPool
Conv(30,20,5)
Batchnorm + MaxPool
FC(180,170)
FC(170,70)
FC(70,1)
Sigmoid

Appendix S2-A – Segmentation

A deep Convolutional Neural Network (CNN), based on a U-Net architecture[1], was trained to identify candidate regions of interest (ROI). Training of this network, known as *dhSegment*[2], involves using pre-trained weights for the "encoding part" of the network and fine-tuning the "expansive part" of the network (which is mapping the encoded feature maps to full resolution feature maps). The raw output undergoes post-processing consisting of connected component analysis and small region removal, providing a simplified output segmentation. The network is implemented in TensorFlow.

We randomly selected an initial 498 LFIA images for developing the segmentation CNN (Originally 500, however there were file format issues with two images). Regions were manually labelled using VGG Image Annotator (VIA)[3] by 2 people, providing a representative sample for training and testing. We split the dataset into fixed train, test, and validation sets: 373 for training, 42 for validation and 83 for testing. Dice scores were found to be: 0.973 for the LFIA cassette, 0.943 for the read-out window, and 0.921 for the blood well.

Appendix S2-B – Classification ground truth

Ground truth for the read-out status was obtained by having a team of 6 trained observers review images. Since weakly positive cases are rare, the training dataset evolved over 8 months, and particularly improved – during Phase 3 – with the inclusion of weaker cases of immunity from vaccinated individuals (Main manuscript, Table 2). This was important as the weakly positive cases were identified as a source of false negative readings by participants of the REACT 2 study when assessed by human experts.

Appendix S2-C – Classification 1D CNN (Phase 2 Read-out Training)

A popular deep neural network architecture for image classification is the 2D CNN. However, as shown in Supplementary Information 1 (Figure S1-3), the LFIA read-out window has a strong linear structure that can be collapsed into 1D signatures. 1D signatures are easily human interpretable, providing semi-quantitative amplitudes, which will be used in subsequent work. Peak detection algorithms can also be applied to detect locations corresponding to lines representing control and readout, with minimal data. 1D CNNs typically consume far less data for training than 2D CNNs, making them suitable to bootstrap the labelling process, and particularly for finding rare but important cases to balance larger training sets to support more sophisticated architectures. We used the ten signatures (Supplementary Information 1, Figure S1-3) to explore five different candidate 1D CNNs (implemented in Pytorch) varying in the number of convolutional filters, filter lengths and number of layers. The final 1D CNN architecture selected for read-out interpretation is shown in Table S2-1.

Appendix S2-D – Classification 2D CNN (Phase 3)

A second 2D CNN (implemented in Tensorflow) is trained for accurate read-out interpretation and yielded the best sensitivity. This CNN utilises a MobileNetV2 architecture with pretrained weights learned on the ImageNet dataset. The final 2D network is based on previous work by the McKendry group and the i-sense interdisciplinary research collaboration (IRC) (www.i-sense.org.uk) at University College London; the main paper describes performance in two separate experiments (CE1 and CE2) which tradeoff slightly the specificity and sensitivity of readout interpretation.

The 2D CNNs for segmentation and classification, as well as the (more recent) transformer network, were trained and fine-tuned on an RTX6000 with 24GB of memory. Training time is usually no more than half a day for any one network.

References

1. Ronneberger, O., Fischer, P. & Brox T. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, 234-241 (2015).
2. Oliveira, S.A., Seguin, B. & Kaplan, F. dhSegment: A generic deep-learning approach for document segmentation. In 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR) 2018 Aug 5 (pp. 7-12). IEEE.
3. Dutta, A. & Zisserman, A. The VIA annotation software for images, audio and video. In Proceedings of the 27th ACM international conference on multimedia 2019 Oct 15 (pp. 2276-2279) https://doi.org/10.1145/3343031.3350535.
4. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. & Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition 2009 Jun 20 (pp. 248-255).