Classification of nucleic acid amplification on ISFET arrays using spectrogram-based neural networks

Prateek Tripathi [a,*], Costanza Gulli [a], Joseph Broomfield [a,b], George Alexandrou [a], Melpomeni Kalofonou [a], Charlotte Bevan [b], Nicolas Moser [a], Pantelis Georgiou [a]

[a] Centre for Bio-Inspired Technology, Department of Electrical and Electronic Engineering, Faculty of Engineering, Imperial College London, SW7 2AZ, London, UK
[b] Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, SW7 2AZ, London, UK

ARTICLE INFO

ABSTRACT

The COVID-19 pandemic has highlighted a significant research gap in the field of molecular diagnostics. This has brought forth the need for AI-based edge solutions that can provide quick diagnostic results whilst maintaining data privacy, security and high standards of sensitivity and specificity. This paper presents a novel proof-of-concept method to detect nucleic acid amplification using ISFET sensors and deep learning. This enables the detection of DNA and RNA on a low-cost and portable lab-on-chip platform for identifying infectious diseases and cancer biomarkers. We show that by using spectrograms to transform the signal to the time–frequency domain, image processing techniques can be applied to achieve the reliable classification of the detected chemical signals. Transformation to spectrograms is beneficial as it makes the data compatible with 2D convolutional neural networks and helps gain significant performance improvement over neural networks trained on the time domain data. The trained network achieves an accuracy of 84% with a size of $30\,kB$ making it suitable for deployment on edge devices. This facilitates a new wave of intelligent lab-on-chip platforms that combine microfluidics, CMOS-based chemical sensing arrays and AI-based edge solutions for more intelligent and rapid molecular diagnostics.

1. Introduction

The COVID-19 pandemic has highlighted the need for rapid and accurate diagnosis of infectious diseases. AI-based edge solutions have shown significant promise towards the real-time classification of sensor data and have the potential to realise fully integrated sensing and processing systems which operate intelligently with their data. The use of Lab-on-Chip (LoC) devices can enable tests to be brought to the point of need and therefore stand to benefit from the use of edge-based AI. Towards this goal, this work introduces a new methodology involving spectrograms and AI for portable CMOS-based LoC diagnostic platforms that can rapidly identify infectious diseases and cancers by using arrays of electrochemical sensors.

Infectious disease detection commonly relies on identifying two types of biomarkers. The first involves, detecting antigens specific to the pathogen and is achieved by reading the signal associated with antigen–antibody binding. Lateral flow devices (LFD) are the most accessible readout technology, achieving low-cost diagnostics at the point of care, thanks to their simplicity of use and manufacture, at the expense of accuracy of the test [1]. The second involves, detecting nucleic acids using molecular methods to target and amplify a known

sequence. Quantification can be achieved by analysing the speed of the reaction. Common nucleic acid amplification tests (NAATs) include the polymerase chain reaction (PCR) and loop-mediated isothermal amplification (LAMP). PCR is the gold-standard laboratory technique and is preferred for high-accuracy diagnostics. Both techniques are highly specific and commonly only result in amplification when their specific nucleic acid target sequence is present in the solution.

There are several readout methods associated with biomarker detection. Fluorescence methods are currently very common and well suited for applications in specialised laboratories as they rely on bulky and expensive optical instrumentation [2]. However, they fall short of the requirements of point-of-care diagnostics in low-resource settings. In this context, electrochemical sensing in CMOS technology is a promising alternative that leverages the economies of scale of the semiconductor industry and combines scalability, miniaturisation and high accuracy. Ion-Sensitive Field-Effect Transistors (ISFETs) [3] are electrochemical sensors which carry the potential to bring the accuracy of molecular methods with the portability of the LFD by leveraging the evolution of CMOS technology in accordance with Moore's law [4]. IS-FET sensors were first introduced by Bergveld in 1972 [3] by replacing
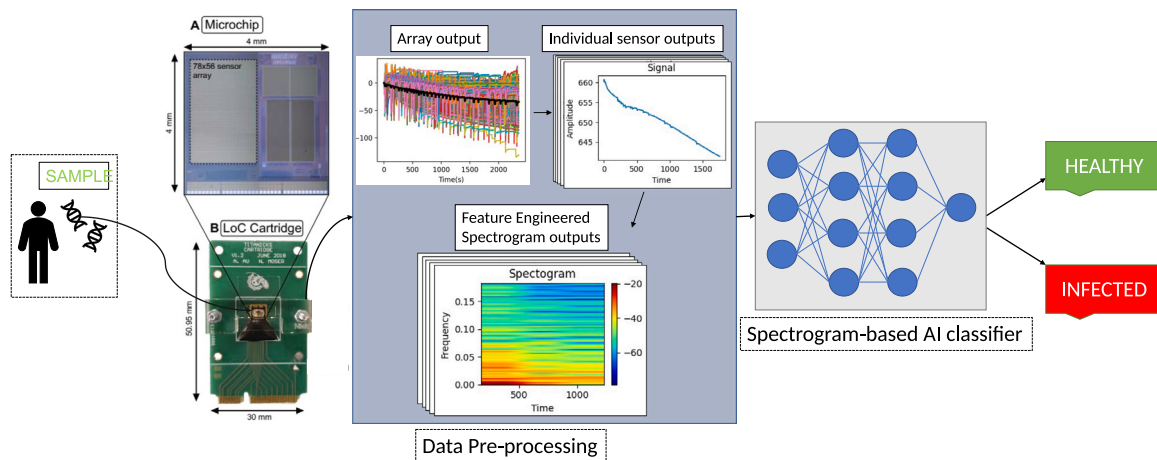

Fig. 1. Concept workflow for the proposed method: ML for infectious disease diagnostics using LoC platform. Sample from the patient is loaded onto microchip A in cartridge B. The data is preprocessed and passed through a spectrogram-based AI classifier to identify the presence of an infection.

the gate oxide of a transistor with an insulating membrane for sensing pH changes in the solution at the sensor surface. Once their implementation was achieved in unmodified CMOS technology [5], ISFETs have been adopted as an industry-standard method for next-generation DNA sequencing (Ion Torrent) [6] and recently have been reported for infectious disease and cancer diagnostics [7–9]. Coupled with LAMP, this technology allows for the detection of amplification in samples of small volume [10] and with minor alterations to the assay (the reaction mix is only weakly buffered to allow for proton release during the reaction) without the need for complex thermocycling. ISFETs also allow for the integration of instrumentation within the same system-on-chip. This technique has previously been used to mitigate device non-idealities (such as offset and drift) and further allow for accurate detection.

ISFETs suffer from non-idealities with drift posing the main challenge to extract the signal generated by nucleic acid amplification. The frequency range of drift and the signal of interest are within 10 mHz to 100 mHz approximately. So far previous works have focussed on standard signal processing methods, identifying an inflection point in the time-series output signal as the amplification event [11,12]. However, this method is compromised by a high drift rate and low signal amplitude and is highly sensitive to any instability or additional noise in the sensor output.

Deep neural networks are a popular tool in the field of healthcare, with medical imaging applications ranging from diagnostics to assistive surgery [13–18]. Deep learning [19] has become one of the greatest success stories in the field of computer vision [20]. However, a vast amount of data is required to effectively train neural networks [21, 22]. This highlights a huge challenge in data collection for the deep learning community, as the crucial role of patient privacy and security needs to be respected [23]. Deep learning is helping physicians by demonstrating promising results for complex diagnostics in radiology [24], dermatology [25,26], ophthalmology [27], ureteroscopy [16] and pathology [28]. These advances bring forth the opportunity to extend deep learning solutions to point-of-care diagnostic devices for infectious diseases as illustrated in Fig. 1.

This paper proposes a framework to efficiently identify nucleic acid amplification from ISFET arrays, towards a fully integrated, portable diagnostic platform, using a novel combination of spectrogram processing and neural networks. It is the first proof-of-concept study for the use of deep learning techniques to improve on the accuracy of classification of ISFET-based nucleic acid amplification data by relying on its frequency spectrum. A dataset of nucleic acid amplification was selected from two categories of diagnostic assays, namely SARS-CoV-2 and three cancer biomarkers. To benchmark our work in the absence of earlier research on this dataset, a time-based approach is first proposed and then

compared against the spectrogram approach to understand how neural networks can best visualise ISFET data for biomarker detection. Finally, an edge implementation with the potential to become a tool in routine diagnostics and future pandemics is proposed. The contributions of this work include the following:

- a benchmark for combining the field of deep learning with point-of-care lab-on-chip (LoC) diagnostics for infectious diseases and cancer
- a novel spectrogram-based approach that facilitates the application of image-processing ML techniques, including 2D-CNNs, for the classification of ISFET data
- an approach that can help train neural networks with a limited dataset of infectious diseases for high precision and accuracy
- a framework for machine intelligent LoC diagnostic platforms by leveraging the combined advances in the fields of spectrogram, deep learning and electrochemical-based detection to (1) improve upon the diagnostic accuracy of conventional methods, (2) transfer learning for faster deployment of tests, (3) reduce bandwidth requirement for sensor data transfer, (4) reduce carbon footprint, (5) have a lower cost and (6) promote data privacy and security.

2. Related work

2.1. ISFETs

$$V_{th(ISFET)} = V_{th(MOSFET)} + \gamma + \alpha S_N \; pH \tag{1}$$

Various circuits have been proposed to achieve sensor readout, i.e. to translate the threshold voltage variation into a signal. An exhaustive review can be found in [29]. In the context of this paper, the ISFET interface leads to an output signal V_o that is linearly related to pH.

Measurement of pH in a solution allows for the detection of DNA amplification on-chip: this is because DNA amplification results in a release of protons during nucleotide incorporation (Eq. (2)). Hence, detecting a pH change indicates the amplification of DNA on the chip surface [7,30].

$$DNA_n + dNTP \rightarrow DNA_{n+1} + PP_i + H^+ \tag{2}$$
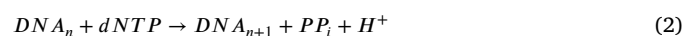
However, despite the great potential, ISFETs present a number of non-idealities that set great challenges for the interpretation of the data. Apart from the electrical noise, ISFETs suffer from certain sources of noise that are specific to the technology. Refer to [29] for an exhaustive discussion on the limitations of ISFET technology in relation to the physical properties of the sensor. In this context, the impact of the non-idealities on the output signal is described as:
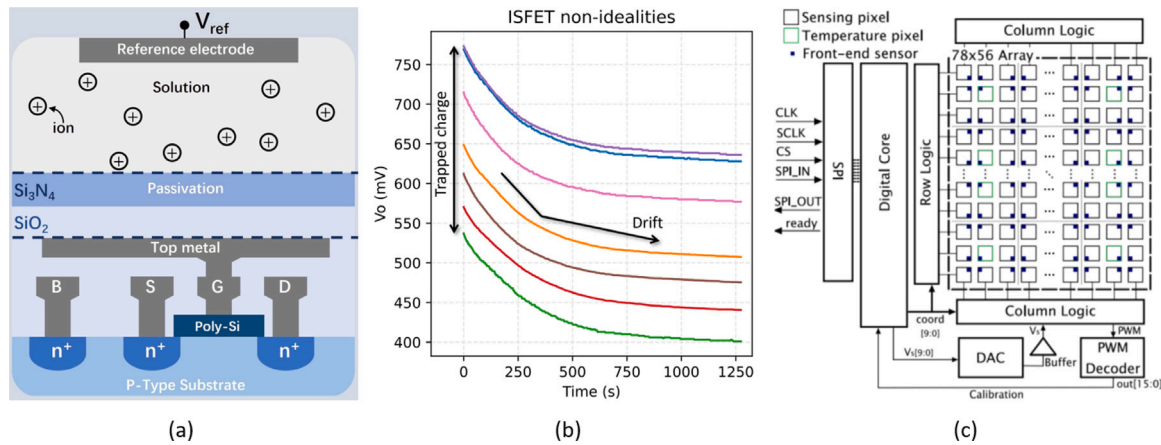
Fig. 2. ISFET overview. (a) Structure of an ISFET sensor fabricated in unmodified CMOS technology [5]. (b) Illustration of main sensor non-idealities (different curves are output from sensors on the same array). (c) Schematic of the circuit required for array integration of the ISFET.

- *Trapped charge* during fabrication at the interfaces of the sensing materials [31] causes an offset of the threshold voltage of the transistor. This offset induces a mismatch between sensors and may drive some out of the range of readout Fig. 2(b).
- *Drift* is a modification of the threshold voltage of the transistor over time, caused by the chemical interaction between the ISFET surface and the solution [32]. It is a complex phenomenon that is challenging to model mathematically because of its highly unpredictable behaviour. A dependence upon several factors has been shown, such as surface geometry and material, sensor size, pH of the solution and readout configuration [33]. In practice, drift results in a monotonic exponential decay in the signal, as shown in Fig. 2(b). The drift behaviour has been shown to demonstrate high stochasticity, and may often exceed the chemical signal in amplitude, adding to the challenge of output classification.

The limitations in ISFET performance motivate the integration of the sensors onto large arrays [4]. This allows the collection of a large amount of spatio-temporal data for each experiment to compensate for the inaccuracies of the individual sensor outputs. The output can be interpreted as a series of frames of pH values, where each frame is equivalent to a chemical image and each sensor acts as a pixel. The data used in this paper is the output of a 78×56 sensing array fabricated in 350 nm CMOS technology with 1 μm SiO_2 with 1 μm SiN_4 on top as the sensing surface obtained using Serial Parallel Interface (SPI) as shown in Fig. 2(c).

2.2. AI for ISFETs

AI techniques are slowly becoming a topic of interest in the field of ISFET sensing. Whilst circuit design has been the foremost method of improving sensor performance [29], recent efforts have been reported using deep learning networks and neuromorphic architectures to combat ISFET non-idealities. The use of Multi-layer perceptron, linear regression, support vector machine, decision trees, random forests, LSTM and GRU have been proposed to compensate for temperature and temporal drift in ISFETs [34,35]. Virtual training feature generation and subsequent training of SVM and ANNs have also been proposed for separating light and pH signals in dual-gated ISFETs [36]. An alternative approach is embedding neuromorphic architectures and using backpropagation as part of the sensor front-end to help compensate for non-ideal effects [37] by using the spatial correlation among neighbouring sensors [38]. While the majority of these efforts have been focused on improving sensor performance, we propose a method to extend the benefits of AI to be used as a classification tool for the ISFET data.

2.3. Spectrograms and AI

Section 2.1 presents a possible analogy between the ISFET array output and a set of frames, i.e. a video. In contrast, we propose an alternative analogy between an ISFET output and an audio signal: where each sensor output is treated independently as a time series. This allows us to explore an entirely new set of processing techniques for experiment classification. When treating audio signals, the choice of how to treat a signal is non-trivial. The signal can be processed in the original time-domain or transformed with a number of methods, including Mel Frequency Cepstral Coefficients (MFCCs), Magnitude spectra and Spectrograms [39,40]. Spectrograms have been used for animal audio classification in combination with Siamese neural networks (SNN), clustering techniques and Support vector machines (SVM) [41]. In the context of cough recognition from audio recordings, Mel-spectrograms have been proposed as data pre-processing technique for CNN classification [42].

The same question has been raised for ECG signals, where 99% classification accuracy in arrythmia classification has been shown by transforming the signal in time–frequency domain with Short-Time Fourier Transform (STFT)-based spectrograms, and using 2D image processing techniques for classification [43]. Similarly for EEG signals, stacked multi-channel EEG spectrograms have been used for training DCNNs for REM Behaviour Disorder (RBD) diagnosis [44]. The EEG spectrograms have also been used for Autism Spectrum disorder using SVM classifiers [45]. The successful application of these algorithms to biological signals poses the justification for the approach proposed in this paper, that is employing spectrograms and CNNs on ISFET outputs for classification.

3. Data collection

Nucleic acid amplification tests, such as the polymerase chain reaction (qPCR), result in amplification events where a specific nucleic acid sequence is detected. These amplification events result in the formation of large quantities of DNA. With each addition of a nucleic acid to a double-stranded DNA sequence, one proton is released into the solution [46]. qPCR usage is limited within point-of-care settings, predominantly due to the requirement for thermal cycling [47]. Loop-mediated isothermal amplification (LAMP) is a rapid, isothermal and quantitative amplification technique [48]. The addition of reverse transcriptases into LAMP assays (RT-LAMP) also results in the detection of RNA sequences. As a result, RT-LAMP has been utilised in the detection of SARS-CoV-2 mRNA, prostate cancer mRNA and influenza mRNA for diagnostics [7,49,50]. Augmentation of LAMP assays to generate a pH readout has previously been established [51]. This work utilises augmented RT-LAMP assays (RT-pHLAMP) for detecting the N gene from

Table 1

Dataset distribution.

Biomarker	No. of samples	No. of active pixels	In training	In testing
N gene (SARS-CoV-2)	24	23193	✓	✓
Cancer marker 1 mRNA	9	14198	✓	✓
Cancer marker 2 mRNA	12	24259	✓	✓
Cancer marker 3 DNA	4	5320	X	✓
Negatives	34	43876	✓	✓
Dataset-1 (Negatives, Cancer Marker 1 and 2 mRNA)	5	7897	X	✓

SARS-CoV-2. The specificity of this assay has been robustly ratified [7]. Cancer markers 1,2 (mRNA) [8] and 3 (DNA) [9] show potential as circulating biomarkers for cancer diagnosis and prognosis. As such, RT-pHLAMP and pHLAMP assays for these markers were implemented on the ISFET sensing array; their data is also included in the dataset.

Microfluidic manifolds were used to house the reaction mixtures over the ISFET sensing array. SARS-CoV-2 samples were contained within a bio-compatible resin (MED-AMB-10) with two 5 μL chambers, one with a positive sample (i.e, would render a release of protons) and one negative sample. The RT-pHLAMP and pHLAMP assays for cancer marker detection were housed in an acrylic manifold with one 20 μL chamber. Positive and negative samples in this instance were run separately. Each chamber of each assay type was affixed to the ISFET sensing array with an adhesive gasket (Double-sided smooth lamination filmic tape, Tesa®) to avoid leakage of the RT-pHLAMP or pHLAMP solution. The reference electrode (AgCl/Ag, 0.03 mm chloridised silver wire) was secured between the adhesive gasket and the ISFET sensing array. A 100 mV voltage spike from the reference electrode occurs at the beginning of the experiment. This indicates if the reference electrode is connected to the circuit and illustrates which pixels are responding to variations in pH change. The reaction was heated to 63 °C with a peltier heating module in contact with the chip.

The chambers were initially filled with nuclease-free water for 700 s to set a common voltage for the array surface. At 700 s the water was replaced with the RT-pHLAMP assay solution. All positive reactions were then run for a further 30 min. SARS-CoV-2 negative reactions were run for 30 min, negative cancer biomarker assays were run for 20 min. Data was recorded in real-time on a mobile phone. Table 1 presents the dataset distribution between training and testing based on the experiments.

4. Method and results

4.1. Data pre-processing

Data preparation is required to address sensor non-idealities prior to network classification. The steps proposed in this method are individually discussed in the following paragraphs and illustrated in Fig. 3: identification of active pixels, extraction of relevant time interval, trapped charge compensation, spectrogram generation, and high-frequency noise compensation.

Identification of active pixels. These filtering steps are computationally inexpensive operations that allow filtering of the sensors that carry significant information about the chemical reaction in the well. Pixels are generally considered as inactive if they are covered by the microfluidic manifold or out of the readout range due to non-ideal effects like trapped charge and drift. These pixels are identified as the sensors that respond to a variation in reference electrode voltage at the beginning of each experiment. Moreover, the offset caused by uneven trapped charge in every sensor causes some of the pixels to have an output that cannot be expressed by the voltage output range $V_o = 0 - 1$ V. These clearly need to be disregarded.

Relevant time interval. As highlighted in Section 3, the relevant section of the experiment has to be identified after the sample is inserted in the well. Any earlier signal is then disregarded as irrelevant. Similarly, the underlying biological properties allow identifying the end of the experiment after 450 samples. In fact, it is possible to infer from the standard curve of Covid-19 and Cancer Biomarkers 1, 2, and 3 LAMP amplification that any positive sample would lead to detectable DNA amplification within the first 22 min [7]. Despite the uneven sampling across experiments, selecting 450 samples allows us to always capture the relevant interval whilst avoiding fitting irrelevant data after the 22 min timestamp.

Trapped charge compensation. Trapped charge compensation consists in subtracting the initial offset from the signal to ensure that all experiments start at the same voltage. However, this approach does not account for the effect of trapped charge on sensor drift, which has not been fully characterised and therefore is not addressed in this work. In reality, offset subtraction does not completely mitigate the effect of trapped charge, which has an influence on the drift behaviour and pH sensitivity of the ISFET, but any additional compensation would be non-trivial because of the lack of complete mathematical models that account for these dependencies.

Spectrogram generation. To explore the frequency content of a discrete ISFET output signal $v(n)$ of length N, the average behaviour of the entire signal could be explored by taking the DFT over the entire period N. However, a more appropriate tool for the non-stationary ISFET output is to look for local information by dividing the signal into small overlapping windows, where the signal can be assumed to be stationary. The time–frequency spectrogram is then found with the STFT as

$$STFT\{v(n)\} = V(m, \omega)$$
$$= \sum_{n=0}^{N-1} v(n + mH)w(n)\,exp\left(-2\pi ik\frac{n}{N}\right) \quad (3)$$

where N is the window size, H is the shift between successive windows, $m \in [0 : M], k \in [0 : K]$ and $w(n)$ is a Hanning window (4) with window size $N = 22$ and window shift $H = 11$.

$$w(n) = 0.5\left(1 - \cos\left(2\pi\frac{n}{N-1}\right)\right) \quad (4)$$

The spectrogram is individually obtained for every pixel, so that each 450-sample signal is transformed in a 12×39 spectrogram array $V(m, \omega)$, with frequency estimates $\in [0 : f_s/2]$ and resolution of f_s/N, where f_s is the sampling frequency.

The spectrogram represents the ISFET readout in a redundant manner, such that each time sample is used twice to obtain a frequency estimate. This in turn allows training a neural network that can learn to classify these features. The underlying idea is that the frequency of drift at the time of amplification is lower than the sigmoid indicating amplification, thus aiding in classification.

High-frequency noise compensation. The chemical reaction associated with DNA hybridisation is a slow process, hence the expected ISFET signal is equivalently expected not to show abrupt changes. As such, the high frequencies are not expected to carry any significant information for classification. Then, the 10×39 section of the spectrum corresponding to the lowest frequencies is used as input to the CNN to avoid fitting data that is known to be irrelevant for classification.
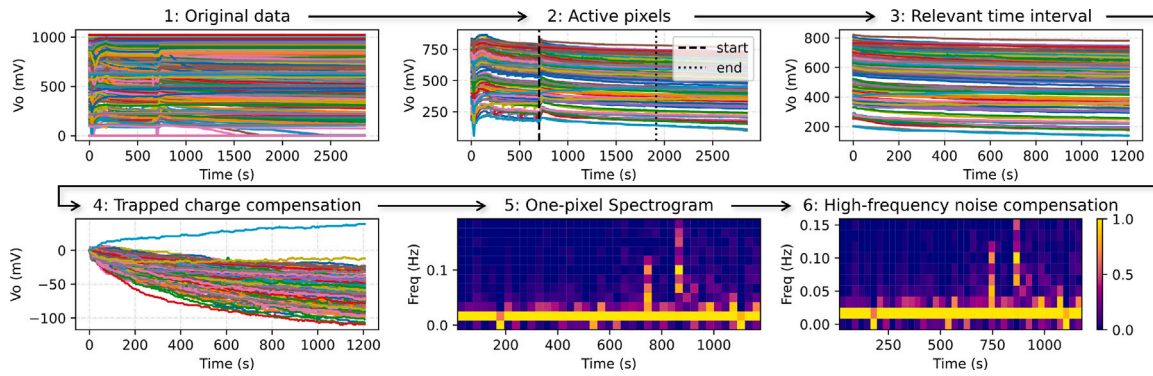
Fig. 3. Schematic of the steps involved in the data pre-processing. Note that the spectrograms are generated for each pixel individually although the output is only illustrated here for a single sensor.
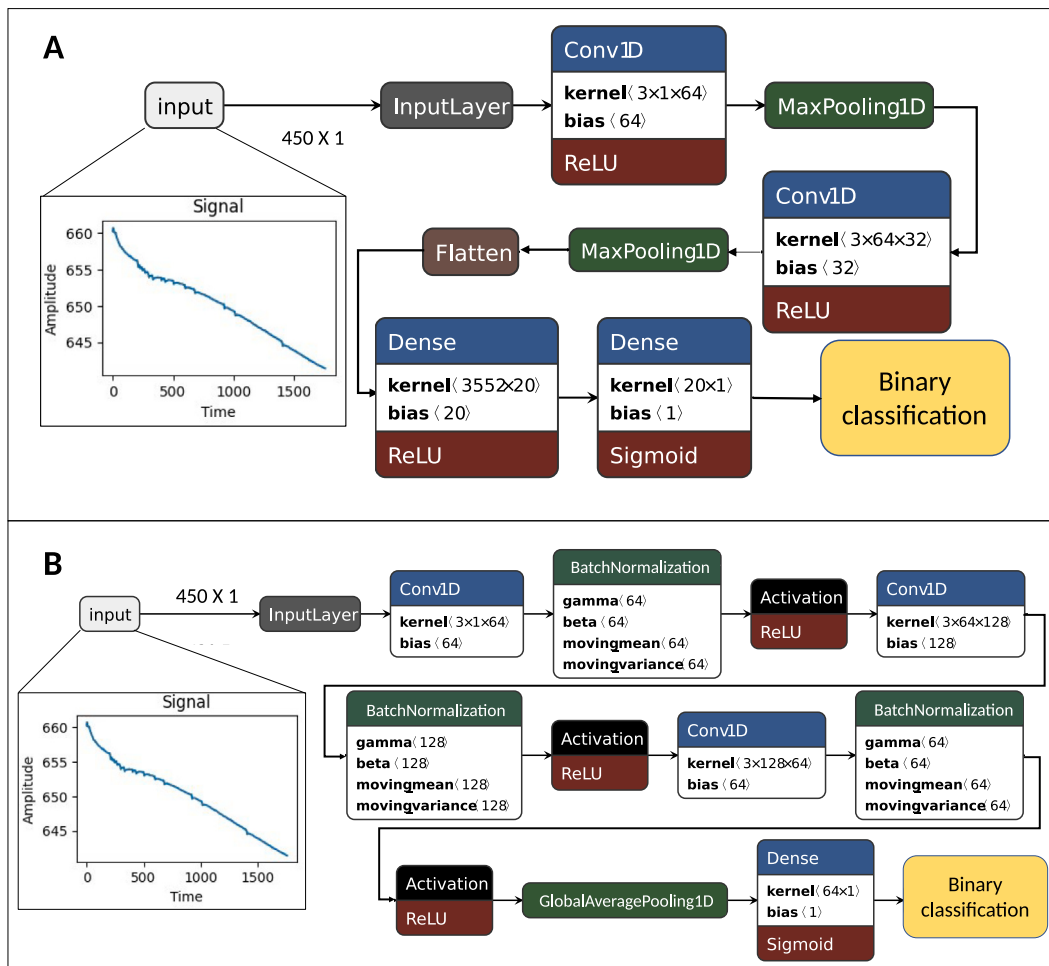

Fig. 4. Time domain neural networks: (a) DCNN, (b) FCN.

4.2. Implemented networks and results

The aim to implement a neural network for binary classification involving the use of both time domain and spectrogram-based approaches to train neural networks. All networks are trained and compared by employing the same seed value, learning rate, epoch, batch size and *adam* as the optimiser for a fair comparison. For the chosen spectrogram-based network, the hyperparameter optimisation is reported in Appendix C. The first approaches explored to achieve the binary classification of the ISFET data are based on processing the time domain data directly. There are sophisticated methods proposed for

time-series-based classification in literature such as FCN [53], Inception Time [52], and RESNET [53] that are proven to perform better than image classification using CNN based models. This was the motivation behind the implementation of the following time-based networks on the time series data from ISFETs:

- *Feed forward neural network*: a basic network with 64 nodes in the input layer followed by 32 and 16 nodes in the hidden layer. They all use Rectified Linear Unit (*ReLU*) as the activation function. These are then flattened and the output *Sigmoid* layer makes the classification.
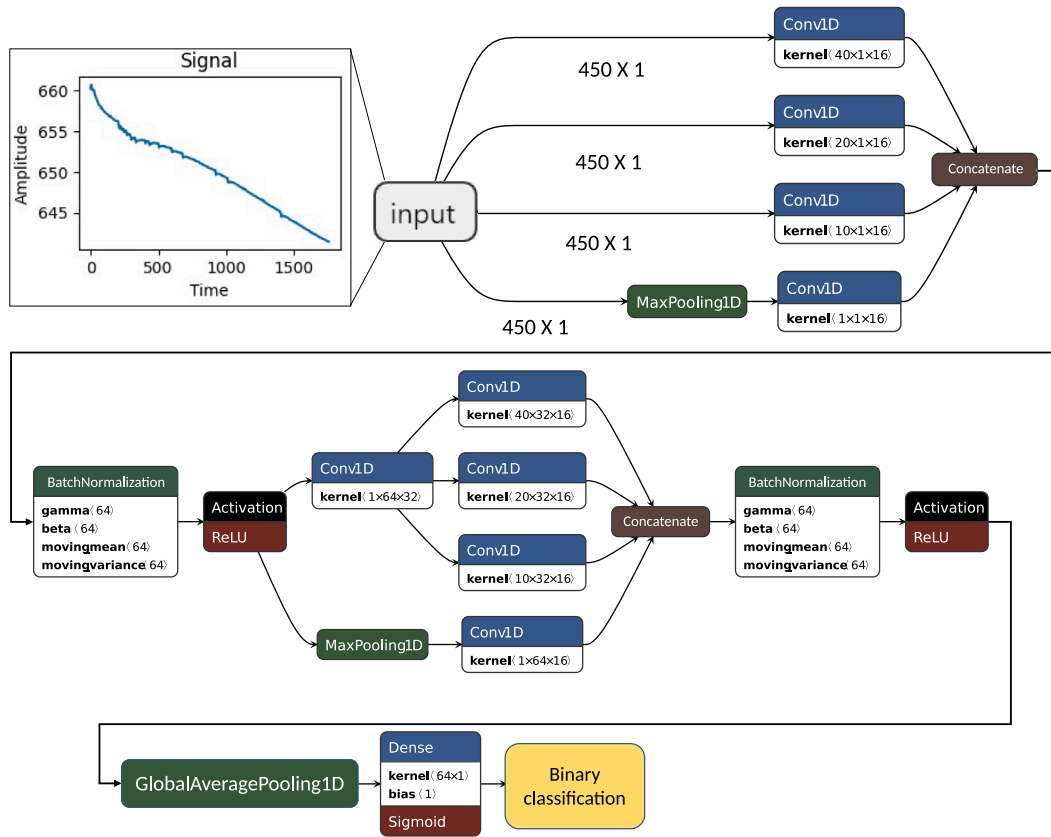
Fig. 5. InceptionTime: This implementation uses multiple filters to automatically extract relevant features on time series data [52].

- *1D-DCNN*: a convolutional network was implemented as shown in Fig. 4a, showing some improvement in performance with the addition of convolutional filters. From Table 2, it can be inferred that 1D-DCNN has the highest accuracy for all the time domain approaches but the lowest recall, suggesting a very high probability of false negatives.
- *FCN*: this was implemented as shown in Fig. 4b From Table 2, it can be inferred that not only does FCN take a longer training runtime but also has a very low recall making it unsuitable for medical applications like the 1D-DCNN.
- *InceptionTime*: this was implemented as shown in Fig. 5. From Table 2, it can be inferred that InceptionTime has the highest accuracy of all the time domain approaches. However, the extremely high training runtime and a large number of trainable parameters make it unsuitable for edge implementations.
- *RESNET*: this was implemented as shown in Fig. 6. From Table 2, it can be inferred that while the RESNET has a more balanced performance across all the metrics, it still needs a longer training runtime and manages to reach an accuracy of only 66.89
- *Autoencoder*: this was implemented as from Fig. 7 and trained with *Mean Absolute Error* [54] loss function, given in Eq. (5), where y_i is the predicted value and $\hat{y}_i$ is the true value and N is the number of samples.

$$MAE = \frac{1}{N} \sum_{j=1}^{N} |y_i - \hat{y}_i| \qquad (5)$$

This started showing some improvements in the recall of the network but leading to low accuracy and precision, essentially showing a bias towards the classification of samples as positive.

This is due to the network being trained only with the positively labelled data and then applying the principle of anomaly detection by using a threshold to treat the negative samples as anomalies. The encoder input, decoder output and reconstruction error for eight pixels in the test set is shown in Fig. 8.

The overall unsatisfactory classification performance of the time-based networks reported in Table 2 resulted in the decision to feature engineer the sensor data into spectrograms, thus allowing the exploration of image-based approaches for training neural networks. A 2D-CNN was designed for classification as shown in Fig. 9. This consists of two convolutional layers with *ReLU* activation function, with each convolutional layer being followed by a 2×2 *Max-Pooling* function. The layer is then flattened and passed through a dense layer with *ReLU* activation function. The output layer is based on a *Sigmoid* activation function for the ultimate classification of a healthy or infected patient. Because the aim is to solve a binary classification problem, the loss function is *Binary Cross-entropy* [55], reported in Eq. (6), with mean-based threshold classification in the neural networks.

$$L_{BCE}(y, \hat{y}) = -(y log(\hat{y}) + (1 - y) log(1 - \hat{y})) \qquad (6)$$

The hyperparameter optimisation performed on the network led to the decision to train with a Batch Size of 16, 40 Epochs, an initial learning rate of 0.001 with a step decay of 75% every 10 epochs.

As the model is implemented to achieve point-of-care diagnostic on a microcontroller, the neural network implementation is aimed at compact structures that are compatible with the application, and that can additionally be trained quickly to be deployed in emergencies. Furthermore, quantisation-aware training was performed with
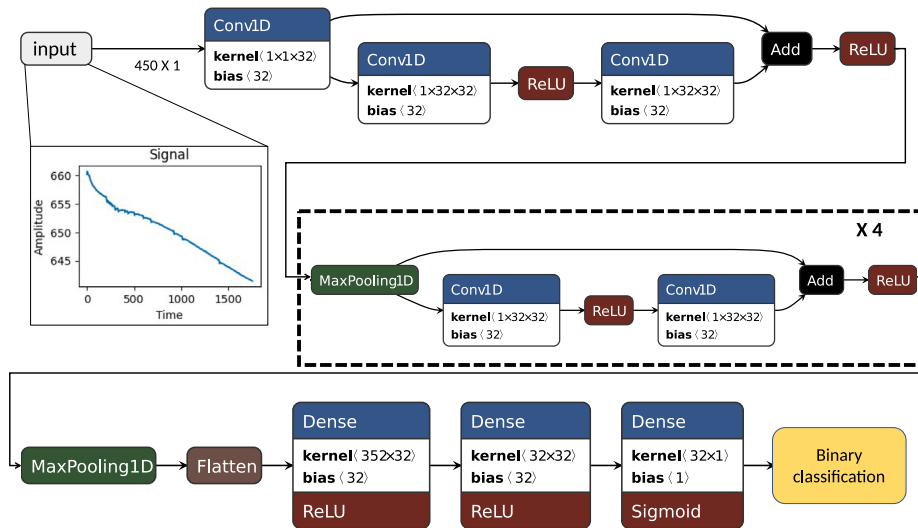
Fig. 6. RESNET: These use skip connections that can help train deeper neural networks by tackling the vanishing gradient.
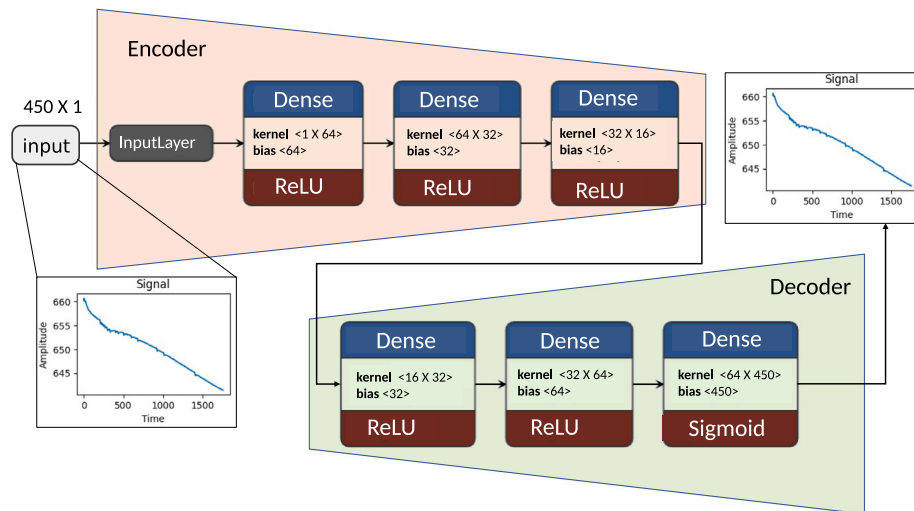

Fig. 7. Autoencoder: The approach is based on training a network to reconstruct just the infected samples and using reconstruction error as a metric for classification.
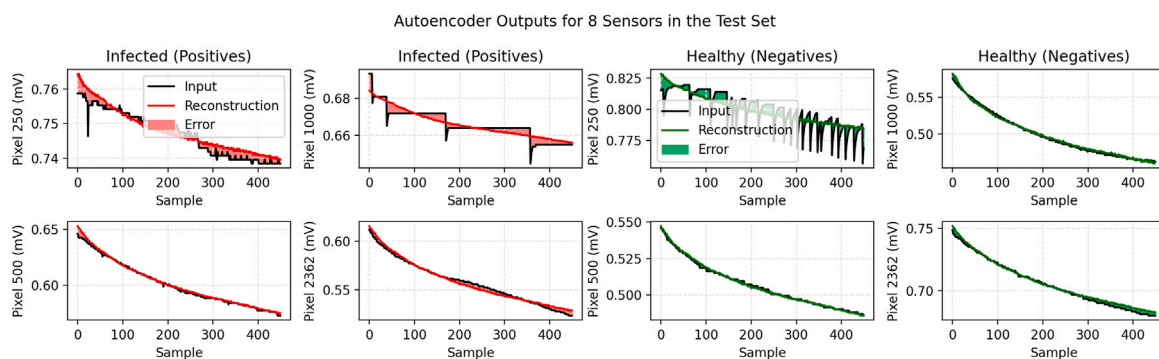

Fig. 8. Autoencoder Processing. Input, Reconstruction and Error for 8 time-domain test pixel signals using autoencoders.
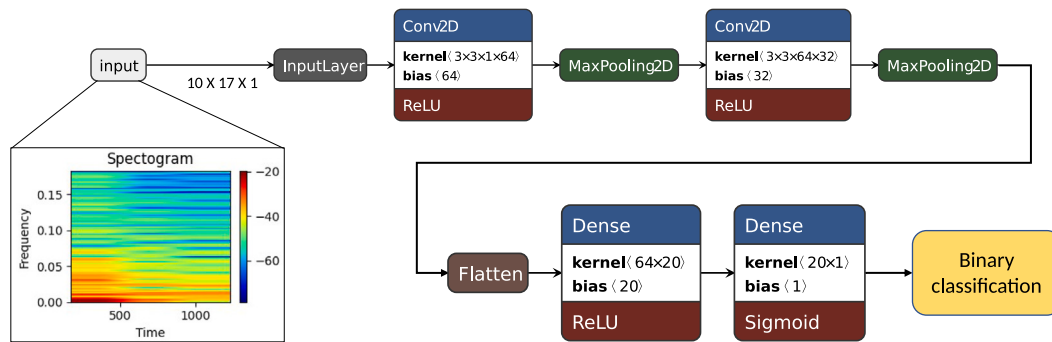
Fig. 9. DCNN using spectrograms. This consists of two convolutional layers with *ReLU* activation function, with each convolutional layer being followed by a 2×2 *Max-Pooling* function. The layer is then flattened and passed through a dense layer with *ReLU* activation function. The output layer is based on a *Sigmoid* activation function for the ultimate classification of a healthy or infected patient.
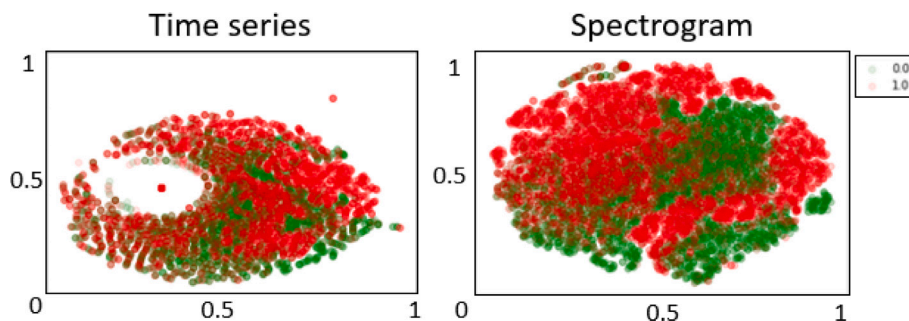

Fig. 10. T-SNE plot. T-SNE plot of the features from the penultimate layer in DCNNs for time series and spectrogram.

Table 2
Evaluation metrics values for comparison of the proposed methods. I-B Precision stands for inter-biomarker testing precision on experiment data.

Metrics	Time-based						Spectrogram-based 2D-DCNN				
	ANN	1D-DCNN	Autoencoder	FCN	Resnet	InceptionTime	Fold1	Fold2	Fold3	Fold4	Fold5
Precision (%)	58.15	74.87	57.91	59.11	66.89	69.77	85.04	85.76	85.01	84.54	85.45
Accuracy (%)	57.58	60.33	55.8	57.24	69.67	69.05	84.21	84.78	84.84	83.79	84.84
Recall (%)	54.05	31.1	89.16	46.93	77.87	67.23	83.21	83.67	84.43	82.59	83.85
F1 score (%)	56.03	43.95	70.21	52.32	71.96	68.48	84.11	84.70	84.72	83.55	84.64
Accuracy (Dataset-1) (%)	–	–	–	–	–	–	71.98	72.16	73.26	72.66	73.08
Precision (Dataset-1)	–	–	–	–	–	–	5/5	5/5	5/5	5/5	5/5
I-B Precision	–	–	–	–	–	–	4/4	4/4	4/4	4/4	4/4
Training runtime(hrs)	0.22	0.38	0.12	2.4	0.5	3.95	0.26	0.27	0.25	0.26	0.2
Trainable Parameters	9937	77513	34578	50177	23009	40369	24265	24265	24265	24265	24265

the spectrogram-based DCNN model. This in turn allows conversion of the model to *TFLite* and *TFLite-Micro* for compatibility with the microcontroller.

Training approach. Two conflicting arguments were considered in the training set preparation. On one hand, as seen in Section 2.1, the single-pixel signal is highly non-ideal: that is why experiments are often treated as the average of the active pixels, essentially using spatial averaging to compensate for the noise of the isolated signal. This is in contrast with the requirement for large amounts of data to successfully train a network. The proposed method thus uses all pixel data as independent samples to train the model. The underlying assumption is that all sensors in contact with a positive sample will eventually detect the reaction at some point. Thus, we present a data augmentation approach for LoC devices to help train with fewer experiments. The limitation in the number of experiments available (83) lead to the decision to consider the single-pixel data in the model instead of the experiment average, leading to 70,208 sets used for training and 17552 for testing. In fact, each experiment is comprised of thousands of pixels that provide a reading independently; the pixel signals from the same experiment do present some level of correlation, thus introducing some redundancy in the dataset but were sufficient for training. In

this method, the dataset is kept balanced in the training by selecting an equal number of positive and negative active pixels. All ISFET pixels are independent and only share resources for communication protocols. The dataset is thus trained and validated using a 5-fold cross-validation strategy by treating each pixel as an independent training data point with its own spectrogram. This is followed by testing on Dataset-1 which comprises 5 new patient samples which have 3 mRNA-based samples and 2 negatives. To further test the generalisability, Interbiomarker precision is further defined for testing of 4 additional Cancer samples which is the cancer marker 3 and is DNA based sample.

Testing approach. 5-Fold Cross-Validation is used for testing, resulting in an 80/20 split for train/test sets with 21105 pixels in turn used to evaluate the performance of the model. Performance metrics are introduced in Appendix B of the supplementary material and values are reported for each fold in Table 2. Bar error plots with tolerance intervals with respect to the random seed are also provided in Appendix C. Appendix C presents the hyperparameter optimisation with hyperparameter tuning done through the scikit grid search. The performance metrics are reported across multiple hyperparameter values and the order of optimisation is also defined. Fig. 10 shows the T-SNE plot of the features from the penultimate layer in the DCNN applied
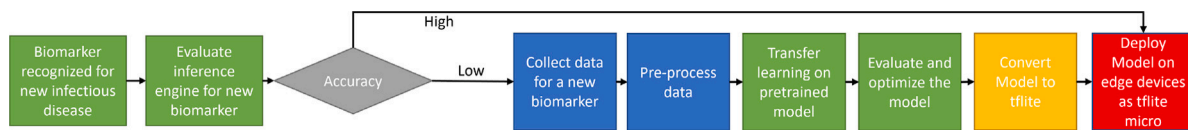
Fig. 11. Proposed edge implementation framework for new infectious diseases. Note that here edge devices will be microcontrollers like the Arduino Nano 33 BLE Sense.

to the time series (left) and on the spectrogram (right). This succinctly depicts that the features extracted from the time–frequency data allow for better distinguishability of the negative and positive experiments, respectively labelled 0 and 1.

Although the spectrogram-based DCNN has proven to be a powerful approach for single-pixel classification, in the context of LoC diagnostics the interest is towards the classification of data from entire experiments, that is, identifying samples as positive/negative using data from the array. Table 1 shows that the SARS-COV-2, Cancer Biomarker 1, Cancer Biomarker 2 and Negatives are used for both training and testing. On the other hand, 4 samples of Cancer Biomarker 3 are just used for testing. This gives rise to the Inter-Biomarker Precision metric in Table 2. Here, the model is evaluated on the basis of single pixel classification and then a majority voting approach is implemented to determine whether the entire experiment is classified as positive or negative. This was tested across the 5 folds for positive samples of Biomarker 3. The percentage of positive active pixels in the positive cancer Biomarker 3 experiments was 59.01% on average, with a standard deviation of 5.57%, with the majority voting always in favour of a positive outcome. This shows that the algorithm is robust to allow diagnostics of data not directly used in the training, thus potentially resulting in a universal diagnostics platform where only the assay preparation needs to be updated.

5. Edge implementation

When treating medical data, such as patient samples in the proposed application, patient privacy and security is essential to avoid malicious intent. Therefore, the use of edge devices is critical to support diagnostic testing while respecting patient confidentiality. The quantisation-aware training of the proposed spectrogram-based DCNN model showed limited drop in performance, achieving an accuracy of 84.57% in the *TFLite* implementation, thus showing good metrics for deployment of devices on the edge. The converted *TFLite-Micro* model of the Spectrogram based DCNN is 30 kB, making it suitable for deployment on most microcontrollers. The microcontroller of choice is the Nano 33 BLE Sense which is compatible with *TFLite-Micro* and can have a standby power consumption of just 0.9 nA. It is expected that the edge implementation of the neural network-based inference engine will have a negligible contribution compared to the heater required for the LAMP reaction that is present on the current system and takes more than 80% of the power consumption. This will further be tested in future work with a full system implementation. Fig. 11 illustrates a framework for the use of the proposed platform in a generic diagnostics context by leveraging the power of *TinyML* [56].

6. Conclusion

This paper presents a novel spectrogram-based DCNN approach that leverages the redundant feature space representation achieved by spectrograms for the classification of ISFET array data of nucleic acid amplification on LoC platforms. The proposed method provides a significant improvement from 57.58%, 60.33%, 55.8%, 57.24%, 69.67%, and 69.05% for the time-domain-based ANN, 1D-DCNN, autoencoder, FCN, RESNET, and InceptionTime approaches respectively, to 84.84% for the spectrogram based 2D-DCNN approach. Our proposed method is the preferred approach because it improves accuracy, reduces bandwidth requirement and allows faster time to market as a more generalised approach without the need for complex feature engineering.

Thus allowing for more efficient diagnostics. Furthermore, medical data from ISFET sensors is very different from the EMG-based and ECG-based time-domain data that has been used to prove the accuracy of time-domain-based networks in the past. This is due to the fact that the presence of an infected sample is plagued by drift, offset and trapped charge in the CMOS-based ISFET sensors, thus making it difficult to highlight a constant feature in the time series data. In addition, we have proposed a preprocessing approach to augment the dataset to allow researchers to train with limited datasets of electrochemical readouts from nucleic acid amplification for infectious disease diagnosis. We also present an edge implementation framework that can help provide tests with high sensitivity and specificity to patients while addressing the need for data security. The accurate classification of Cancer Biomarker 3, which is excluded from the training set, proves that the model can be generalised for future diagnostic applications. This work sets a benchmark for future AI approaches to point-of-care diagnostic devices using potentiometric sensors and presents spectrograms as a benchmark transform for implementing neural networks in the field of diagnostics for infectious diseases and cancers.

7. Limitations and future work

- Although some methods are used for active pixel identification, the performance of individual sensors has not been characterised in the overall positive and negative experiments. This means a greater amount of time spent on the annotation of individual sensor data could result in higher accuracy of the models.
- The pre-processing of data needs to be converted from Python to C++ for integration with microcontrollers. We intend to work on this and show real-time results in the future.
- Treating the sensors as pixels forming electrochemical image frames can help us apply image processing techniques without much pre-processing. This would require a large amount of frame-wise annotation and will be pursued in the future to compare performance.
- The ISFET output signal is low frequency and has information in terms of time, frequency and amplitude. This makes a time–frequency and amplitude transform like the spectrogram, a perfect approach for ISFET based diagnostics, allowing the authors to present it as a benchmark. We expect to carry out future work to identify the pros and cons of alternative transforms. Among the alternative transforms, Wigner Ville [57,58] and scalograms [59] have shown promise in the past in biomedical applications. However, Wigner Ville is known to introduce cross terms between several frequency components. As is visible in the signal from ISFETs, there are a large amount of frequency components residing in the low-frequency region, making Wigner Ville unsuitable for this application. Scalograms on the other hand are computationally expensive and are not suitable for implementation on microcontrollers. Currently, the constant window constant length in STFT for spectrograms appears as a more suitable candidate.
- The models used in this paper have been chosen by keeping in mind that the size of the model should be compact enough for deployment on a Nano BLE 33. This means an effort will be made in the future to explore more sophisticated neural networks such as recurrent neural networks (RNNs) and transformer-based models.

Declaration of competing interest

None Declared.

Acknowledgements

Appendix A. Supplementary data

The Supplementary section contains the data collection and ethics approval as part of Appendix A. Appendix B defines the performance metrics used for comparison of the neural networks. Appendix C presents the hyperparameter optimisation.

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compbiomed.2023.107027.

References

[1] B. Merrick, M. Noronha, R. Batra, S. Douthwaite, G. Nebbia, L. Snell, S. Pickering, R. Galao, J. Whitfield, A. Jahangeer, R. Gunawardena, T. Godfrey, R. Laifa, K. Webber, P. Cliff, E. Cunningham, S. Neil, H. Gettings, J. Edgeworth, H. Harrison, Real-world deployment of lateral flow SARS-CoV-2 antigen detection in the emergency department to provide rapid, accurate and safe diagnosis of COVID-19, Infect. Prev. Pract. 3 (4) (2021) 100186.

[2] A. Priye, S.W. Bird, Y.K. Light, C.S. Ball, O.A. Negrete, R.J. Meagher, A smartphone-based diagnostic platform for rapid detection of zika, chikungunya, and dengue viruses, Sci. Rep. 1 (12) (2017).

[3] P. Bergveld, Development, operation, and application of the ion-sensitive field-effect transistor as a tool for electrophysiology, IEEE Trans. Biomed. Eng. BME-19 (1972).

[4] N. Moser, L. Keeble, J. Rodriguez-Manzano, P. Georgiou, ISFET arrays for lab-on-chip technology: A review, in: 2019 26th IEEE International Conference on Electronics, Circuits and Systems, ICECS, 2019, pp. 57–60, http://dx.doi.org/10.1109/ICECS46596.2019.8965034.

[5] J. Bausells, J. Carrabina, A. Errachid, A. Merlos, Ion-sensitive field-effect transistors fabricated in a commercial CMOS technology, Sensors Actuators B 57 (1) (1999) 56–62.

[6] J.M. Rothberg, W. Hinz, T.M. Rearick, J. Schultz, W. Mileski, M. Davey, J.H. Leamon, K. Johnson, M.J. Milgrew, M. Edwards, J. Hoon, J.F. Simons, D. Marran, J.W. Myers, J.F. Davidson, A. Branting, J.R. Nobile, B.P. Puc, D. Light, T.A. Clark, M. Huber, J.T. Branciforte, I.B. Stoner, S.E. Cawley, M. Lyons, Y. Fu, N. Homer, M. Sedova, X. Miao, B. Reed, J. Sabina, E. Feierstein, M. Schorn, M. Alanjary, E. Dimalanta, D. Dressman, R. Kasinskas, T. Sokolsky, J.A. Fidanza, E. Namsaraev, K.J. McKernan, A. Williams, G.T. Roth, J. Bustillo, An integrated semiconductor device enabling non-optical genome sequencing, Nature 475 (2011) 348–352.

[7] J. Rodriguez-Manzano, K. Malpartida-Cardenas, N. Moser, I. Pennisi, M. Cavuto, L. Miglietta, A. Moniri, R. Penn, G. Satta, P. Randell, F. Davies, F. Bolt, W. Barclay, A. Holmes, P. Georgiou, Handheld Point-of-Care system for rapid detection of SARS-CoV-2 extracted RNA in under 20 min, ACS Cent. Sci. 7 (2021) 307–317, http://dx.doi.org/10.1021/acscentsci.0c01288.

[8] J. Broomfield, M. Kalofonou, T. Pataillot-Meakin, S.M. Powell, R.C. Fernandes, N. Moser, C.L. Bevan, P. Georgiou, Detection of YAP1 and AR-V7 mRNA for prostate cancer prognosis using an ISFET lab-on-chip platform, ACS Sensors (2022).

[9] G. Alexandrou, N. Moser, K.T. Mantikas, J. Rodriguez Manzano, S. Ali, R.C. Coombes, J. Shaw, P. Georgiou, C. Toumazou, M. Kalofonou, Detection of multiple breast cancer ESR1 mutations on an ISFET based lab-on-chip platform, IEEE Trans. Biomed. Circuits Syst. 15 (3) (2021) 380–389.

[10] S. Kurbanoglu, B. Dogan-Topal, E.P. Rodriguez, B. Bozal-Palabiyik, S.A. Ozkan, B. Uslu, Advances in electrochemical DNA biosensors and their interaction mechanism with pharmaceuticals, J. Electroanal. Soc. 775 (2016) 8–26.

[11] N. Moser, L.S. Yu, J. Rodriguez Manzano, K. Malpartida-Cardenas, A. Au, P. Arkell, C. Cicatiello, A. Moniri, L. Miglietta, W.-H. Wang, S.F. Wang, A. Holmes, Y.-H. Chen, P. Georgiou, Quantitative detection of dengue serotypes using a smartphone-connected handheld lab-on-chip platform, Front. Bioeng. Biotechnol. 10 (2022).

[12] L.S. Yu, J. Rodriguez-Manzano, N. Moser, A. Moniri, K. Malpartida-Cardenas, N. Miscourides, T. Sewell, T. Kochina, A. Brackin, J. Rhodes, A.H. Holmes, M.C. Fisher, P. Georgiou, Rapid detection of azole-resistant *Aspergillus fumigatus* in clinical and environmental isolates by use of a lab-on-a-chip diagnostic system, J. Clin. Microbiol. 58 (11) (2020) e00843–20, URL https://journals.asm.org/doi/abs/10.1128/JCM.00843-20.

[13] G. Wang, J.C. Ye, B. De Man, Deep learning for tomographic image reconstruction, Nat. Mach. Intell. 2 (12) (2020) 737–748.

[14] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, Nat. Med. 25 (1) (2019) 24–29.

[15] J. van der Laak, G. Litjens, F. Ciompi, Deep learning in histopathology: the path to the clinic, Nat. Med. 27 (5) (2021) 775–784.

[16] S. Gupta, S. Ali, L. Goldsmith, B. Turney, J. Rittscher, MI-UNet: Improved segmentation in ureteroscopy, in: 2020 IEEE 17th International Symposium on Biomedical Imaging, ISBI, 2020, pp. 212–216, http://dx.doi.org/10.1109/ISBI45749.2020.9098608.

[17] E. Messner, M. Fediuk, P. Swatek, S. Scheidl, F.-M. Smolle-Jüttner, H. Olschewski, F. Pernkopf, Multi-channel lung sound classification with convolutional recurrent neural networks, Comput. Biol. Med. 122 (2020) 103831.

[18] A. Verma, S.B. Amin, M. Naeem, M. Saha, Detecting COVID-19 from chest computed tomography scans using AI-driven android application, Comput. Biol. Med. 143 (2022) 105298.

[19] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., ImageNet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.

[21] G. Chartrand, P.M. Cheng, E. Vorontsov, M. Drozdzal, S. Turcotte, C.J. Pal, S. Kadoury, A. Tang, Deep learning: A primer for radiologists, RadioGraphics 37 (7) (2017) 2113–2131, http://dx.doi.org/10.1148/rg.2017170077, PMID: 29131760.

[22] J. De Fauw, J.R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, D. Visentin, et al., Clinically applicable deep learning for diagnosis and referral in retinal disease, Nat. Med. 24 (9) (2018) 1342–1350.

[23] W.G. van Panhuis, P. Paul, C. Emerson, J. Grefenstette, R. Wilder, A.J. Herbst, D. Heymann, D.S. Burke, A systematic review of barriers to data sharing in public health, BMC Public Health 14 (1) (2014).

[24] G. Rodrigues, C.M. Barreira, M. Bouslama, D.C. Haussen, A. Al-Bayati, L. Pisani, B. Liberato, N. Bhatt, M.R. Frankel, R.G. Nogueira, Automated large artery occlusion detection in stroke: A single-center validation study of an artificial intelligence algorithm, Cerebrovasc. Dis. 51 (2) (2021) 259–264.

[25] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.

[26] H. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A.B.H. Hassen, L. Thomas, A. Enk, et al., Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists, Ann. Oncol. 29 (8) (2018) 1836–1842.

[27] D.S.W. Ting, L.R. Pasquale, L. Peng, J.P. Campbell, A.Y. Lee, R. Raman, G.S.W. Tan, L. Schmetterer, P.A. Keane, T.Y. Wong, Artificial intelligence and deep learning in ophthalmology, Br. J. Ophthalmol. 103 (2) (2019) 167–175.

[28] A. Echle, N.T. Rindtorff, T.J. Brinker, T. Luedde, A.T. Pearson, J.N. Kather, Deep learning in cancer pathology: a new generation of clinical biomarkers, Br. J. Cancer 124 (4) (2020) 686–696.

[29] N. Moser, T.S. Lande, C. Toumazou, P. Georgiou, ISFETs in CMOS and emergent trends in instrumentation: A review, IEEE Sens. J. 16 (2016) 6496–6514.

[30] K. Malpartida-Cardenas, N. Miscourides, J. Rodriguez-Manzano, L.S. Yu, N. Moser, J. Baum, P. Georgiou, Quantitative and rapid *Plasmodium falciparum* malaria diagnosis and artemisinin-resistance detection using a CMOS lab-on-chip platform, Biosens. Bioelectron. 145 (2019).

[31] P. Georgiou, C. Toumazou, ISFET characteristics in CMOS and their application to weak inversion operation, Sensors Actuators B 143 (2009) 211–217.

[32] S. Jamasb, S. Collins, R.L. Smith, A physical model for drift in pH ISFETs, Sensors Actuators B 49 (1998) 146–155.

[33] M. Sohbati, C. Toumazou, Dimension and shape effects on the ISFET performance, IEEE Sens. J. 15 (2015) 1670–1679.

[34] S. Sinha, R. Bhardwaj, N. Sahu, H. Ahuja, R. Sharma, R. Mukhiya, Temperature and temporal drift compensation for Al2O3-gate ISFET-based pH sensor using machine learning techniques, Microelectron. J. 97 (2020) 104710.

[35] N. Sahu, R. Bhardwaj, H. Shah, R. Mukhiya, R. Sharma, S. Sinha, Towards development of an ISFET-based smart pH sensor: Enabling machine learning for drift compensation in IoT applications, IEEE Sens. J. 21 (17) (2021) 19013–19024.

[36] W.E. Hsu, Y.H. Chang, C.T. Lin, A machine-learning assisted sensor for chemo-physical dual sensing based on ion-sensitive field-effect transistor architecture, IEEE Sens. J. 19 (21) (2019) 9983–9990.

[37] P. Tripathi, N. Moser, P. Georgiou, Multiple ion-channel ISFET neuron for lab-on-chip applications, in: 2021 IEEE Biomedical Circuits and Systems Conference, BioCAS, 2021, pp. 01–05, http://dx.doi.org/10.1109/BioCAS49922.2021.9645040.

[38] P. Tripathi, N. Moser, P. Georgiou, A neuron-based ISFET array architecture with spatial sensor compensation, in: 2019 IEEE International Symposium on Circuits and Systems, ISCAS, 2019, pp. 1–5, http://dx.doi.org/10.1109/ISCAS.2019.8702711.

[39] L. Wyse, Audio spectrogram representations for processing with convolutional neural networks, in: Proceedings of the IEEE International Conference on Deep Learning and Music, 2017, pp. 37–41.

[40] M. Mazumder, C. Banbury, J. Meyer, P. Warden, V.J. Reddi, Few-shot keyword spotting in any language, in: Proc. Interspeech 2021, 2021, pp. 4214–4218, http://dx.doi.org/10.21437/Interspeech.2021-1966.

[41] L. Nanni, A. Rigo, A. Lumini, S. Brahnam, Spectrogram classification using dissimilarity space, Appl. Sci. (Switzerland) 10 (2020).

[42] Q. Zhou, J. Shan, W. Ding, C. Wang, S. Yuan, F. Sun, H. Li, B. Fang, Cough recognition based on mel-spectrogram and convolutional neural network, Front. Robot. AI 8 (2021).

[43] J. Huang, B. Chen, B. Yao, W. He, ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network, IEEE Access 7 (2019) 92871–92880.

[44] G. Ruffini, D. Ibañez, M. Castellano, L. Dubreuil-Vall, A. Soria-Frisch, R. Postuma, J.F. Gagnon, J. Montplaisir, Deep learning with EEG spectrograms in rapid eye movement behavior disorder, Front. Neurol. 10 (2019).

[45] M. Baygin, S. Dogan, T. Tuncer, P. Datta Barua, O. Faust, N. Arunkumar, E.W. Abdulhay, E. Emma Palmer, U. Rajendra Acharya, Automated ASD detection using hybrid deep lightweight features extracted from EEG signals, Comput. Biol. Med. 134 (2021) 104548.

[46] M.R. Sawaya, R. Prasad, S.H. Wilson, J. Kraut, H. Pelletier, Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: Evidence for an induced fit mechanism, Biochemistry 36 (37) (1997) 11205–11215.

[47] H. Deng, Z. Gao, Bioanalytical applications of isothermal nucleic acid amplification techniques, Anal. Chim. Acta 853 (1) (2015) 30–45.

[48] T. Notomi, H. Okayama, H. Masubuchi, T. Yonekawa, K. Watanabe, N. Amino, T. Hase, Loop-mediated isothermal amplification of dna, Nucleic Acids Res. 28 (12) (2000).

[49] L.X. Wang, J.J. Fu, Y. Zhou, G. Chen, C. Fang, Z.S. Lu, L. Yu, On-chip RT-LAMP and colorimetric detection of the prostate cancer 3 biomarker with an integrated thermal and imaging box, Talanta 208 (2020).

[50] S.J. Ahn, Y.H. Baek, K.K.S. Lloren, W.-S. Choi, J.H. Jeong, K.J.C. Antigua, H.i. Kwon, S.J. Park, E.-H. Kim, Y.i. Kim, Y.J. Si, S.B. Hong, K.S. Shin, S. Chun, Y.K. Choi, M.-S. Song, Rapid and simple colorimetric detection of multiple influenza viruses infecting humans using a reverse transcriptional loop-mediated isothermal amplification (RT-LAMP) diagnostic platform, BMC Infect. Dis. 19 (1) (2019) 676.

[51] C. Toumazou, L.M. Shepherd, S.C. Reed, G.I. Chen, A. Patel, D.M. Garner, C.J.A. Wang, C.P. Ou, K. Amin-Desai, P. Athanasiou, H. Bai, I.M. Brizido, B. Caldwell, D. Coomber-Alford, P. Georgiou, K.S. Jordan, J.C. Joyce, M.L. Mura, D. Morley, S. Sathyavruthan, S. Temelso, R.E. Thomas, L. Zhang, Simultaneous DNA amplification and detection using a pH-sensing semiconductor system, Nature Methods 10 (2013) 641–646.

[52] H. Ismail Fawaz, B. Lucas, G. Forestier, C. Pelletier, D.F. Schmidt, J. Weber, G.I. Webb, L. Idoumghar, P.A. Muller, F. Petitjean, InceptionTime: Finding AlexNet for time series classification, Data Min. Knowl. Discov. 34 (6) (2020) 1936–1962.

[53] Z. Wang, W. Yan, T. Oates, Time series classification from scratch with deep neural networks: A strong baseline, in: 2017 International Joint Conference on Neural Networks, IJCNN, 2017, pp. 1578–1585.

[54] S. Dangoury, S. Abouzahir, A. Alali, M. Sadik, Impacts of losses functions on the quality of the ultrasound image by using machine learning algorithms, in: 2021 IEEE International Conference on Automatic Control Intelligent Systems, I2CACIS, 2021, pp. 380–385, http://dx.doi.org/10.1109/I2CACIS52118.2021.9495878.

[55] S. Jadon, A survey of loss functions for semantic segmentation, in: 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB, 2020, pp. 1–7, http://dx.doi.org/10.1109/CIBCB48159.2020.9277638.

[56] D. Situnayake, P. Warden, TinyML, O'Reilly Media, Inc., 2019.

[57] F. Akdeniz, İ. Kayikçioğlu, İ. Kaya, T. Kayikçioğlu, Using Wigner–Ville distribution in ECG arrhythmia detection for telemedicine applications, in: 2016 39th International Conference on Telecommunications and Signal Processing, TSP, 2016, pp. 409–412, http://dx.doi.org/10.1109/TSP.2016.7760908.

[58] I. Shafi, A. Aziz, S. Din, I. Ashraf, Reduced features set neural network approach based on high-resolution time-frequency images for cardiac abnormality detection, Comput. Biol. Med. 145 (2022) 105425.

[59] M. Loey, S. Mirjalili, COVID-19 cough sound symptoms classification from scalogram image representation using deep learning models, Comput. Biol. Med. 139 (2021) 105020.