



Fluid classification with integrated flow and pressure sensors using machine learning

D. Alveringh^{a,*}, D.V. Le^a, J. Groenesteijn^b, J. Schmitz^a, J.C. Lötters^{a,b}

^a University of Twente, Enschede, The Netherlands

^b Bronkhorst High-Tech BV, Ruurlo, The Netherlands

ARTICLE INFO

Keywords:

Microfluidics
Sensors
Microfabrication
Neural network
Chips
Lab-on-a-chip
Artificial intelligence
Deep learning
Machine learning
Electronics
IoT
Coriolis
Mass flow

ABSTRACT

This paper describes fluid classification methods using machine learning applied on a microfabricated Coriolis mass flow sensor with integrated pressure sensors. The latter are positioned upstream and downstream of the Coriolis mass flow sensor, which enables the measurement of the viscosity-dependent pressure drop. The Coriolis mass flow sensor itself is particularly sensitive to the mass flow and density of the fluid. Five different liquids (nitrogen, water, isopropanol, ethanol and acetone) are applied to the sensor system in different combinations of mass flow rate, pressure and temperature. For each combination, the raw signals from all sensors are amplified, demodulated, digitized, sampled and stored. Then BiLSTM and CNN neural networks were trained and tested by using train-test split validation and K-fold cross-validation. With both methods, the classification accuracy is determined using a different part of the dataset than for learning. For mass flow rates up to 5 g/h, pressures between 4 bar and 6 bar and temperatures between 288 K and 308 K. BiLSTM performs best with a cross-validated accuracy of 77% up to 100%, dependent on the inclusion of low-flow data.

1. Introduction

Liquids and gases play leading roles in organisms and machines and their interactions with their environments. Sensing fluid parameters, flow rates and pressures has therefore numerous applications in industry, healthcare and the environment. Miniaturization of fluid sensing principles used in these applications, by e.g., microtechnology, have shown significant advantages in mass fabrication and performance [1, 2].

Coriolis mass flow sensors form a category of these sensors that use a flow-dependent mechanism to detect the true mass flow rate. These sensors have successfully been miniaturized [3,4] and have been integrated with in-line pressure sensors using the same microtechnology [5]. This evolved into actual multi-parameter fluid measurement systems that are capable of measuring real-time mass flow, pressure, density and viscosity [6].

Sensing structures that are directly or indirectly sensitive to fluid parameters like viscosity and density can be used to recognize the fluid using a simple decision tree, e.g., for oil in water [7]. When situations become more complicated, e.g., substances with temperature-dependent viscosities and varying flows, more elaborate decision making processing is required. This can be developed automatically by a

computer system that learns from examples: machine learning [8]. The use of machine learning might even enable the detection of information from sensor combinations that was not detectable using conventional data processing [9,10].

The latter has potential, as sensors are generally not sensitive to just one physical quantity. For example, a simple steel ruler does not only measure distance, but also temperature as the steel expands for higher temperatures. Dependence on other physical quantities can be reduced by a smart design and sensor fusion [11]. Although designers try to optimize the linearity and sensitivity to one quantity of interest, it is mathematically more relevant that the output signals are distinguishably and repeatably sensitive to all quantities of interest. A specific example is presented in [12], in which the side-effects of a Coriolis mass flow sensor were used to measure the differential pressure. This enables the sensor to indirectly detect the kinematic viscosity and, thanks to the integrated density sensing within the same structure, dynamic viscosity.

Application-oriented machine learning has been used many times in combination with sensor data, examples being image recognition [13], speech recognition [14,15], farm optimization [16] and oil spill detection [17]. Research on machine learning is mostly focused on the software algorithms or on a novel application-oriented features using

* Corresponding author.

E-mail address: d.alveringh@utwente.nl (D. Alveringh).

off-the-shelf sensors. This trend is also applicable in fluidics, where e.g., the detection of gases using gas sensors has been improved in sensitivity and reduction of drift by machine learning [18]. More specific, machine learning techniques have also been applied to improve the performance of Coriolis mass flow sensors [19–21], e.g., by reducing the uncertainty for multi-phase flow measurements.

The use of machine learning to process the data from such experimental microfluidic devices is an emerging field, varying from biotechnological applications, e.g., cell classification [22] to device design optimization [23]. To our knowledge machine learning has never been used in combination with microfabricated Coriolis-based multi-parameter fluid measurement systems. Yet, it has potential to improve performance, understanding and functionality when directly applied to the raw sensor signals that contain as many as possible (including non-ideal) aspects of the sensing mechanisms. This paper describes the first step, in which the raw signals are used for fluid classification.

A feasibility study on a simulated dataset has been performed preliminary to this research [24]. In this paper, we present fluid classification using a microfabricated Coriolis mass flow sensor with integrated pressure sensors. First, the theory and design of the sensors and machine learning algorithms are described. Then, the fabrication, experimental setup and preprocessing are discussed. The classification performance is presented in the Results section. The paper concludes with a discussion, summary and outlook.

2. Theory and design

The sensor design is presented in [6] and is kept unchanged. It consists of a Coriolis mass flow sensor and two pressure sensors. The latter are positioned in the upstream and downstream channels of the Coriolis mass flow sensor as illustrated in Fig. 1a.

2.1. Mass flow and density sensing

The Coriolis mass flow sensor consists of a rectangular frame of silicon nitride channels with semi-circular cross section. It is mechanically fixed in the center at one of the sides of the rectangular frame. The channel can be actuated by Lorentz actuation, i.e., an alternating current through one of the wires on top of the channel in a magnetic field causes the channel to move in so-called twist mode. A mass flow through the vibrating channel causes a Coriolis force which forces the channel to vibrate in a second mode at the same frequency. This so-called swing mode is a measure for the mass flow rate through the channel. The actuation and detection modes are illustrated in Fig. 1b.

Two gold electrodes are positioned at each side of the twist mode axis as indicated in Fig. 1b. Without mass flow, the phase shift between the electrodes is 180° . The swing mode, due to a mass flow, causes the electrodes to move less than 180° shifted from each other. The phase shift $\Delta\phi$ can be derived from the ratio between the swing mode amplitude \hat{z}_{swing} and the twist mode amplitude \hat{z}_{twist} at the electrode's position x_e :

$$\Delta\phi = \arctan\left(\frac{\hat{z}_{\text{swing}}(x_e)}{\hat{z}_{\text{twist}}(x_e)}\right) \propto \arctan(\dot{m}) \approx \dot{m}, \quad (1)$$

which is approximately proportional to the mass flow \dot{m} for small phase shifts, and thus low mass flows.

The fluid inside the channel is a significant part of the total mass of the vibrating structure. The resonance frequency f_0 of the filled channel is therefore not only dependent on the mass of the channel m_c , but also on the density of the fluid ρ_f :

$$f_0 \propto \frac{1}{2\pi} \sqrt{\frac{k_{\text{eff}}}{m_c + V_f \rho_f}}, \quad (2)$$

with k_{eff} the effective spring constant of the channel for the twist mode and V_f the volume inside the vibrating channel. The density of the fluid can therefore be measured with the same structure and in parallel with the mass flow. Fig. 1e shows an illustration of the signals from the electrodes.

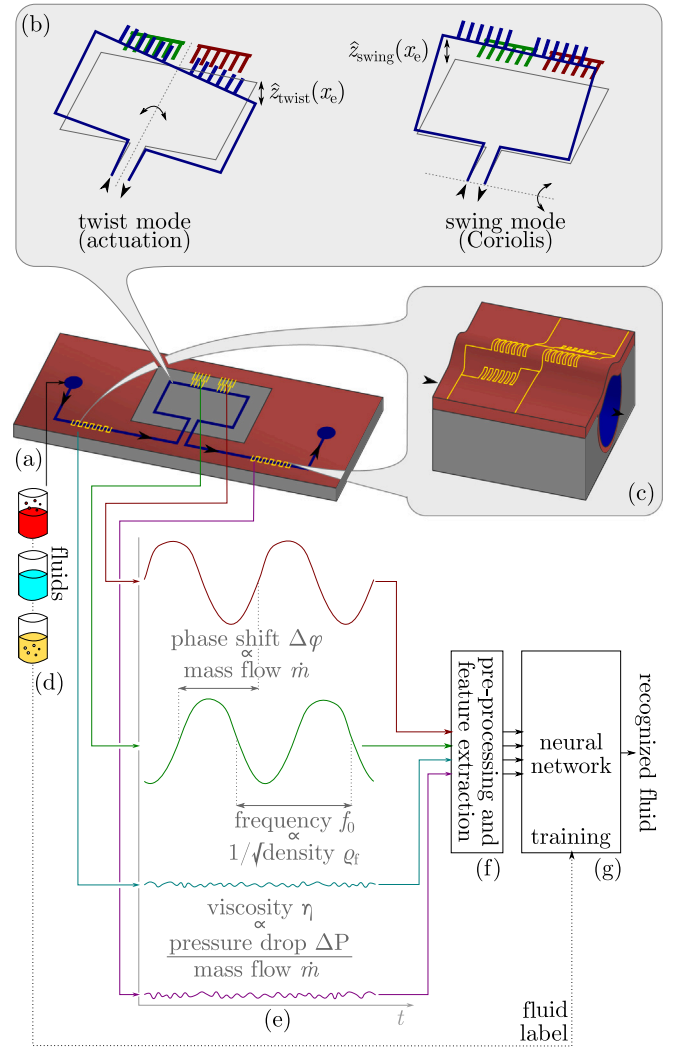


Fig. 1. Principle of fluid classification using an integrated mass flow sensor and pressure sensors, with (a) the chip, (b) the two modes of the Coriolis mass flow sensor, (c) the pressure sensor structure with deforming membrane, (d) different applied fluids, (e) the four output signals, (f) pre-processing and (g) the neural network that recognizes the fluid after training.

Source: Adapted from [5].

2.2. Pressure and viscosity sensing

The pressure sensors consist of the same type of semi-circular silicon nitride channels, but are fixed in the silicon bulk as illustrated in Fig. 1c. The flat silicon nitride ceiling of the channels forms a membrane that deforms due to a pressure inside the channel that is different than the pressure outside the channel. Meandering gold electrodes are positioned on top of the channel. Two of these electrodes elongate and two electrodes compress when the membrane deforms. These pressure-dependent resistors are configured in a Wheatstone bridge configuration. The output voltage V_{bridge} can be approximated by the following model:

$$V_{\text{bridge}} \propto PV_{\text{supply}}, \quad (3)$$

with P the gauge pressure and V_{supply} supply voltage of the Wheatstone bridge. Both pressure sensors combined can be used to measure the pressure drop

$$\Delta P = P_{\text{upstream}} - P_{\text{downstream}} \quad (4)$$

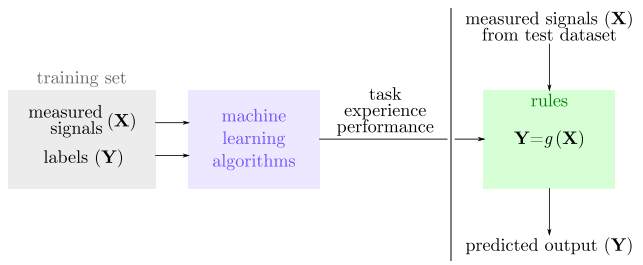


Fig. 2. Illustration of machine learning where a set of rules $g(\mathbf{X})$ will be found by algorithms to predict output values \mathbf{Y} based on input values \mathbf{X} .

and the average pressure

$$\langle P \rangle = \frac{P_{\text{upstream}} + P_{\text{downstream}}}{2} \quad (5)$$

over/in the Coriolis mass flow sensor channel.

The combination of sensing structures, i.e., the Coriolis mass flow/density sensor and the pressure sensors, can be used to estimate the viscosity of the fluid. Following to the Hagen–Poiseuille law, the relation between the pressure drop ΔP and the mass flow \dot{m} is:

$$\Delta P = \frac{8\eta L}{\rho\pi r_{\text{eff}}^4} \dot{m} = \frac{8\nu L}{\pi r_{\text{eff}}^4} \dot{m}, \quad (6)$$

with η the dynamic viscosity, ν the kinematic viscosity, L the length of the channel, ρ the density of the fluid and r_{eff} the effective radius [25]. This is a rough estimation, as it assumes an incompressible Newtonian fluid with uniform density without acceleration with a laminar fluid flow. Nevertheless, also in non-ideal cases, e.g., when the fluid is compressible, the viscosity still has a positive effect on the pressure drop. Using the Hagen–Poiseuille law as a function of the kinematic viscosity compared to dynamic viscosity is convenient when mass flow is measured, since the relation is independent of density.

Summarized, the sensing structures are together able to measure the fluid’s mass flow, pressure drop, density, and viscosity. Although rough linear estimations of the physical sensing principles are given, the actual relations are more complex. This holds especially for compressible fluids and fluids with a very low density (e.g., gases). Furthermore, other complex physical effects might give interesting information of the fluid. The magnitudes of the higher order harmonics of the Coriolis mass flow sensor are for example pressure dependent [12]. Since the raw signals of these sensors contain fluid-dependent and fluid-independent information, they could be used to classify the fluid.

2.3. Machine learning algorithms

As depicted in Fig. 2, a machine learning algorithm is an algorithm that is able to find rules $\mathbf{Y} = g(\mathbf{X})$ from training data, usually pairwise data of input vector \mathbf{X} and output vector \mathbf{Y} . Traditionally, scientists are supposed to find these (physical) rules, as e.g., given by Eqs. (1), (2), (3), and (6). In contrast, machine learning researchers design algorithms to approximate these rules for scientists.

In 1997, Mitchell [8] provided a solid definition “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E ”. In this paper, we formulate our fluid classification problem using the aforementioned definition of machine learning.

Given the sampled discrete-time signals from the Coriolis mass flow sensor and two pressure sensors, we aim at training a machine learning model which can classify four different fluid types, such as nitrogen, water, isopropanol, ethanol, and acetone. Note that it is necessary to incorporate a settling time between changing conditions, so the discrete-time sampled signals have been measured in steady-state. Regarding the definition of machine learning above, the model

is asked to specify which of n ($= 5$) fluid types the test sample belongs to (task T). By comparing the predicted fluid types and their actual labels (the ‘truth’), a performance measure P can be defined. The model improves performance with experience E after each iteration of comparing the predicted fluid type with the label.

This type of task is so-called classification [26,27]. To solve this task, the learning algorithm has to learn a way to produce a function g

$$\mathbf{Y} = g(\mathbf{X}) \quad (7)$$

with the input $\mathbf{X} \in \mathbb{R}^{m \times \tau}$ of m signals with each τ samples, and $\mathbf{Y} \in \mathbb{N}^{1 \times n}$. In other words, the model assigns the sensory measurements, described by the input vector \mathbf{X} to a fluid type identified by numeric code y . By using the labeled fluid types (experience E), identifying the fluid types (T), and evaluating performance (P), the model learns to produce the function g which forms a probability distribution over the four fluid types.

In this paper, we use two recent machine learning algorithms, i.e., convolutional neural networks (CNN) [28] and Bidirectional Long Short-Term Memory (BiLSTM) [29,30], to produce the function g . Traditional deep neural networks (DNN) use matrix multiplications to describe the correlation between input and output. For each DNN layer, if there are $m \times \tau$ inputs and n outputs, the matrix multiplication requires $m \times \tau \times n$ parameters and the runtime is proportional to $m \times \tau \times n$. With CNN, the number of neurons can be significantly reduced while it is still possible to obtain good accuracy. Using CNN also allows us to train the network models quicker since the learned signatures that represent the fluid type in one part of the signal can be reused in another part of the signal. In addition, we expect that the signatures are sparse in the input signals. CNN will learn to just focus on the signature, not the entire window of the signals.

While CNN is a very effective model for classification, it does not explore the temporal information in sequential data, whereas the sensing structures can provide data continuously. Exploring the temporal information might improve the classification accuracy. In fact, CNN also can share parameters across time when a sliding window is used. However, performance is still limited to the size of the sliding window. Therefore, in this paper we also apply recurrent neural networks (RNN) [29,30]. An RNN processes a sequence of input values \mathbf{x}^t , $\mathbf{x}^t \in \mathbb{R}^m$, where $t = 1, 2, \dots, \tau$, with τ the number of sequential input samples within a window of observation \mathbf{X} , $\mathbf{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^\tau\}$. The key idea to enable a recurrent neural network to exploit the temporal information is sharing parameters θ across time. In an RNN, the output of state \mathbf{S}^t is a function of the previous state \mathbf{S}^{t-1} , signal input \mathbf{x}^t and the sharing parameters output θ used to parametrize g ,

$$\mathbf{S}^t = g(\mathbf{S}^{t-1}, \mathbf{x}^t; \theta), \quad (8)$$

where the states \mathbf{S}^t contains information about the whole past sequence, which are the fluid types in our application.

In practice, RNN is also not able to capture the information which happened further back in the past, i.e., detecting long-term dependencies. The basic problem is the vanishing gradient problem when that gradients propagated over many states. The exponentially smaller weights given to long-term interactions also remove long-term dependencies as a weight w is multiplied by itself many times. The product w^t will either vanish or explode (rarely) depending on the magnitude of w . To overcome the limitations of a regular RNN on long-term dependencies, Long Short-Term Memory (LSTM) [29] is adopted in addition to CNN.

Future input signals coming up later than the current time $t + t_1$ is usually also useful for prediction. With a recurrent neural network, this can be achieved by delaying the output of RNN and LSTM. In practice, this method is not expected to be effective, especially when t_1 is very large. The RNN could increasingly concentrates on remembering the input information up to $t + t_1$ to predict the output \mathbf{Y}^t . In other words, less knowledge from previous input vectors was used for the prediction.

To this end, Bidirectional Long Short-Term Memory (BiLSTM) [31] is adopted as most effective method. BiLSTM is capable of learning long-term dependencies in the past and future of a specific time frame [31, 32].

To investigate whether the predictions made by the models are reliable and reasonable, the classification decisions of the trained models should be analyzed after training. Gradient-weighted class activation mapping (Grad-CAM) technique [33] utilizes the gradient of the network's classification score with regards to its determined convolutional features to identify the crucial aspects of the data for classification. In the case of our time-series sensor data, Grad-CAM determines the time steps that have the greatest impact on the network's classification decision.

3. Fabrication and experimental setup

Before the mentioned classification methods can be applied, the described sensor chip has to be fabricated and connected to a fluidic and electronic interfacing setup. Then, the signals from the sensor chip together with the fluid labels need to be digitized and stored for different conditions of the fluid.

3.1. Fabrication

The device is fabricated using Surface Channel Technology [34,35]. Fig. 3 shows an illustration of the most relevant steps of this process. Microchannels are realized by low-pressure chemical vapor deposition of a low-stress silicon-rich silicon nitride (SiRN) layer (Fig. 3b), etching of a slit pattern in this layer by reactive ion etching (Fig. 3c) and isotropic plasma etching in the highly p-type doped silicon wafer (Fig. 3d). Inlets are realized by wafer-through deep reactive-ion etching from the backside (Fig. 3e). A second deposition step of SiRN is performed to form the channel and inlet walls (Fig. 3f). Metal layers are deposited using sputtering (chromium and gold) to form the electrodes (Fig. 3g) and are etched using ion beam etching. A final isotropic plasma etch is performed to release the channels from the silicon. A microscopy image is shown in Fig. 4.

When the wafer is separated into chips, the chips are adhesively bonded onto a printed circuit board and wirebonded. Interfacing to the chip board is performed by the lab's standard 3D-printed fluidic connector and pogo-pin connections [37].

3.2. Experimental setup

Fig. 5 shows an illustration of the experimental setup for fluid and temperature control. The liquids are applied to the sensor using a reservoir in a pressure chamber. A pressure controller (Bronkhorst® EL-PRESS P-602CV) is used to control the pressure of the liquid at the inlet of the chip. A flow sensor with control valve (Bronkhorst® mini CORI-FLOW™ ML120V21) controls the flow through the chip. Extra pressure sensors (uncalibrated Honeywell sensors from the 24PC series) have been included to record the pressure changes in real-time together with the on-chip measurements, but the data has not been used in the analysis. The chip with electronic and fluidic interfacing is placed in an incubator (Peltier-cooled incubator IPP55 from Memmert) to apply different temperatures (see Fig. 6). A photograph of the setup is shown in Fig. 6.

The sensing structures on the chip are resistive (Wheatstone bridge) for the pressure sensors and capacitive for the Coriolis mass flow sensor. The Wheatstone bridge signals are recorded using a custom-built synchronous detector, consisting of a differential amplification stage, demodulation at the carrier wave frequency with offset correction and a low-pass filter.

The Coriolis mass flow sensor is placed in a constant magnetic field and actuated using Lorentz actuation. The actuation current is provided by a custom made oscillator circuit that amplifies signal

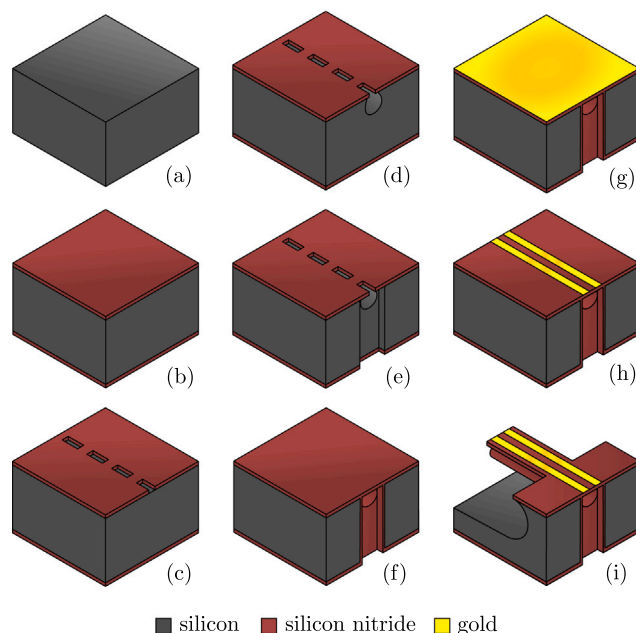


Fig. 3. Illustration of the most relevant fabrication steps of Surface Channel Technology [34,35]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Source: Reprinted from [36].

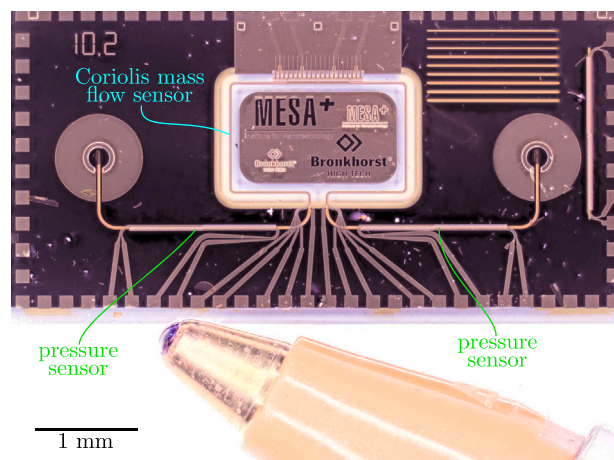


Fig. 4. Microscope image of the integrated upstream pressure sensor and part of the Coriolis channel.

from the electrodes of the Coriolis channel. This circuit in combination with the Coriolis channel form an electromechanical oscillator which frequency is only dependent on the mechanics as described by Eq. (2). The capacitive signals of the Coriolis mass flow sensor are detected using a charge amplifier with virtual ground to reduce parasitic effects. Similar to the resistive readout circuitry, demodulation of the carrier wave and low-pass filtering is performed to reduce the signals outside the band of interest. All analog signals are recorded using an analog-to-digital converter (ADC) card (NI PCI-6143) and stored on a hard drive. Effectively, the Coriolis mass flow sensor signals are sampled with 250 kS s^{-1} for 1 s. One sample is captured for each pressure sensor and is copied 250,000 times to match the length of the Coriolis mass flow sensor signals. A schematic representation of the readout circuitry is shown in Fig. 7.

Table 1

Overview of the five different fluids used in the experiments with values for the density, kinematic viscosity and dynamic viscosity for a gauge pressure of 5 bar. Similar values are highlighted. All parameters are estimated using Bronkhorst® FLUIDAT® [38].

Fluid	Abbreviation	Phase	Temperature (K)	Density (kg m^{-3})	Kinematic viscosity ($\text{mm}^2 \text{s}^{-1}$)	Dynamic viscosity (mPas)
Nitrogen	N ₂	Gas	288	7.05	2.44	$1.73 \cdot 10^3$
Nitrogen	N ₂	Gas	308	6.58	2.76	$1.82 \cdot 10^3$
Water	H ₂ O	Liquid	288	999	1.16	1.15
Water	H ₂ O	Liquid	308	995	0.74	0.74
Isopropanol	IPA	Liquid	288	800	3.40	2.72
Isopropanol	IPA	Liquid	308	778	1.94	1.51
Ethanol	EtOH	Liquid	288	812	1.53	1.27
Ethanol	EtOH	Liquid	308	791	1.10	0.87
Acetone	Ace	Liquid	288	805	0.42	0.34
Acetone	Ace	Liquid	308	800	0.36	0.28

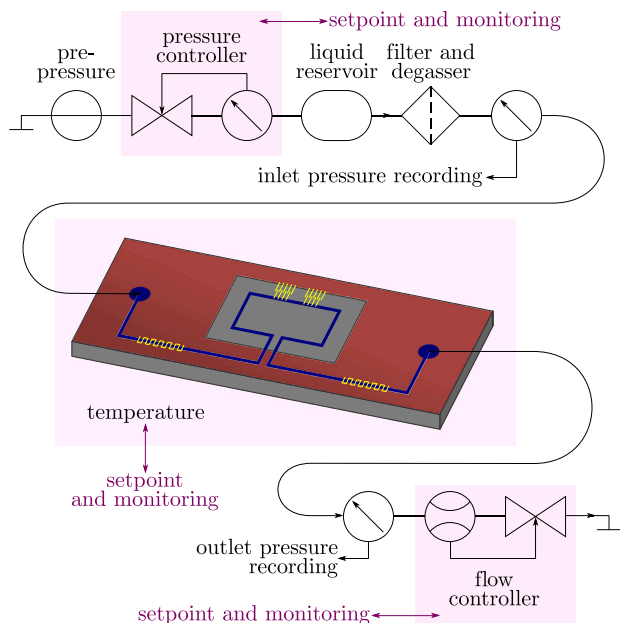


Fig. 5. Illustration of the experimental setup for fluid and temperature control. Note that there are four variables that can be controlled in this setup: the type of fluid, the temperature, the pressure and the mass flow.

3.3. Conditions

Various environmental and fluid conditions are applied to test the performance of the sensor and classification algorithms. All combinations of temperatures between 288 K and 308 K in steps of 5 K, pressures between 4 bar and 6 bar in steps of 0.5 bar, and mass flows between 0 g h^{-1} and 5 g h^{-1} in steps of 0.3 g h^{-1} to 1 g h^{-1} have been applied. The applied conditions are measured by the reference sensors. The measured conditions are presented in Fig. 8.

All experiments have been conducted with five different fluids: nitrogen (N₂), water (H₂O), isopropanol (IPA, propan-2-ol), ethanol (EtOH) and acetone (Ace, propan-2-one). Relevant parameters for these experiments, i.e., density and viscosity, for 288 K and 308 K for 5 bar are presented in Table 1. Isopropanol, ethanol and acetone have similar densities, which forces any successful classification method to be based on more than density sensing alone. Furthermore, water and ethanol have similar dynamic viscosities, although their kinematic viscosities are different.

3.4. Pre-processing

The dataset consists of 5540 streams, with each 4 signals of 250,000 samples captured in 1 s. First, the data as indicated by the brown-shaded areas in Fig. 8 is removed from the set. This data corresponds

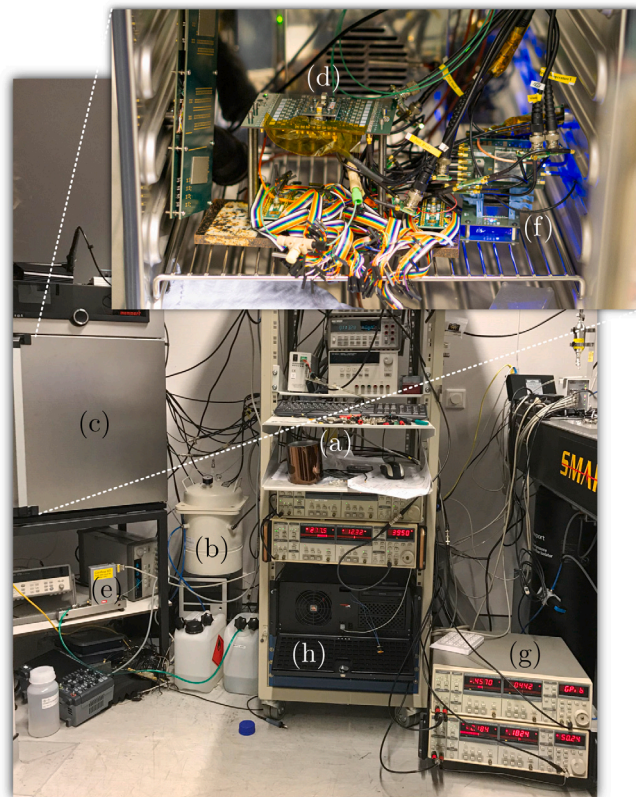


Fig. 6. Photograph of the experimental setup with (a) pressure controller (not clearly visible), (b) pressure vessel with liquid reservoir, (c) incubator, (d) sensor chip, (e) flow controller, (f) electronics as illustrated in Fig. 7, (g) lock-in amplifiers for debugging, (h) computer with digitizers and storage.

to measurements where fluids have just been exchanged and redundant measurements. Since a stream of one second at 250 kS s^{-1} is longer than needed for signals in the kHz-regime, each stream is split into 50 parts, with each part being 5000 samples (i.e., 20 ms, > 40 periods at resonance). A low mass flow can cause a pressure drop that is too low to sense. Therefore, five different datasets with different thresholds for the mass flow are extracted, i.e., $\geq 4 \text{ g h}^{-1}$ (all data with mass flows lower than 4 g h^{-1} is removed), $\geq 3 \text{ g h}^{-1}$, $\geq 2 \text{ g h}^{-1}$, $\geq 1 \text{ g h}^{-1}$ and no threshold (leaving all data in the dataset). This means that no classification has been performed for nitrogen when low mass flows ($< 2 \text{ g h}^{-1}$) are removed from the dataset.

The four signals are then subjected to the feature extraction algorithm. The average pressure and the pressure drop of the Coriolis mass flow sensor channel is calculated from the two pressure sensor signals. The magnitude spectrum and phase spectrum are calculated

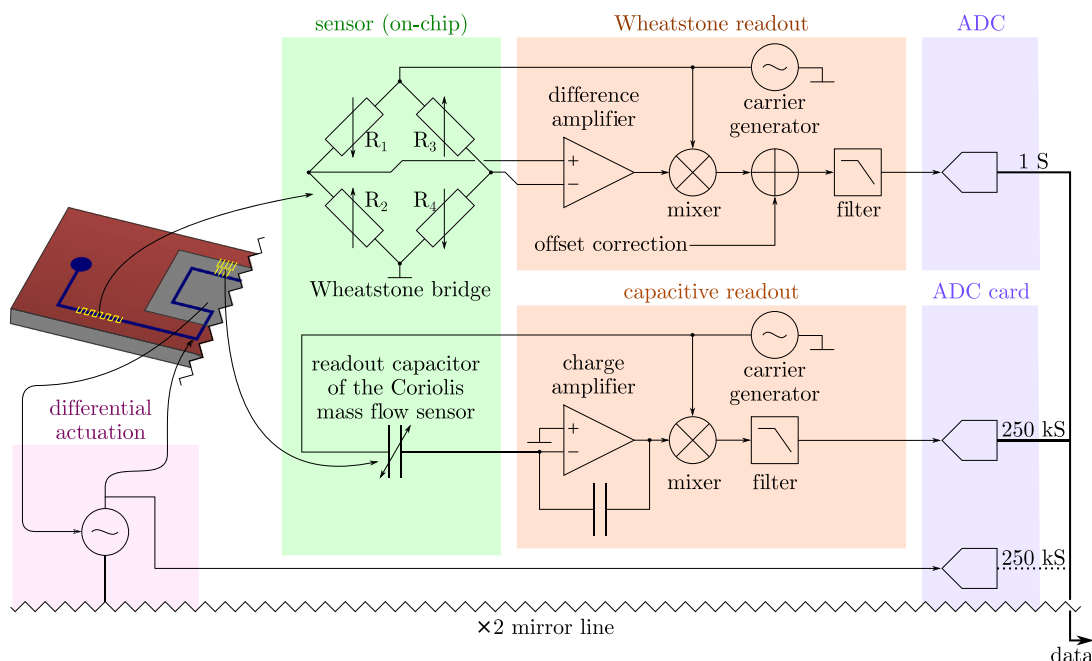


Fig. 7. Illustration of the electronic measurement setup. Everything above the $\times 2$ mirror line has been implemented twice.

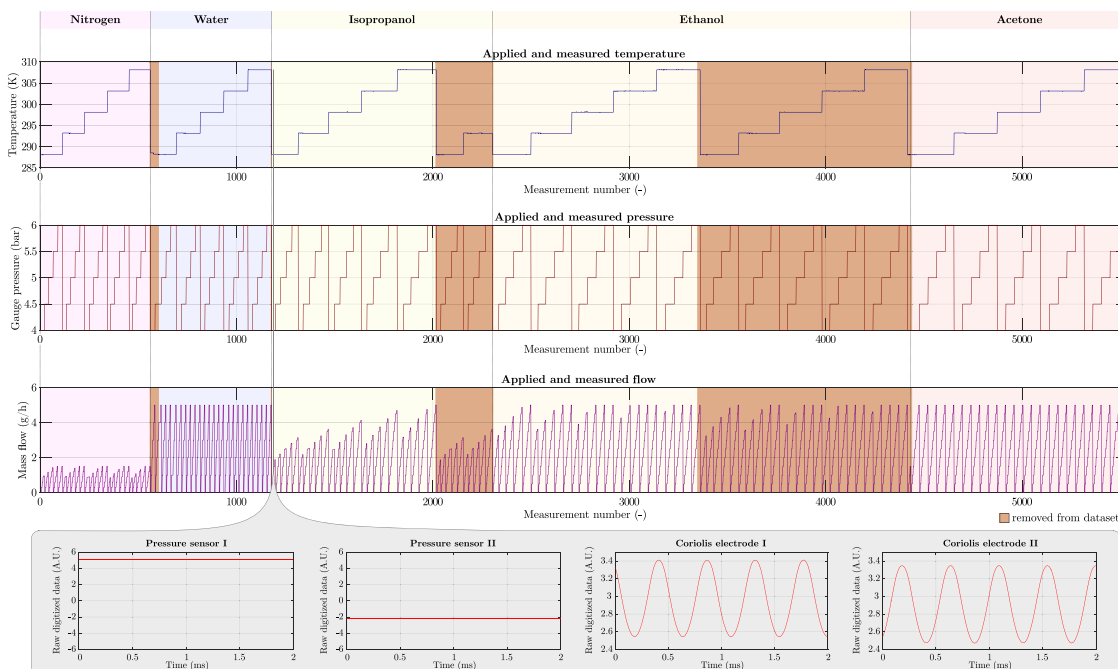


Fig. 8. Illustration of the applied conditions to the chip, based on measurement data from the reference sensors. All combinations of fluids, temperature, pressure and mass flow have been applied. The brown-shaded intervals have been removed from the dataset. The bottom gray area show (partly) the four signals that are recorded during each condition. A settling time between measurements without data acquisition has been incorporated. The pressure sensor signals are based on one repeated sample to match the sample length of the signals from the Coriolis electrodes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from the Coriolis mass flow sensor electrodes by Fast Fourier Transforms (FFT). A notch filter at resonance and harmonics is applied to the phase spectrum to emphasize the relevant Coriolis-induced phase shift. A schematic overview of the pre-processing algorithm, including the feature extraction, is shown in Fig. 9.

3.5. Training

Neural networks based on both BiLSTM and CNN have been trained using MATLAB[®] R2023a with the Deep Learning Toolbox[™]. An

overview of the CNN layer stack is given in Table 2 [27,39]. The architecture consists of a sequence input layer, two one-dimensional convolution layers with rectifying linear units and layer normalization layers, one-dimensional global average pooling layer, fully connected layer, softmax layer and classification layer. The CNN architecture results in 7076 learnable parameters. On the other hand, the BiLSTM layer stack consists of a sequence input layer, BiLSTM layer, fully connected layer, softmax layer and classification layer 3. The BiLSTM architecture results in 84804 learnable parameters, 12X more than the CNN model. To make the training lightweight and generalized

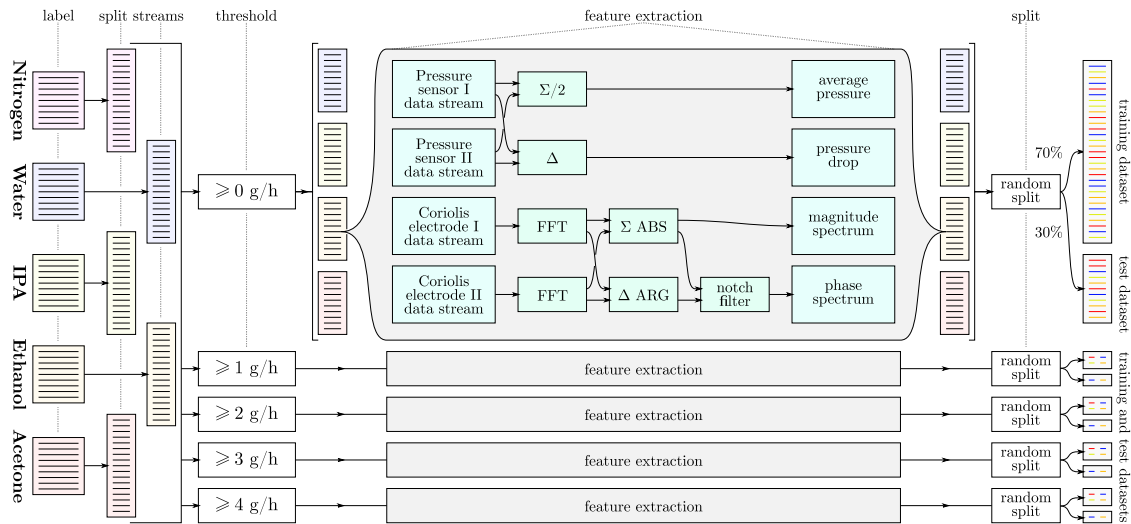


Fig. 9. Schematic overview of the pre-processing algorithm. The streams of data are split into multiple streams and then filtered for two different mass flow thresholds. Features are then extracted from each stream. The resulting dataset is then split into a set for learning and a set for testing. In addition to this train-test split validation method, also K-fold cross-validation is performed.

Table 2
Used CNN architecture with eleven layers based on MATLAB[®] syntax [27,39].

Layer	Type	Description
1	Sequence input	Sequence input with 4 dimensions (2 when pressure sensor signals are excluded)
2	1D Convolution	32 3×4 convolutions with stride 1 and padding
3	ReLU	Rectified linear unit layer
4	Layer normalization	Layer normalization with 32 channels
5	1-D Convolution	64 3×32 convolutions with stride 1 and padding 'causal'
6	ReLU	Rectified linear unit layer
7	Layer normalization	Layer normalization with 64 channels
8	1D Global average pooling	1D global average pooling
9	Fully Connected	Fully connected layer with output size 4 or 5, dependent on the inclusion of nitrogen
10	Softmax	Softmax layer
11	Classification output	Crossentropyex of '1' and 3 other classes

with very noisy and sparse input signals, we use the Adam optimizer [40]. The Adam optimizer combines the advantages of two other extensions of stochastic gradient descent, Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp). AdaGrad maintains a per-parameter learning rate that improves performance on problems with sparse gradients. Meanwhile, RMSProp adapts per-parameter learning rates based on the average of recent magnitudes of the gradients for the weight (e.g., how quickly it is changing). This means that the algorithm can handle online and non-stationary problems (e.g. noise). We have fine-tuned the models to a learning rate of 0.001. The maximum epochs were set to 3, mini-batch size is set to 150. To train and test the models, we have performed two different validation methods: train-test split using 70% of the data for training (X_{tr}) and 30% for testing (X) and K-fold cross-validation with $K = 10$. With the latter method, the dataset is split into ten groups. In ten cycles, each tenth of the dataset is used for testing and the rest for training. This results in ten accuracies of which the mean and standard deviation can be calculated as a measure for the reliability of the method.

3.6. Analytical benchmark

Parallel to testing the neural network, a simple fluid classification algorithm based on Eqs. (2) and (6) has been applied to the same testing parts of the dataset (X) as for the neural network training. First, measures for the resonance frequency and for the viscosity are extracted

Table 3
Used LSTM architecture with five layers based on MATLAB[®] syntax [27,39].

Layer	Type	Description
1	Sequence input	Sequence input with 4 dimensions (2 when pressure sensor signals are excluded)
2	BiLSTM	BiLSTM with 100 hidden units
3	Fully connected	Fully connected layer with output size 4 or 5, dependent on the inclusion of nitrogen
4	Softmax	Softmax layer
5	Classification output	Crossentropyex of '1' and 3 other classes

from the data. Based on these two parameters, decision trees with a maximum depth of four are automatically generated using MATLAB[®] R2023a. The decision trees are tested using the same testing part of the dataset (X_{tr}) as for the neural network testing. The results of these analytical benchmarks are labeled 'Analytical' in the results.

4. Results

A quick characterization has been performed to verify the condition of the Coriolis mass flow sensor. Fig. 10a shows the phase shift at resonance between the two capacitive signals for different mass flows for different fluids. A line, based on the linear relation in Eq. (1), has been fitted through the results and show a sensitivity of $2.7^\circ \text{h g}^{-1}$. Fig. 10b shows the resonance frequency for the different densities of the applied fluids. A line, based on the relation in Eq. (2), has been fitted through the results.

The classification analyses that have been performed can be categorized into the following combinations:

- Analytical, all sensor signals included, using train-test split validation;
- BiLSTM, all sensor signals included, using train-test split validation;
- BiLSTM, all sensor signals included, using K-fold cross-validation;
- BiLSTM, pressure sensor signals excluded, using train-test split validation;
- CNN, all sensor signals included, using train-test split validation;
- CNN, all sensor signals included, using K-fold cross-validation;
- CNN, pressure sensor signals excluded, using train-test split validation.

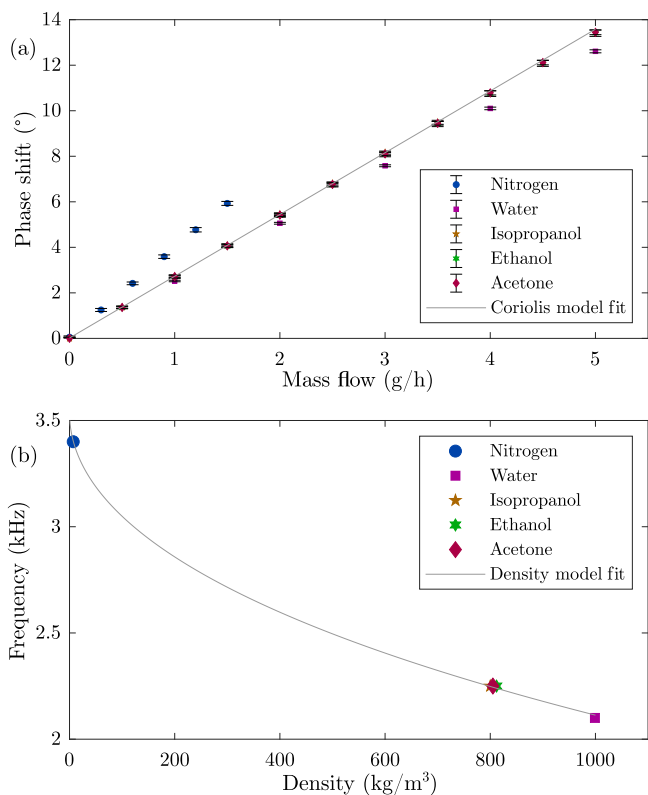


Fig. 10. Characterization plots for the Coriolis mass flow sensor. The Coriolis model fit and density model fit are based on Eqs. (1) and (2) respectively. The graphs are based on the same dataset as for classification, with data for 20 °C, for all applied pressures and split for different fluids. Reference values for the density are from Table 1, reference values for the flow are from the reference flow meter (Bronkhorst® mini CORI-FLOW™ ML120V21), rounded to the nearest setpoint to enable statistical analysis. Errorbars represent three times the standard deviation.

Then, for each combination, analyses are performed for a mass flow threshold of $\geq 4 \text{ g h}^{-1}$, $\geq 3 \text{ g h}^{-1}$, $\geq 2 \text{ g h}^{-1}$ and $\geq 1 \text{ g h}^{-1}$ and for no threshold, which sums up to a total of 35 analyses. Most important results, i.e. BiLSTM and CNN including all sensor signals using train-test split validation for $\geq 4 \text{ g h}^{-1}$, $\geq 2 \text{ g h}^{-1}$ and no threshold are presented in Tables 4, 5–9. The accuracies for all results (including K-fold cross-validation and excluding pressure sensors) are summarized in Table 10. Fig. 11 shows a graphical comparison of the different methods and their accuracies using train-test split validation.

To show how the model pays attention to the input signals, five random test examples from the five different fluid types have been analyzed using Grad-CAM [33]. Fig. 12 displays the four extracted feature signals with highlighted Grad-CAM importance for both BiLSTM and CNN, which emphasizes the areas utilized by the network in making its classification decision. In most cases, there is high importance at resonance. Yet, a high importance at resonance is not directly related to classification by density of the fluid, since the viscosity has influence at the phase at resonance as well.

5. Discussion

In this paper, we have shown that the functionality of existing microfabricated Coriolis mass flow sensors can be extended by machine learning. By post-processing, it is proven that fluid classification is possible without sensor redesign. After training, the model may even be implemented in embedded processing electronics for real-time fluid classification. This enables non-invasive physics-based fluid classification in in-line medical or industrial systems.

Table 4

Predicted values for four different liquids for flows of 4 g h^{-1} and higher using BiLSTM. Accuracy: 100%.

BiLSTM	Predicted					Recall
	$\geq 4 \text{ g h}^{-1}$	N ₂	H ₂ O	IPA	EtOH	
True	N ₂	–	–	–	–	–
	H ₂ O	–	2740	0	0	0
	IPA	–	0	745	0	0
	EtOH	–	0	0	3272	0
	Ace	–	0	0	0	4195
Precision	–	100%	100%	100%	100%	100%
F ₁ score	–	100%	100%	100%	100%	100%

Table 5

Predicted values for four different fluids for flows of 4 g h^{-1} and higher using CNN. Accuracy: 100%.

CNN	Predicted					Recall
	$\geq 4 \text{ g h}^{-1}$	N ₂	H ₂ O	IPA	EtOH	
True	N ₂	–	–	–	–	–
	H ₂ O	–	2740	0	0	0
	IPA	–	0	745	0	0
	EtOH	–	0	0	3272	0
	Ace	–	0	0	0	4195
Precision	–	100%	100%	100%	100%	100%
F ₁ score	–	100%	100%	100%	100%	100%

Table 6

Predicted values for four different fluids for flows of 2 g h^{-1} and higher using BiLSTM. Accuracy: 92.2%.

BiLSTM	Predicted					Recall
	$\geq 2 \text{ g h}^{-1}$	N ₂	H ₂ O	IPA	EtOH	
True	N ₂	–	–	–	–	–
	H ₂ O	–	5025	0	477	0
	IPA	–	3	3963	2231	0
	EtOH	–	1730	1037	6579	0
	Ace	–	101	0	0	10 146
Precision	–	73%	79%	71%	100%	100%
F ₁ score	–	81%	71%	71%	100%	100%

Table 7

Predicted values for four different fluids for flows of 2 g h^{-1} and higher using CNN. Accuracy: 76.6%.

CNN	Predicted					Recall
	$\geq 2 \text{ g h}^{-1}$	N ₂	H ₂ O	IPA	EtOH	
True	N ₂	–	–	–	–	–
	H ₂ O	–	2991	0	1811	700
	IPA	–	0	3888	2309	0
	EtOH	–	1049	1021	6947	329
	Ace	–	97	0	0	10 150
Precision	–	72%	79%	63%	91%	95%
F ₁ score	–	62%	70%	68%	95%	95%

Table 8

Predicted values for five different fluids for all mass flows using BiLSTM. Accuracy: 84.4%.

BiLSTM	Predicted					Recall
	All flows	N ₂	H ₂ O	IPA	EtOH	
True	N ₂	8268	0	0	26	0
	H ₂ O	0	6189	93	879	1358
	IPA	0	764	8691	87	2955
	EtOH	0	1133	1292	12 438	834
	Ace	0	35	47	56	16 312
Precision	100%	76%	86%	92%	76%	76%
F ₁ score	100%	74%	77%	85%	86%	86%

From Table 10, it appears that train-test split validation results in similar accuracies as K-fold cross-validation. This gives a strong

Table 9
Predicted values for five different fluids for all mass flows using CNN. Accuracy: 50.4 %.

CNN	Predicted					Recall	
All flows	N ₂	H ₂ O	IPA	EtOH	Ace		
True	N ₂	734	36	847	3604	3073	9 %
	H ₂ O	0	3198	0	2200	3121	38 %
	IPA	486	533	3561	5172	2745	28 %
	EtOH	360	1351	898	7490	5598	48 %
	Ace	0	310	0	130	16010	97 %
Precision	46 %	59 %	67 %	40 %	52 %		
F ₁ score	15 %	46 %	40 %	44 %	68 %		

Table 10
Overview of the performances for the discussed methods. Train-test split validation (TTS) as illustrated by Fig. 9 and K-fold cross validation. In case of K-fold cross-validation, also the standard deviation is given.

Method	Sensors	Validation	Mass flow threshold	Accuracy	Standard deviation
Analytical	All	TTS	≥ 4 g h ⁻¹	100 %	
BiLSTM	All	TTS	≥ 4 g h ⁻¹	100 %	
BiLSTM	All	K-fold	≥ 4 g h ⁻¹	100 %	0.3 %
BiLSTM	Cori. only	TTS	≥ 4 g h ⁻¹	47 %	
CNN	All	TTS	≥ 4 g h ⁻¹	100 %	
CNN	All	K-fold	≥ 4 g h ⁻¹	100 %	0.0 %
CNN	Cori. only	TTS	≥ 4 g h ⁻¹	47 %	
Analytical	All	TTS	≥ 3 g h ⁻¹	96 %	
BiLSTM	All	TTS	≥ 3 g h ⁻¹	94 %	
BiLSTM	All	K-fold	≥ 3 g h ⁻¹	93 %	2.3 %
BiLSTM	Cori. only	TTS	≥ 3 g h ⁻¹	46 %	
CNN	All	TTS	≥ 3 g h ⁻¹	93 %	
CNN	All	K-fold	≥ 3 g h ⁻¹	92 %	0.6 %
CNN	Cori. only	TTS	≥ 3 g h ⁻¹	40 %	
Analytical	All	TTS	≥ 2 g h ⁻¹	80 %	
BiLSTM	All	TTS	≥ 2 g h ⁻¹	82 %	
BiLSTM	All	K-fold	≥ 2 g h ⁻¹	81 %	5.5 %
BiLSTM	Cori. only	TTS	≥ 2 g h ⁻¹	31 %	
CNN	All	TTS	≥ 2 g h ⁻¹	77 %	
CNN	All	K-fold	≥ 2 g h ⁻¹	77 %	1.0 %
CNN	Cori. only	TTS	≥ 2 g h ⁻¹	41 %	
Analytical	All	TTS	≥ 1 g h ⁻¹	72 %	
BiLSTM	All	TTS	≥ 1 g h ⁻¹	85 %	
BiLSTM	All	K-fold	≥ 1 g h ⁻¹	77 %	9.6 %
BiLSTM	Cori. only	TTS	≥ 1 g h ⁻¹	76 %	
CNN	All	TTS	≥ 1 g h ⁻¹	63 %	
CNN	All	K-fold	≥ 1 g h ⁻¹	64 %	0.6 %
CNN	Cori. only	TTS	≥ 1 g h ⁻¹	35 %	
Analytical	All	TTS	-	69 %	
BiLSTM	All	TTS	-	84 %	
BiLSTM	All	K-fold	-	88 %	7.6 %
BiLSTM	Cori. only	TTS	-	64 %	
CNN	All	TTS	-	50 %	
CNN	All	K-fold	-	51 %	0.8 %
CNN	Cori. only	TTS	-	30 %	

positive indication for (a) the reliability of the used neural network configurations, (b) the reliability for using provided dataset for learning and (c) the reliability of the train-test split method.

As shown in Table 10, high accuracies cannot be achieved when excluding the pressure sensor data for high mass flows. Since isopropanol, ethanol and acetone have very similar densities (as specified in Table 1, classification cannot be performed well on resonance frequency only. Viscosity of the fluid has more influence on the pressure drop when the mass flow is higher. Data with only high mass flows show significantly higher accuracies, up to 100 % when all low-flow data of < 4 g h⁻¹ is removed. For low-flow data where (low-density) nitrogen comes into play, so statistically more accurate classifications can be performed on frequency alone as can be concluded from the confusion matrix in Table 8. Better pressure sensing could help in increasing the classification accuracy for low flows (and thus low pressure drops). The gold strain gauges of the pressure sensors may drift due to the high

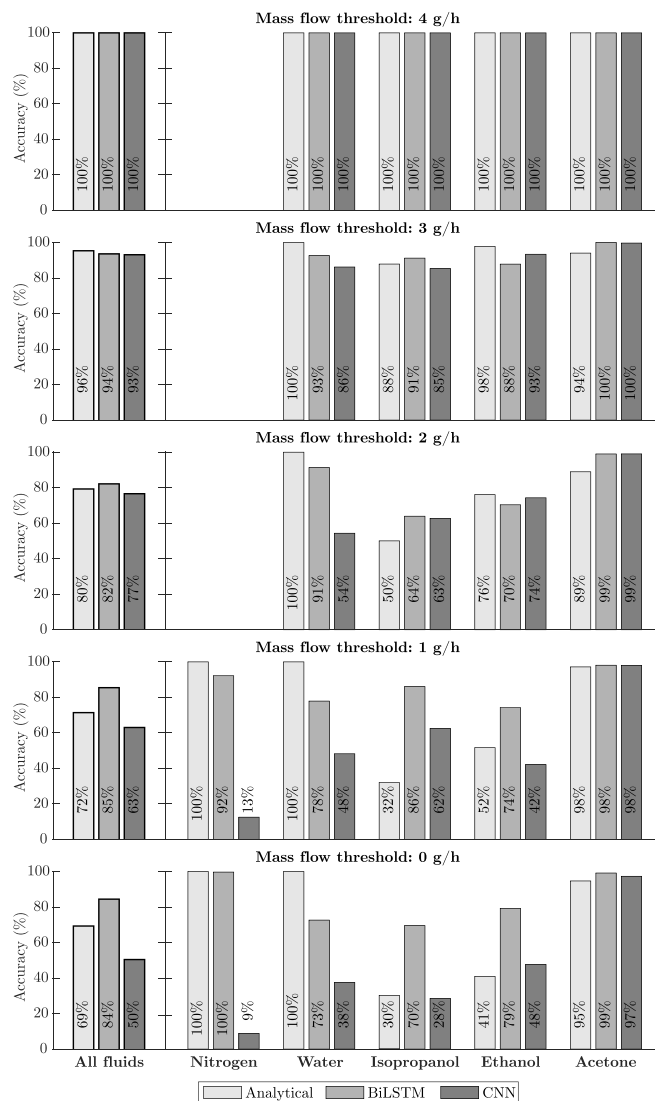


Fig. 11. Graphical representation of the accuracies from Table 10 for the three different classification methods and for the five different mass flow thresholds using test-train split validation.

malleability of gold [41], i.e., its ability to easily plastically deform under pressure-induced stresses. Replacing gold with platinum in the fabrication process might reduce these drift issues.

The functionality may further be extended, since the raw signals from the Coriolis mass flow sensor contain valuable information that is related to physical quantities other than mass flow. The magnitudes of the higher harmonics are, for example, pressure dependent [12]. In contrast to the conventional criteria to design the sensor as sensitive as possible for one specific parameter, the machine learning method presented in this article could benefit from a vibrating channel with dependencies on multiple parameters. As can already be concluded from the Grad-CAM analyses in Fig. 12, the relatively high importance at the higher harmonics show potential for the use in classification or more accurate fluid parameter measurements. Focusing more on higher harmonics in the feature extraction process might render the integrated pressure sensors redundant.

From Fig. 11, it can be concluded that BiLSTM neural networks perform better than convolutionary neural networks, especially for low flows. This is expected due to BiLSTM's better performance for sequential data, apparently even in the frequency domain. Although our BiLSTM model generally performs better than the CNN model, the

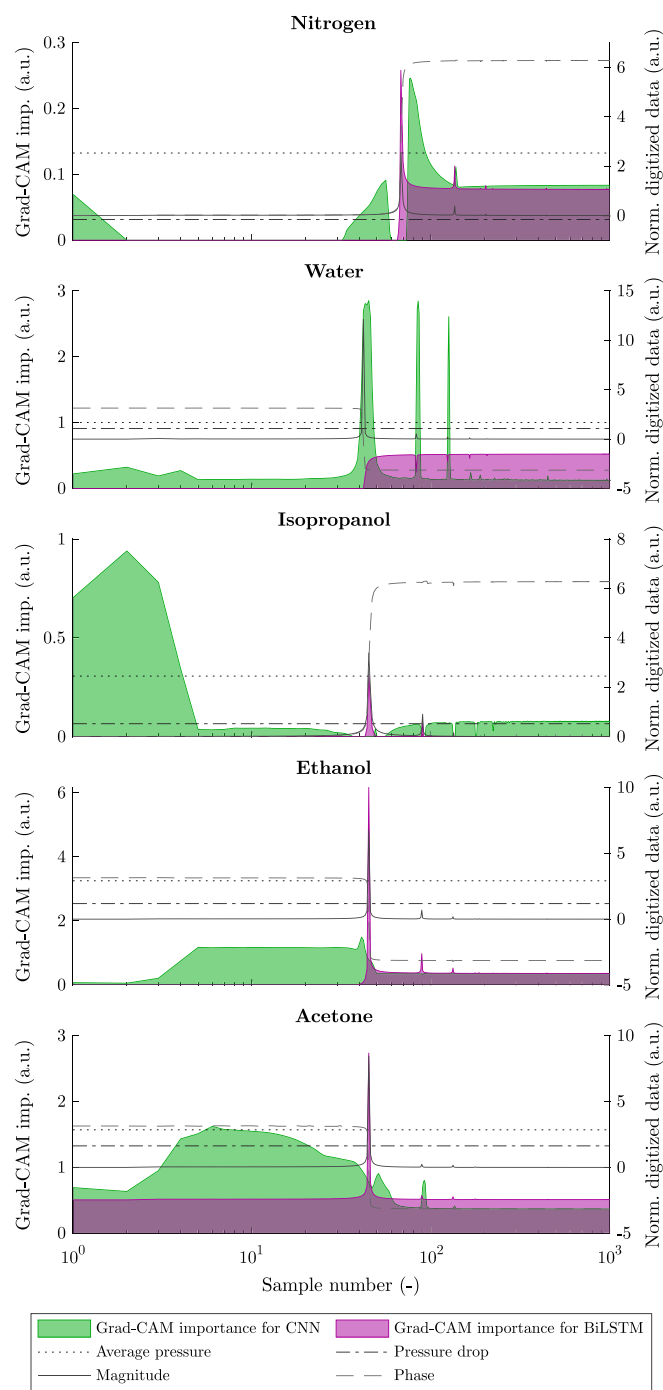


Fig. 12. Gradient-weighted class activation mapping analysis for a random extracted feature signal from the 0 g h^{-1} threshold set for the five classes. Note that there might be difference in conditions (i.e., mass flow, pressure and temperature) for the plotted analyses.

CNN model is $12\times$ lighter than the BiLSTM model. This is critical when deploying the models on tiny embedded devices which have limited memory capacity and computation power. Nevertheless, it requires further investigation on the trade-off of accuracy-latency and accuracy-resource consumption to come up with a decision on which model is more preferred. There is also room for improving the AI models in several ways. First, the designed models were simple and not fully optimized. Using a Neural Architecture Search (NAS) [42] would result in better CNN and BiLSTM models in terms of accuracy. Second, combining CNN and BiLSTM in one neural network architecture has been

shown more effective [43,44]. Finally, empirical study [45] showed that even LSTM can only maintain effectively the long term memory up to about 1000 previous states. In other words, the information of about the first 4000 samples of the observation windows in our experiments were not effectively retrieved. A solution is finding compromise of using shorter windows, lower sampling rates, and replacing BiLSTM with Transformer [46]. With the attention-based model of Transformer, all such 5000 samples in the observation windows will be considered together for the classification task.

Future work will focus on the classification of additional types of fluids and fluid compositions. Optimizations of the machine learning hyperparameters and the use of algorithms that might give more insight into the sensor's limitations (e.g., Bayesian neural networks) are planned. Furthermore, a new generation of sensing structures optimized for machine learning should be designed, possibly even computer-generated. Another challenge for the future lies in the regression of the physical parameters, e.g., density and viscosity.

6. Conclusion

Novel methods to classify fluids using a microfabricated Coriolis mass flow sensor with integrated pressure sensors and a trained machine learning model has been realized. Thanks to the chip's mechanisms that are sensitive to multiple fluid parameters, correct classification proved to be mostly independent on mass flow, pressure and temperature. A comparison has been made between machine learning methods BiLSTM and CNN and an analytical method, of which BiLSTM performs best with a cross-validated accuracy of 77% up to 100%, dependent on the inclusion of low-flow data. Improved pressure sensing can further reduce the lower flow limit.

Future work will focus on physics-based machine learning algorithms for fluid composition classification, physical parameter regression and computer-generated sensor designs.

CRediT authorship contribution statement

D. Alveringh: Conceptualization, Methodology, Software, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **D.V. Le:** Conceptualization, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **J. Groenesteijn:** Methodology, Software, Investigation, Resources, Writing – review & editing. **J. Schmitz:** Supervision, Funding acquisition, Writing – review & editing. **J.C. Lötters:** Supervision, Resources, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors gratefully acknowledge the reviewing support of K. Wintreaecken and the technical support of T. S. J. Lammerink, R. Zubavičius and R. G. P. Sanders. Furthermore, the authors also gratefully acknowledge the fruitful discussions with R. J. Wiegerink and P. J. M. Havinga and the financial support of the Dutch Ministry of Education, Culture and Science through Sectorplan Bèta en Techniek.

References

- [1] S. Silvestri, E. Schena, Micromachined flow sensors in biomedical applications, *Micromachines* 3 (2) (2012) 225–243.
- [2] F. Ejeian, S. Azadi, A. Razmjou, Y. Orooji, A. Kottapalli, M.E. Warkiani, M. Asadnia, Design and applications of MEMS flow sensors: A review, *Sensors Actuators A* 295 (2019) 483–502.
- [3] P. Enoksson, G. Stemme, E. Stemme, A Coriolis mass flow sensor structure in silicon, in: *Proceedings of Ninth International Workshop on Micro Electromechanical Systems*, IEEE, 1996, pp. 156–161.
- [4] J. Haneveld, T.S.J. Lammerink, M.J. de Boer, R.G.P. Sanders, A. Mehendale, J.C. Lötters, M. Dijkstra, R.J. Wiegerink, Modeling, design, fabrication and characterization of a micro Coriolis mass flow sensor, *J. Microelectromech. Syst.* 20 (12) (2010) 125001.
- [5] D. Alveringh, T.V.P. Schut, R.J. Wiegerink, W. Sparreboom, J.C. Lötters, Resistive pressure sensors integrated with a Coriolis mass flow sensor, in: *2017 19th International Conference on Solid-State Sensors, Actuators and Microsystems, TRANSDUCERS, IEEE*, 2017, pp. 1167–1170.
- [6] T.V.P. Schut, D. Alveringh, W. Sparreboom, J. Groenesteijn, R.J. Wiegerink, J.C. Lötters, Fully integrated mass flow, pressure, density and viscosity sensor for both liquids and gases, in: *2018 IEEE Micro Electro Mechanical Systems, MEMS, IEEE*, 2018, pp. 218–221.
- [7] S.F. Wong, J.S. Lim, S.S. Dol, Crude oil emulsion: A review on formation, classification and stability of water-in-oil emulsions, *J. Pet. Sci. Eng.* 135 (2015) 498–504.
- [8] T.M. Mitchell, *Machine Learning*, vol. 1, (9) McGraw-Hill New York, 1997.
- [9] Z. Ballard, C. Brown, A.M. Madni, A. Ozcan, Machine learning and computation-enabled intelligent sensor design, *Nat. Mach. Intell.* 3 (7) (2021) 556–565.
- [10] J. Zhu, X. Liu, Q. Shi, T. He, Z. Sun, X. Guo, W. Liu, O.B. Sulaiman, B. Dong, C. Lee, Development trends and perspectives of future sensors and MEMS/NEMS, *Micromachines* 11 (1) (2019) 7.
- [11] H. Xue, M. Zhang, P. Yu, H. Zhang, G. Wu, Y. Li, X. Zheng, A novel multi-sensor fusion algorithm based on uncertainty analysis, *Sensors* 21 (8) (2021) 2713.
- [12] D. Alveringh, R.J. Wiegerink, J.C. Lötters, Integrated pressure sensing using capacitive Coriolis mass flow sensors, *J. Microelectromech. Syst.* 26 (3) (2017) 653–661.
- [13] M. Egmont-Petersen, D. de Ridder, H. Handels, Image processing with neural networks—A review, *Pattern Recognit.* 35 (10) (2002) 2279–2301.
- [14] D.R. Reddy, Speech recognition by machine: A review, *Proc. IEEE* 64 (4) (1976) 501–531.
- [15] L. Deng, X. Li, Machine learning paradigms for speech recognition: An overview, *IEEE Trans. Audio Speech Lang. Process.* 21 (5) (2013) 1060–1089.
- [16] S. Neethirajan, The role of sensors, big data and machine learning in modern animal farming, *Sens. Bio-Sens. Res.* 29 (2020) 100367.
- [17] R. Al-Ruzouq, M.B.A. Gibril, A. Shanableh, A. Kais, O. Hamed, S. Al-Mansoori, M.A. Khalil, Sensors, features, and machine learning for oil spill detection and monitoring: A review, *Remote Sens.* 12 (20) (2020) 3338.
- [18] U. Yaqoob, M.I. Younis, Chemical gas sensors: Recent developments, challenges, and the potential of machine learning—A review, *Sensors* 21 (8) (2021) 2877.
- [19] R.P. Liu, M.J. Fuent, M.P. Henry, M.D. Duta, A neural network to correct mass flow errors caused by two-phase flow in a digital Coriolis mass flowmeter, *Flow Meas. Instrum.* 12 (1) (2001) 53–63.
- [20] L. Wang, J. Liu, Y. Yan, X. Wang, T. Wang, Gas-liquid two-phase flow measurement using Coriolis flowmeters incorporating artificial neural network, support vector machine, and genetic programming algorithms, *IEEE Trans. Instrum. Meas.* 66 (5) (2016) 852–868.
- [21] Y. Wang, S. Wang, C. Decès-Petit, A Bayesian method for on-line evaluation of uncertainty in measurement of Coriolis flow meters, *Measurement* 179 (2021) 109448.
- [22] J. Riordon, D. Sovilj, S. Sanner, D. Sinton, E.W. Young, Deep learning with microfluidics for biotechnology, *Trends Biotechnol.* 37 (3) (2019) 310–324.
- [23] E.A. Galan, H. Zhao, X. Wang, Q. Dai, W.T.S. Huck, S. Ma, Intelligent microfluidics: The convergence of machine learning and microfluidics in materials science and biomedicine, *Matter* 3 (6) (2020) 1893–1922.
- [24] R.R.A. Groen, Applying Artificial Intelligence on an Integrated Pressure and Mass Flow Sensor (BSc. thesis), University of Twente, 2020.
- [25] S.P. Sutera, R. Skalak, The history of Poiseuille’s law, *Annu. Rev. Fluid Mech.* 25 (1) (1993) 1–20.
- [26] S.B. Kotsiantis, I. Zaharakis, P. Pintelas, et al., Supervised machine learning: A review of classification techniques, *Emerg. Artif. Intell. Appl. Comput. Eng.* 160 (1) (2007) 3–24.
- [27] N.K. Chauhan, K. Singh, A review on conventional machine learning vs deep learning, in: *2018 International Conference on Computing, Power and Communication Technologies, GUCON, IEEE*, 2018, pp. 347–352.
- [28] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* 86 (11) (1998) 2278–2324.
- [29] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780.
- [30] M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks, *IEEE Trans. Signal Process.* 45 (11) (1997) 2673–2681.
- [31] A. Graves, J. Schmidhuber, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural Netw.* 18 (5–6) (2005) 602–610.
- [32] T. Osogami, H. Kajino, T. Sekiyama, Bidirectional learning for time-series models with hidden units, in: *International Conference on Machine Learning, PMLR*, 2017, pp. 2711–2720.
- [33] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, GradCam: Visual explanations from deep networks via gradient-based localization, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 618–626.
- [34] J. Groenesteijn, M.J. de Boer, J.C. Lötters, R.J. Wiegerink, A versatile technology platform for microfluidic handling systems, part I: Fabrication and functionalization, *Microfluid. Nanofluid.* 21 (7) (2017) 1–14.
- [35] J. Groenesteijn, M.J. de Boer, J.C. Lötters, R.J. Wiegerink, A versatile technology platform for microfluidic handling systems, part II: Channel design and technology, *Microfluid. Nanofluid.* 21 (7) (2017) 1–12.
- [36] D. Alveringh, *Integrated Throughflow Mechanical Microfluidic Sensors* (Ph.D. thesis), University of Twente, 2018.
- [37] D. Alveringh, R.G.P. Sanders, J. Groenesteijn, T.S.J. Lammerink, R.J. Wiegerink, J.C. Lötters, Universal modular fluidic and electronic interfacing platform for microfluidic devices, in: *Proceedings of the 3rd Conference on Microfluidic Handling Systems, MFHS 2017*, 2017, pp. 106–109.
- [38] Bronkhorst High-Tech B.V., FLUIDAT on the net, mass flow and physical properties calculations, 2023, <https://www.fluidat.com/>. (Accessed 7 August 2023).
- [39] MathWorks, List of deep learning layers, 2023, <https://www.mathworks.com/help/deeplearning/ug/list-of-deep-learning-layers.html>. (Accessed 15 February 2023).
- [40] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.
- [41] M.R. Neuman, W.G. Sutton, Structural dependence of strain gauge effect and surface resistivity for thin gold films, *J. Vac. Sci. Technol.* 6 (4) (1969) 710–713.
- [42] S. Sinha, D.V. Le, Completely automated CNN architecture design based on VGG blocks for fingerprinting localisation, in: *2021 International Conference on Indoor Positioning and Indoor Navigation, IPIN, IEEE*, 2021, pp. 1–8.
- [43] A. Yasmeeen, F.I. Rahman, S. Ahmed, M.H. Kabir, CSVC-Net: Code-switched voice command classification using deep CNN-LSTM network, in: *2021 Joint 10th International Conference on Informatics, Electronics & Vision and 2021 5th International Conference on Imaging, Vision & Pattern Recognition, ICIEV and ICIVPR, IEEE*, 2021, pp. 1–8.
- [44] P. Kavianpour, M. Kavianpour, E. Jahani, A. Ramezani, A cnn-bilstm model with attention mechanism for earthquake prediction, 2021, arXiv preprint arXiv:2112.13444.
- [45] S. Li, W. Li, C. Cook, C. Zhu, Y. Gao, Independently recurrent neural network (indrn): Building a longer and deeper rnn, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5457–5466.
- [46] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, L. Sun, Transformers in time series: A survey, 2022, arXiv preprint arXiv:2202.07125.

Dr. ir. Dennis Alveringh received a Ph.D. degree on the subject of microfabricated flow and pressure sensors in 2018. He has worked in industry in R&D on semiconductor and MEMS testing for multiple years before returning back to academia. He is currently an assistant professor in the Integrated Devices and Systems group (member of the MESA+ Institute at the University of Twente) and focuses on MEMS- and CMOS-based sensor systems.

Dr. Le Viet Duc received a Ph.D. on the subject of symbiotic sensing and AI in 2016. He is Assistant Professor at the Department of Computer Science at the University of Twente. Inspired by the way human brain learns the world, he is eager to develop sustainable AI models for IoT systems that is robust to change in data distribution caused by new environments, bridging the gap between lab and world. His focus is on finding efficient methods to learn, transfer and distil knowledge hidden in different datasets, especially unlabeled ones. As he has 10-years R&D experience in developing smart and wearable and industrial devices, he is also interested in applying his research to industrial products for a sustainable world.

Dr. ir. Jarno Groenesteijn received a Ph.D. degree on the subject of microfabricated Coriolis-based sensor and actuator systems in 2016. Hethen joined Bronkhorst High-Tech B.V. where he works closely together with the Integrated Devices and Systems group (member of the MESA+ Institute at the University of Twente) on next-generation microfabricated flow sensors.

Prof. dr. Jurriaan Schmitz received a Ph.D. degree (1994) in experimental physics at the University of Amsterdam on research carried out at the Nikhef research institute. He worked for Philips Research on CMOS scaling, characterization and reliability before joining the University of Twente as a full professor. He is currently chairing the Integrated Devices and Systems group (member of the MESA+ Institute at the University of Twente).

Prof. dr. ir. Lötters received a Ph.D. degree on the subject of triaxial capacitive accelerometers in 1997. He then joined Bronkhorst High Tech B.V. where he has concentrated his research on flow measurement and control. He currently combines his research for Bronkhorst High Tech B.V. with a part-time full professorship in the Integrated Devices and Systems group (member of the MESA+ Institute at the University of Twente) on microfluidic handling systems.