Applied Acoustics 202 (2023) 109161

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Towards the detection and classification of indoor events using a loudspeaker

Patrick Marmaroli, Mark Allado, Romain Boulandet *

HEPIA, HES-SO University of Applied Sciences and Arts Western Switzerland, Rue de la Prairie 4, CH-1202 Geneva, Switzerland

ARTICLE INFO

Article history:
Received 1 September 2022
Received in revised form 24 November 2022
Accepted 29 November 2022
Available online 15 December 2022

Keywords:
Loudspeaker
Room acoustics
Room modes
Artificial intelligence
Acoustic security
Acoustic impedance
Acoustic monitoring

ABSTRACT

In recent decades, smart building has received considerable research attention due to the increased demand for connected and integrated technology. Based on data collected by sensors, alarms, lighting, access control, heating and cleaning can be adjusted according to human activity and the actual needs of the occupants, resulting in efficient energy management and operating cost savings. However, in most cases, these sensors are application-specific, which limits their usefulness and scalability. Of the available sensing technologies, acoustic methods often rely on the use of microphones, which can lead to privacy issues. In this work, we use an electrodynamic loudspeaker in combination with a convolutional neural network algorithm to extract and classify the features of indoor events from the sound field. We show how the loudspeaker impedance is sensitive, through the modal response of the room, to changes in occupancy or room layout (presence of people, movement or removal of furniture), door or window opening, or temperature variation. This gives the speaker a new functionality in addition to audio broadcasting. Theoretical analysis and experiments in real rooms demonstrate the accuracy and effectiveness of this acoustic-based approach for supervised classification of indoor events.

1. Introduction

Nowadays, research and innovation in smart buildings are attracting more and more interest in the construction and building management society. Sensing technologies are enabling areas such as lighting, heating, ventilation and air-conditioning (HVAC), access control and many others to become more connected and integrated, making buildings truly smart. By detecting indoor occupancy, it becomes possible to adjust thermostat settings and HVAC operations or secure sensitive spaces more efficiently, improving working conditions and safety while saving costs and energy [1,2]. Among them, camera-based detection is very popular, with the disadvantage of having to store large amounts of data and especially of interfering with privacy issues. Passive infrared (PIR) sensing is widely used for lighting management, automatic door opening, intrusion detection, etc. However, this non-intrusive approach has several weaknesses. The sensor is sensitive to changes in the environment and may cause a false alarm if there is heat or cold flow from an HVAC system or direct exposure to sunlight [3]. In addition, people who are not moving cannot be detected. Environmental sensors typically used for measuring temperature, relative humidity, or CO_2 concentration can also be used to estimate occupancy over time [4]. The main drawback, however, is a low accuracy and a strong dependence on external atmospheric conditions. Also worth mentioning are magnetic reed switches used to improve resident safety by detecting any unauthorized entry or intrusion through a door or window. More recently, acoustic methods have been proposed for room occupancy estimation [5], occupant behaviour monitoring and emergency event detection [6], or as a complement to a home security system for intrusion detection [7]. However, these techniques involving microphones can raise privacy concerns in the workplace by inferring personal information from voice or speech recordings.

The alternative approach proposed in this research is to use an electrodynamic loudspeaker as an impedance sensor to relate small changes in the room's modal response to the indoor events that produced them. Generally speaking, when a room is excited by a sound source, the input of acoustic energy at modal frequencies causes standing waves called room modes [8]. This results in an uneven distribution of acoustic energy with large pressure fluctuations influenced by various factors such as room size ratio, room shape, surface absorption and wall configuration. Room modes are known to alter the free-field response of electrodynamic loudspeakers and are one of the main obstacles to accurate sound reproduction in listening rooms [9–12]. The acoustic coupling of the listener/loudspeaker to modal pressure variations is termed

* Corresponding author.
E-mail address: romain.boulandet@hesge.ch (R. Boulandet).

https://doi.org/10.1016/j.apacoust.2022.109161
0003-682X/

modal coupling. Near the pressure nodes, the acoustic load impedance exerted by the room to the loudspeaker diaphragm is very low, resulting in a low radiation efficiency. In contrast, in the vicinity of the anti-nodes, the acoustic load impedance is very large, which greatly increases the radiation efficiency of the loudspeaker. The loudspeaker impedance therefore reflects some characteristics of the room at a given location and can be used to study the behavior and performance of sound sources in rooms at low frequencies.

The acoustic response of the room to a source depends not only on the position of the source, but also on internal damping and scattering by non-absorbing objects [13–16]. The damping of room modes usually comes from the amount of absorbing materials through furniture or acoustic treatment. Well below the Schroeder frequency, however, porous or fibrous materials such as textiles, curtains, carpets, acoustic foam panels or mineral wool have a relatively weak effect. In contrast, the walls can be vibrated by the sound waves, thus dissipating some of the acoustic energy from the room's modes [17]. Adding or removing a piece of furniture, on the other hand, can slightly distort the mode shapes and shift the natural frequencies of the room [17]. This also applies between an occupied and an empty room. Depending on the wavelength, the sound waves will be partly diffracted around the body and partly absorbed. The absorption of sound energy in this case is mainly dominated by the amount and type of clothing [18]. By monitoring the acoustic impedance of the loudspeaker diaphragm over time, it is therefore possible to detect different events resulting from changes in the room, simply by examining the evolution of the room modes. This, coupled with the fact that a loudspeaker cannot record sensitive data, unlike a security camera, makes it a good candidate for use where security and privacy are at stake.

In this work, we investigate how a conventional electrodynamic loudspeaker can be used to 1) monitor room occupancy to know the presence of people in the workplace in real time, 2) detect a change in perimeter (opening of a door or window) to alert if an intruder enters a protected area, 3) track a thermal anomaly to warn of an unusual temperature variation (e.g. fire).

The rest of the paper is structured as follows. Section 2, provides the theoretical background for describing the acoustic coupling between a loudspeaker and the room modes. A laboratory experiment conducted in a meeting room is presented in Section 3 to demonstrate the feasibility of the proposed methodology. Through a 25-h test campaign, we show how the impedance of the loudspeaker diaphragm is affected by room occupancy, perimeter change or thermal anomaly. The classification problem is addressed in Section 4 using a convolutional neural network algorithm trained to recognise within labelled data sets the recurrent patterns specific to the events of interest. We conclude with a discussion and prospects for future work.

2. Theoretical background

In this section, we present the modal theory for studying the dynamic interaction between a loudspeaker and the room modes. In the following model, the boundary surfaces are considered to be locally reacting, so that the impedance of the walls is independent of the angle of incidence of the sound waves. The effect of room walls on internal damping can therefore be fully characterized by an impedance boundary condition which depends on spatial coordinates and frequency.

2.1. Modal decomposition model

Considering a point source of volume flow Q in $r_0 = (x_0, y_0, z_0)$, the pressure field in $r = (x, y, z)$ can be expressed as a series of normal modes [8]:

$$p(x,y,z) = j\omega\rho Q \sum_n \frac{p_n(r)p_n(r_0)}{\Lambda_n\left(k_n^2 - k^2\right)} \tag{1}$$

where ω is the driving angular frequency, ρ is the density of air, c is the speed of sound, k_n, p_n and Λ_n denote respectively the eigenvalues, the eigenfunctions and a coefficient which depend on the assumptions made about the walls.

2.1.1. Rigid wall

With rigid walls and assuming that the mode coupling is small enough to be neglected, these quantities are real and can be written as:

$$k_n = k_{n0} = \sqrt{k_x^2 + k_y^2 + k_z^2} \tag{2}$$

$$p_n(x,y,z) = \cos(k_x x)\cos(k_y y)\cos(k_z z) \tag{3}$$

$$\Lambda_n = \frac{V}{\epsilon_{n_x}\epsilon_{n_y}\epsilon_{n_z}} \tag{4}$$

where $k_x = n_x\pi/L_x, k_y = n_y\pi/L_y, k_z = n_z\pi/L_z$ are the wave numbers associated with each dimension, n_x, n_y, n_z are non-negative integers that indicate the numbers of nodal planes perpendicular to the x-axis, y-axis and z-axis, respectively, and $\epsilon_{n_x} = 1$ for $n_x = 0$ and $\epsilon_{n_x} = 2$ for $n_x > 0$. Similar expressions are used of ϵ_{n_y} and ϵ_{n_z}. The corresponding natural frequency is therefore given by

$$f_n = \frac{c}{2}\sqrt{k_x^2 + k_y^2 + k_z^2} \tag{5}$$

2.1.2. Finite wall impedance

If the walls are not perfectly rigid but have a uniform specific admittance $\beta = \xi + j\sigma$ such that $\beta kL \ll 1$, where L is one of the dimension of the room, it can be shows that [19]:

$$k_n = k_{n0} + j\frac{k}{2k_{n0}}\left(\epsilon_{n_x}\frac{\beta_{x0} + \beta_{xL_x}}{L_x} + \epsilon_{n_y}\frac{\beta_{y0} + \beta_{yL_y}}{L_y} + \epsilon_{n_z}\frac{\beta_{z0} + \beta_{zL_z}}{L_z}\right) \tag{6}$$

$$p_n(x,y,z) = \cos\left(\eta_{xn_x}\frac{\pi x}{L_x} - j\beta_{x0}\frac{kL_x}{\pi\eta_{xn_x}}\right)\cos\left(\eta_{yn_y}\frac{\pi y}{L_y} - j\beta_{y0}\frac{kL_y}{\pi\eta_{yn_y}}\right)$$
$$\cos\left(\eta_{zn_z}\frac{\pi z}{L_z} - j\beta_{z0}\frac{kL_z}{\pi\eta_{zn_z}}\right) \tag{7}$$

$$\Lambda_n = V\left(\frac{1}{2} + \frac{\sin(\pi\eta_{xn_x})}{2\pi\eta_{xn_x}}\right)\left(\frac{1}{2} + \frac{\sin(\pi\eta_{yn_y})}{2\pi\eta_{yn_y}}\right)\left(\frac{1}{2} + \frac{\sin(\pi\eta_{zn_z})}{2\pi\eta_{zn_z}}\right) \tag{8}$$

where,

$$\begin{aligned} \eta_{x0} &= -\tfrac{1}{j\pi}\sqrt{jkL_x(\beta_{x0} + \beta_{xL_x})} \quad n_x = 0 \\ \eta_{xn_x} &= n_x + j\tfrac{kL_x}{\pi^2 n_x}(\beta_{x0} + \beta_{xL_x}) \quad n_x \geqslant 1 \end{aligned} \tag{9}$$

Similar expressions are used for η_{yn_y} and η_{zn_z}.

2.1.3. Walls with low real admittance

Assuming now a real specific admittance at the walls, i.e. $\beta = \xi$, the damping factor δ_n for each mode n is defined as [12,20]

$$\delta_n = \frac{c}{2}\left(\epsilon_{n_x}\frac{\beta_{x0} + \beta_{xL_x}}{L_x} + \epsilon_{n_y}\frac{\beta_{y0} + \beta_{yL_y}}{L_y} + \epsilon_{n_z}\frac{\beta_{z0} + \beta_{zL_z}}{L_z}\right) \tag{10}$$

If now ξ can be considered as uniform for each wall, Eq. (10) can be written

$$\delta_n = \left(\frac{\epsilon_{n_x}}{L_x} + \frac{\epsilon_{n_y}}{L_y} + \frac{\epsilon_{n_z}}{L_z}\right)c\xi \tag{11}$$

Substituting ξ by δ_n in Eqs. (6)–(9), Eq. (1) can be rewritten:

$$p(x,y,z) = j\omega\rho c^2 Q \sum_n \frac{p_n(r)p_n(r_0)}{\Lambda_n(\omega_n^2 - \omega^2 - 2j\omega\delta_n)} \tag{12}$$

The average damping factor δ_n in Eq. (12) can be obtained experimentally without prior knowledge of the impedance of the walls. When the room modal density is low, which is true in a lightly damped room well below the Schroeder frequency [21], δ_n is related to the modal decay time $MT60_n$ of an eigenmode n as [8,20]

$$MT60_n = \frac{3\ln(10)}{\delta_n} = \frac{3\ln(10)}{c\xi}\frac{V}{(\epsilon_{n_x}L_yL_z + \epsilon_{n_y}L_xL_z + \epsilon_{n_z}L_xL_y)} \tag{13}$$

which is analog to the reverberation time required for the sound energy of a single mode to decay by 60 dB according to Sabine's formula.

The modal decomposition method discussed above is applicable when the boundary surfaces of the room have a large and real impedance such as massive or very stiff walls, which are therefore not sound-absorbing. For room with randomly distributed impedance surfaces, moreover, approximation methods must be used to analytically calculate the eigenmodes [19].

2.2. Radiation of a piston in a room

Let us now consider a square piston mounted into a wall in the plane $x = 0$ in the rectangular room, as shown in Fig. 1. The piston coordinates are $y_0 < y < y_0 + a$ and $z_0 < z < z_0 + a$.

The position of the source is such that only the front radiation is taken into account. This excludes many real-world situations but reduces considerably the complexity of the analytical expression. It is assumed that the piston oscillates with a vibratory speed v normal to the wall at the driving frequency ω. In a rigid wall configuration, the boundary conditions are therefore given by:

$$\left.\frac{\partial p}{\partial x}\right|_{x=L_x} = \left.\frac{\partial p}{\partial y}\right|_{y=0} = \left.\frac{\partial p}{\partial y}\right|_{y=L_y} = \left.\frac{\partial p}{\partial z}\right|_{z=0} = \left.\frac{\partial p}{\partial z}\right|_{z=L_z} = 0 \quad \text{and}$$
$$\left.\frac{\partial p}{\partial x}\right|_{x=0} = \begin{cases} j\omega\rho v, & (y,z) \in S \\ 0, & (y,z) \notin S \end{cases} \tag{14}$$

The sound pressure field in the room can be expressed as [22]:

$$p(x,y,z) = j\omega\rho\frac{4v}{L_yL_z}\sum_{n_y,n_z}\frac{\alpha_{n_y}\beta_{n_z}}{\varepsilon_{n_y}\varepsilon_{n_z}}\frac{\cos\left(k_x(x-L_x)\right)}{k_x\sin(k_xL_x)} \times \cos\left(k_yy\right)\cos\left(k_zz\right) \tag{15}$$

where $k_x = \sqrt{k^2 - k_y^2 - k_z^2}$, $\varepsilon_0 = 2, \varepsilon_j = 1$ for $j \neq 0, \alpha_0 = \beta_0 = a$,

$$\alpha_{n_y \geqslant 1} = \frac{\sin\left(k_y(y_0+a)\right) - \sin(k_yy_0)}{k_y} \quad \text{and} \quad \beta_{n_z \geqslant 1} = \frac{\sin\left(k_z(z_0+a)\right) - \sin(k_zz_0)}{k_z} \tag{16}$$

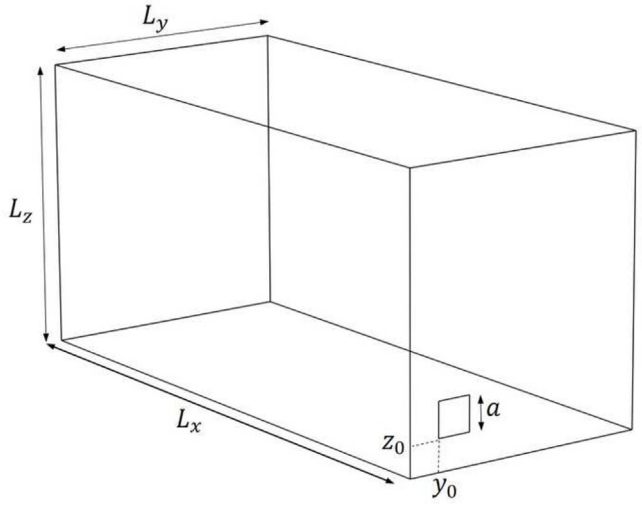

Fig. 1. Location of the square piston in the rectangular room.

The total reaction force acting on the piston of surface area $S = a^2$ is given by:

$$F = \int\int_S p(0,y,z)\,dydz \tag{17}$$

and the radiation impedance of the piston can therefore be obtained from Eqs. (15) and (17) as [22]:

$$Z_{mr}(\omega) = \left.\frac{F}{v}\right|_{x=0} = j\frac{4\rho\omega}{L_yL_z}\sum_{n_y,n_z}\frac{\cot\left(k_xL_x\right)}{\varepsilon_{n_y}\varepsilon_{n_z}k_x}\alpha_{n_y}^2\beta_{n_z}^2 \tag{18}$$

Eq. (18) describes how the room affects the radiation from the loudspeaker at the diaphragm by exerting a complex frequency-dependent additional load. Any changes at the room boundaries such as opening a door or window will therefore affect the load impedance seen by the loudspeaker. Similarly, the speed of sound involved in Eq. (18), which is a function of room temperature, is also expected to have a quantifiable effect on this impedance as we shall see in Section 3.5.

2.3. Computed results

For comparison with the closed-form expression given in Eq. (18), a numerical simulation was performed with COMSOL Multiphysics 5.6 using the Pressure Acoustics, Frequency Domain interface. We have used the 3D geometrical model shown in Fig. 1 where the loudspeaker is modelled as a square piston vibrating with a normal velocity v and a surface area of 0.125 m^2. The walls are modelled as a sound hard boundary in the physics interface. The specific acoustic impedance is calculated as the ratio of the spatial average of the acoustic pressure to the vibratory speed of the piston. The interior of the room is filled with air with $c = 343$m/s $\rho = 1.2$ kg/m^3 for the speed of sound and density, respectively.

Fig. 2 illustrates the frequency response of a rectangular room of dimensions (5.97 m × 2.885 m × 2.615 m) calculated with sound hard walls. It compares the specific acoustic impedance of the square piston calculated using Eq. (18) with that computed numerically using the FEM simulation software. As can be shown in Fig. 2, a good agreement can be observed between the two

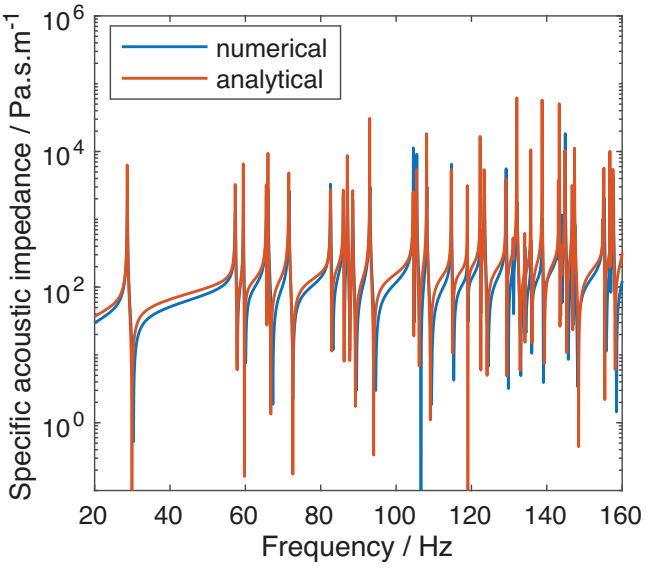

Fig. 2. Plot of numerical versus analytical frequency response functions of the specific acoustic radiation impedance of a square piston in an undamped rectangular room, with a 0.1 Hz frequency resolution.

frequency response functions. It can be seen that the specific acoustic impedance has alternating maxima and minima which reflect the fluctuation of pressure at the position of the loudspeaker in the room.

3. Validation experiments

3.1. Description

In this section, we present the results of a test campaign conducted in a meeting room of dimensions 4.98 m $\times$ 3.24 m $\times$ 3.04 m made of drywall and glazing. The objective is to show how certain events related to human activity can affect the diaphragm impedance of a loudspeaker in the room. In this experiment, the effect of opening/closing doors and windows, an empty or occupied room, and a change in room temperature is specifically evaluated. As can be seen in Fig. 3, the room is furnished with a meeting table and has a door on one side and two windows on the opposite side.

A 6-inch closed-box loudspeaker is placed on the floor in a corner of the room throughout the experiment. To obtain the impedance of the loudspeaker diaphragm, we used a Brüel & Kjaer Type 4954 1/4 inch microphone to pick up the sound pressure at about 1 cm from the diaphragm and a Polytec IVS-500 laser Doppler vibrometer to measure the normal velocity of the membrane near the connection between the dust cap and the cone. Signals are acquired at a sampling rate of 4096 Hz with a Brüel & Kjaer Type 3160 data acquisition hardware. Each measurement is 2 min long. Data acquisition was repeated every 10 min for a total duration of 25 h. This resulted in the recording of 153 measurements. Throughout the experiment, the loudspeaker continuously played low-level white noise fed by an input voltage of about 30 mV RMS.

The status of the parameters during each of the 14 phases of the experiment is given in Table 1. Configurations 1, 4, 6, 8, 10, 12 and 14 are all identical and are referenced by a star symbol ($*$). They correspond to the baseline scenario characterized by closed doors and windows, no chairs or sound absorbing panels, and no people in the room. In configuration 7, we used a single panel of Dinaphon B 810 (bulk density 8.5–11.5 kg m^{-3}, different pore sizes with a maximum of 10 pores of 5–15 mm diameter per square meter) open cell melamine foam with a total effective volume of $1.2 \times 0.6 \times 0.05 = 0.036$ m^3. Configuration 9, consisting of 97 measurements, corresponds to a variation in ambient temperature

Fig. 3. Meeting room used for the validation experiment; a) view of the room from the door (left) and from the windows (right) and b) map of the meeting room.

Table 1
Scenario timeline of the experiment.

scenario	1*	2	3	4*	5	6*	7
date	14/06/21	14/06/21	14/06/21	14/06/21	14/06/21	14/06/21	14/06/21
start time	11.19 am	12.18 pm	1.08 pm	1.48 pm	2.37 pm	3.17 pm	3.57 pm
stop time	12.08 pm	12.58 pm	1.38 pm	2.27 pm	3.07 pm	3.47 pm	4.27 pm
nb. meas.	6	5	4	5	4	4	4
door status	C	O	O	C	C	C	C
window 1	C	O	C	C	C	C	C
window 2	C	C	C	C	C	C	C
body	–	–	–	–	1	–	–
chairs	–	–	–	–	1	–	–
foam panel	–	–	–	–	–	–	1
avg. temp (°C)	22.4	22.6	23.2	23.5	24.2	24.8	25.4
avg. hum (%)	25.8	25	26	27	27.8	28.8	29

scenario	8*	9	10*	11	12*	13	14*
date	14/06/21	14/06/21	15/06/21	15/06/21	15/06/21	15/06/21	15/06/21
start time	4.37 pm	5.16 pm	9.20 am	10 am	10.39 am	11.19 am	11.59 am
stop time	5.06 pm	9.10 am	9.50 am	10.30 am	11.09 am	11.49 am	12.29 pm
nb. meas.	4	97	4	4	4	4	4
door status	C	C	C	O	C	C	C
window 1	C	O	C	C	C	C	C
window 2	C	O	C	C	C	C	C
body	–	–	–	–	–	–	–
chairs	–	–	–	–	–	9	–
foam panel	–	–	–	–	–	–	–
avg. temp (°C)	26.1	25.1	24.7	25.1	25.9	26.5	26.9
avg. hum (%)	30	32	30.5	28.2	27.2	26.5	26.5

obtained by leaving the two windows open overnight. In Table 1, "C" stands for closed and "O" stands for open. The temperature and relative humidity were taken during each measurement and the average values for each configuration are reported in Table 1. Throughout the experiment, the temperature and relative humidity varied between 22.4 and 26.9°C, and between 25% and 32% respectively. The effect of temperature on the modal response of the room is discussed in Section 3.5.

3.2. Event detection

In this section, we examine how a specific event can impact on the loudspeaker impedance, in order to perform a supervised classification.

Fig. 4 depicts in light grey solid lines the 31 measurements of the modulus of the loudspeaker specific acoustic impedance (SAI) - defined as the complex-valued ratio between the total acoustic pressure and the diaphragm velocity - that have been obtained in the baseline configuration described previously. They are displayed between 80 Hz and 650 Hz with a frequency resolution of 0.0625 Hz. Assuming these measurements follow a normal distribution for each frequency bin, about 99.7% of values are contained within three standard deviations of the mean. The upper and lower control limits are plotted as solid black lines and can be thought as the "three-sigma limits" into which an SAI measured in the baseline situation should fall. In the same graphs, an example of an indoor event (opening of two windows) is depicted in solid red line. It is clear that certain room modes are largely affected by this event, as its corresponding SAI falls outside the three-sigma limits. Therefore, this kind of event can be easily detected by a trained algorithm and alert to a potential security issue in the room.

An example of simple detection algorithm is described in the following. Let's call s a Boolean of value 1 in case the room differs from the baseline situation, and 0 otherwise. A possible algorithm for estimating the value of s is:

$$s = \begin{cases} 1, & \text{if } N > \Lambda \times N_t \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

where N_t is the total number of frequency bins contained in the SAI frequency response, N is the number of frequency bins for which the modulus of the SAI is not contained within the upper and lower limits, and Λ is an arbitrary threshold between 0 and 1 (typically 0.1).

The confusion matrix obtained after applying such a decision rule over the 153 available measurements is given in Table 2. For these results, we set $\Lambda = 1/100 \times M$, where M is the number of frequency bins between 80 Hz and 280 Hz.

In Table 2, the missed detection corresponds to a realisation of scenario 13 which simply differs from the baseline by the presence of 9 chairs in the room. This is due to the very small effect that the addition of furniture has on the diaphragm impedance at the loudspeaker's location in the room. The algorithm was able to recognise that the data from the other three measurements in scenario 13 were different from the baseline scenario. All other indoor events involving the opening of windows and doors, the presence of people, the addition of absorbent foam panels were correctly detected.

3.3. Introducing impedograms

As shown in Section 2, the SAI allows to account for the load exerted by the room on the radiation from the loudspeaker. Inspired by *spectrograms* which are a well-known time–frequency representations of concatenated spectrum of signal chunks [23,24], the concatenation of acoustic impedance gives rise to what might be termed acoustic "impedograms" with time in seconds in the x-axis and the frequency in Hz in the y-axis; the color scale representing the module of the specific acoustic impedance, in Pa.s/m.

Raw signals are segmented into chunks of size L_w samples and overlap L_o samples, then (optionally) weighted by an apodization window. The acoustic impedance is estimated by computing the transfer function between normal velocity of the diaphragm sensed by a laser Doppler vibrometer and the microphone signal, for example by using the *tfestimate* function from Matlab. The concatenation of successive signal chunks forms 2D matrices of size $F \times N$ where F is the number of frequency bins, and N is the number of impedance curves, given by the following expression:

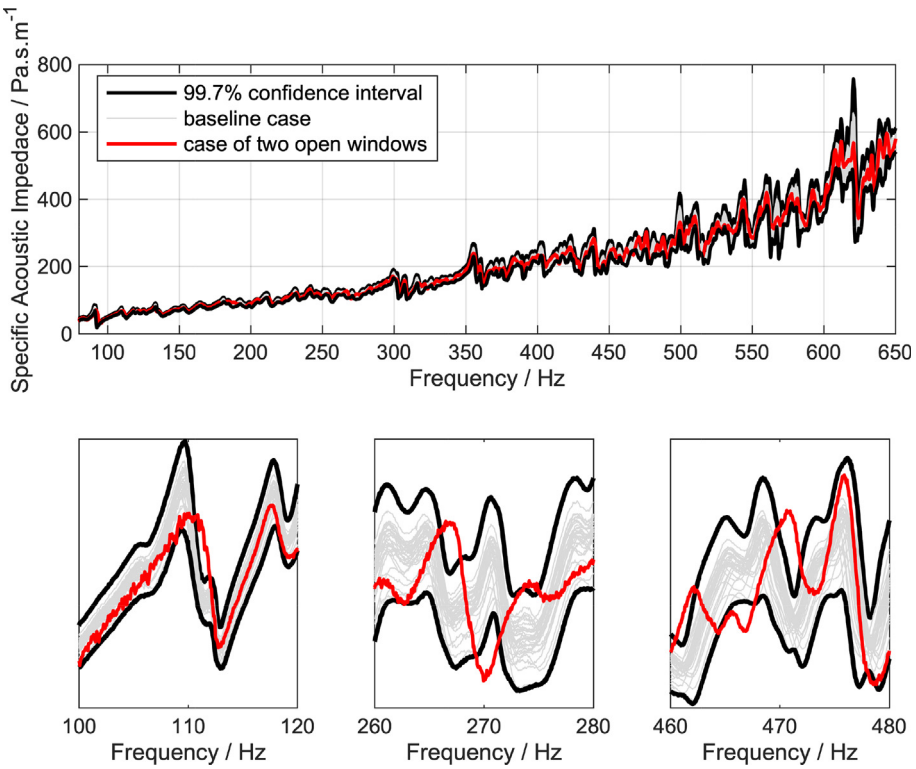

Fig. 4. Frequency response function of the loudspeaker specific acoustic impedance (SAI) showing the typical dispersion (excluding temperature) observed for the baseline scenario from 31 measurements (configurations 1, 4, 6, 8, 10, 12 and 14 in Table 1) compared to the case of the two open windows (one measurement extracted the 97 measurements of configuration 9 in Table 1). Bottom: zoom in on different frequency ranges showing how such a change in boundary conditions affect the modal response of the room, and thus the loudspeaker SAI.

Table 2
Confusion matrix for room change detection.

n = 153	a change actually occurred	no change actually occurred
a change is predicted	True positive: 121 (99.1%)	False alarm: 1 (0.9%)
no change is predicted	Missed detection: 1 (3.2%)	True negative: 30 (96.8%)

$$N = \left\lfloor \frac{L - L_w}{L_w - L_o} \right\rfloor + 1 \tag{20}$$

where L is the length of the raw signal, in samples, and $\lfloor . \rfloor$ stands for the floor operator.

3.4. Examples of impedograms

As an example, Figs. 5–7 show acoustic impedograms measured between 600 Hz and 650 Hz after a change in perimeter, a change in occupancy and a change in temperature, respectively.

Fig. 5 illustrates an intentional perimeter change consisting first of opening a window and a door (between measurements 6 and 7), followed by closing the window (between measurements 11 and 12), then the closing of the door (between measurements 15 and 16). The ambient temperature was stable throughout the measurement, around 26°C. It can be observed that the opening of a door or window substantially affects the frequency response of the room by damping, broadening or frequency shifting the room modes.

The scenario described in Fig. 6 consists of an intentional change of occupation in a meeting room. Successive measurements from 17 to 20 are taken with an empty room with only a table in the center, from 21 to 24 the room is occupied (one person sitting on a chair), and from 25 to 28 the room is empty again. The ambient temperature was relatively stable around 26.5°C throughout this experiment. As shown in Fig. 6, there is a noticeable attenuation of certain modes in the presence of a person and a chair in the meeting room, especially in the frequency range 600–650 Hz.

Fig. 7 illustrates the effect of varying the room temperature by deliberately leaving two windows open overnight, keeping the other parameters unchanged. It is clear that temperature variation produces an increasing frequency shift of all room modes. The frequency offset, in Hz, can be theoretically estimated as explained below.

Figs. 8–10 show the corresponding impedograms in the electrical domain, computed from the voltage and current at the loudspeaker input terminals. Note that the electrical and acoustic data (pressure and velocity at the loudspeaker diaphragm) used to calculate the electrical and acoustic impedograms were measured simultaneously. As shown in Fig. 8–10 b), however, the dynamic effect of the indoor events studied on the room modal response is less noticeable in the electrical impedograms than in the acoustic impedograms. This is due in particular to the principle of electrodynamic transduction, in which the internal impedance of the loudspeaker predominates over that of the acoustic load exerted by the room. The frequency responses shown in Figs. 8–10 c) clearly illustrate the mechanical resonance of the moving part of the loudspeaker (here around $f_s = 125$ Hz) accounts

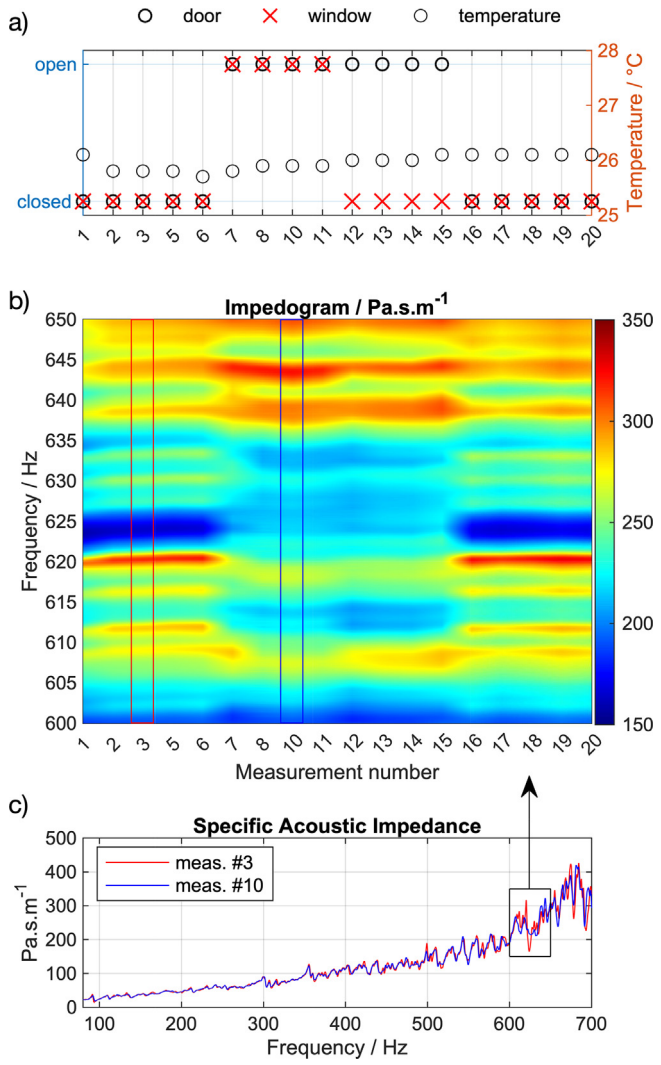

Fig. 5. Evolution of room modes over time depending on a perimeter change; a) gives the state of the room boundary conditions between measurements 1 and 20 and the temperature, b) shows the corresponding impedogram between 600 Hz and 650 Hz resulting from the concatenation of 20 SAIs, and c) shows the frequency response functions of the SAI for measurements 3 (door and windows closed) and 10 (door and two windows open) as an example.

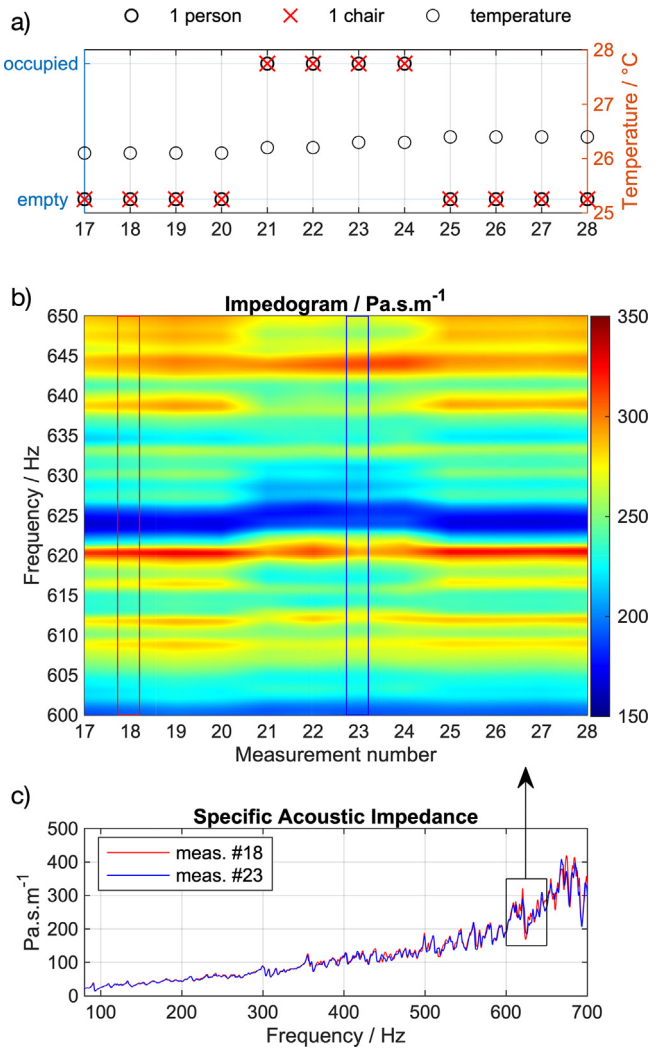

Fig. 6. Evolution of room modes over time due to a change in occupation; a) gives the occupancy status of the room and temperature between measurements 17 and 28, b) shows the corresponding impedogram between 600 Hz and 650 Hz resulting from the concatenation of 12 SAIs, and c) shows the frequency response functions of the SAI for measurements 18 (empty room) and 23 (one person sitting on a chair) as an example.

for most of the dynamics. The effect of the room modes thus contribute moderately to this motional impedance, which is added to the blocked impedance of the voice coil [25]. Nevertheless, it can be seen that the acoustic load exerted by the room on the loudspeaker is able to modulate the electrical input impedance. Even if the effects are small, changes in perimeter, occupancy and temperature can be detected, paving the way for even easier indoor event monitoring.

3.5. On temperature change

Assuming air as an ideal gas, the speed of sound in dry air at normal conditions can be approximated with an error less than 0.2% between −30°C and + 30°C using the following linearization [26]:

$$c_\theta = 331.4 + 0.59 \cdot \theta \tag{21}$$

where θ is the temperature in degrees Celsius. By combining Eq. (21) with Eq. (5), it can be shown that a temperature shift $\Delta\theta$ implies that the natural frequency of all modes is shifted by:

$$\Delta f_n = \frac{0.59}{2}\sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2}\,\Delta\theta \tag{22}$$

As indicated in Eq. (22), the natural frequency of a given room mode therefore varies proportionally with the ambient temperature. The higher the natural frequency, the greater the frequency shift due to a change in temperature. Fig. 11 shows contour curves where Δf_n has constant values for a given mode between 0 and 700 Hz and a temperature increase from 0 to + 5°C.

As an example, Fig. 12 shows the measured specific acoustic impedance of the loudspeaker diaphragm at two different temperatures when the room is in the same configuration (door closed, two windows open, no chair, no additional sound absorbing materials and no one inside). It is thus confirmed that a change in room temperature indeed leads to a frequency shift of room modes which increases with frequency. This experiment shows that the impedance of the loudspeaker membrane is sensitive to small modal variations in the sound field and that the interior events studied have specific identifiable characteristics on the modal response.

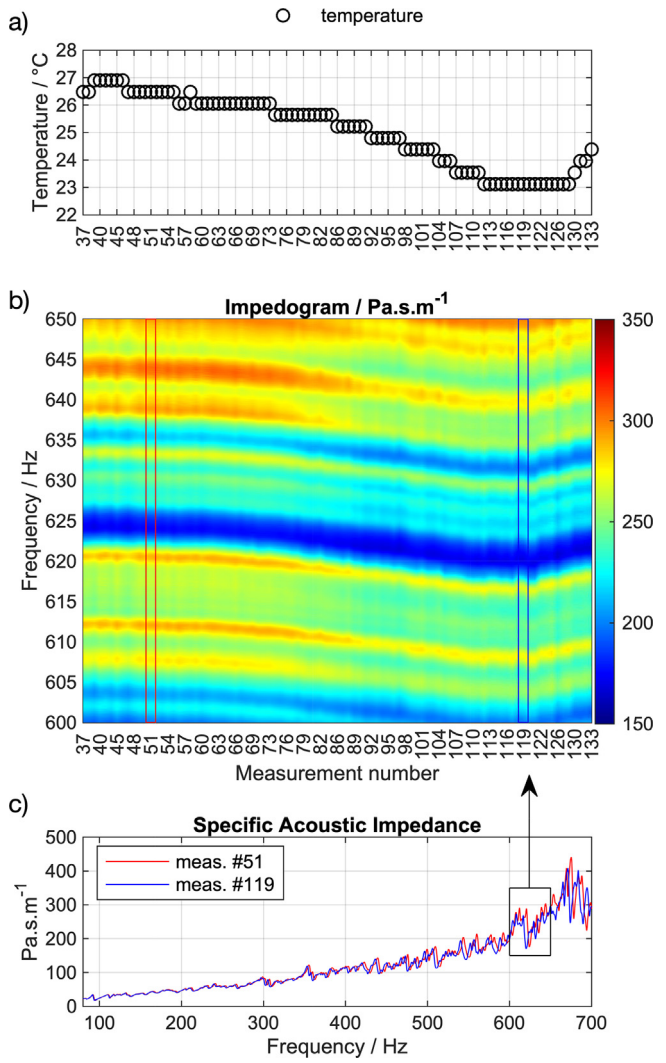

Fig. 7. Evolution of room modes over time due to a change in temperature; a) gives the evolution over 25 h of the room temperature, b) shows the corresponding impedogram between 600 Hz and 650 Hz resulting from the concatenation of 96 segments of 5 min recorded every 10 min and c) gives the SAI frequency response functions for measurements 51 (26.5°C) and 119 (23.2°C) as an example.

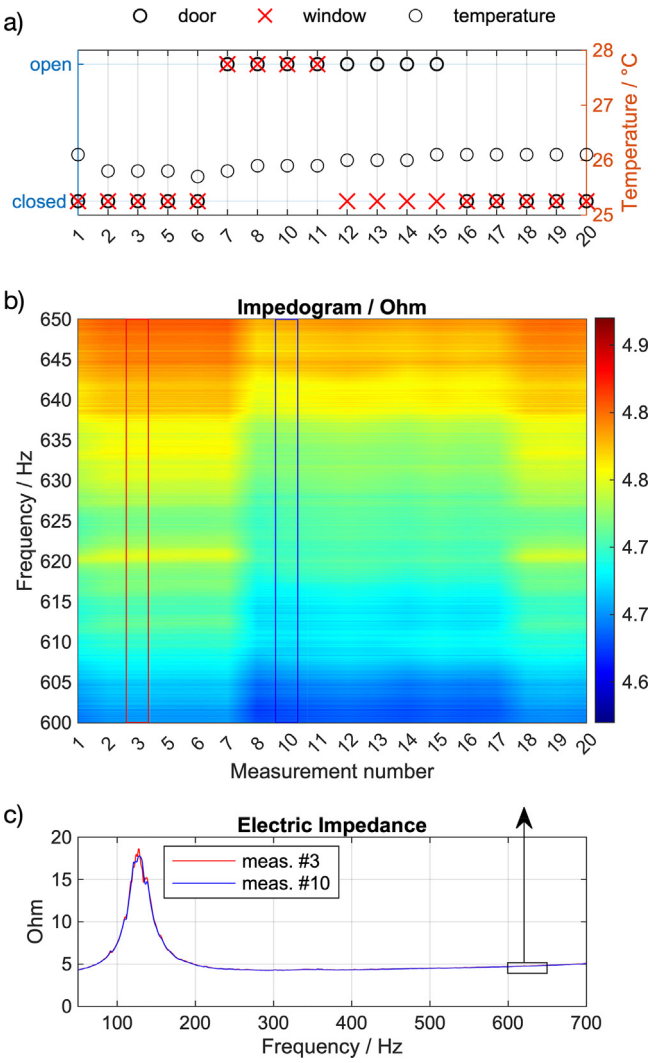

Fig. 8. Same as Fig. 5 but in the electrical domain.

4. Event classification

4.1. Problem statement

In the following, we present the classification process implemented using an algorithm trained to recognize recurrent patterns specific to the events of interest on the impedograms discussed above. The intended application is to predict the classes of measured data that differ from the baseline scenario – identified by the algorithm described in Section 3.2 for example. The general algorithm framework for such a classification problem should be flexible enough to work in real rooms and to not be limited to a specific number of scenario. Indeed, some users may be interested in detecting large changes, such as those that can be observed between an open and a closed door, while others may be interested in fine classification of very acoustically similar situations, such as predicting the number of people in a room.

Regardless of the type of signal or application, a classification problem traditionally involves two steps: feature extraction and pattern recognition. In our context, knowing which feature(s) in the impedance frequency response function is relevant to dissociate two distinct scenarios generally requires very precise room and loudspeaker models, and some expertise in engineering and simulation. In practice, obtaining the mechanical and acoustic properties of the boundaries of the room and the furniture is a cumbersome process, making it almost impossible to identify in advance what really matters in dissociating two scenarios. The reason we chose a solution based on a deep neural network is that they can extract such features automatically when trained on a huge amount of data [27].

In recent years, deep neural networks have been successfully applied to solve similar engineering problems in other application areas such as ECG signal classification [28] or speech emotion classification [29]. The common strategy is to use two-dimensional classification models even if the raw data is one-dimensional signals. This makes it possible to fall back within algorithms framework originally implemented for image classification and whose efficiency is no longer questioned, see [30] for example. One such very famous classifier is the convolutional neural network (CNN), introduced in 1989 by LeCun et al. for handwritten zip codes recognition [31].

In audio or biomedical processing, a common method of converting 1D signals to 2D images is to build (mel-) spectrograms, as in [28]. We applied a similar approach to feed a CNN classifier with impedograms described in Section 3.3. The preliminary study

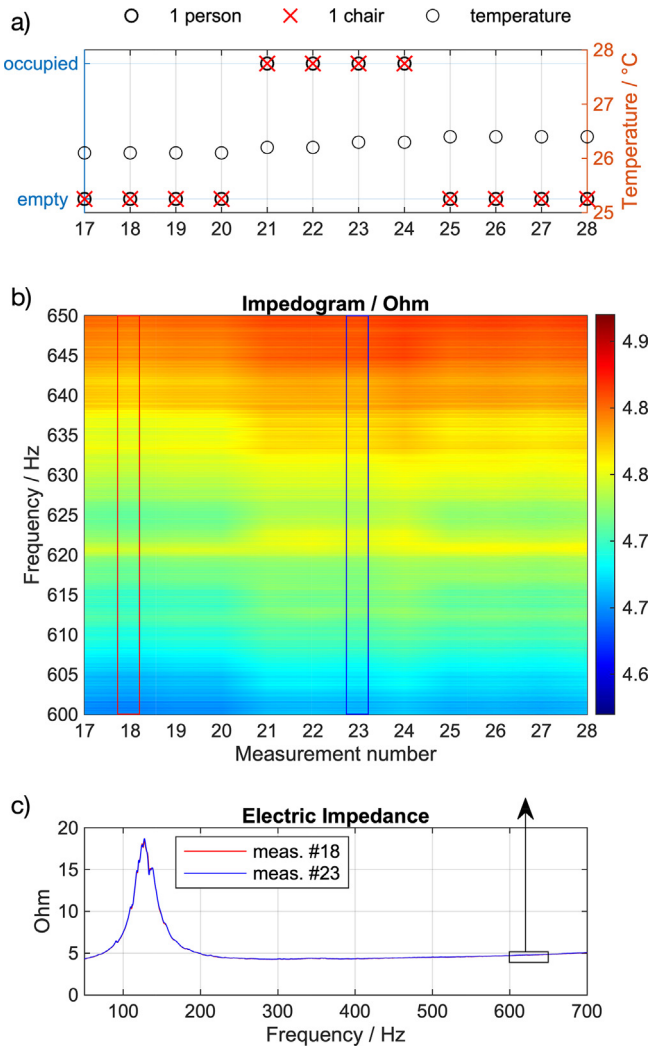

Fig. 9. Same as Fig. 6 but in the electrical domain.

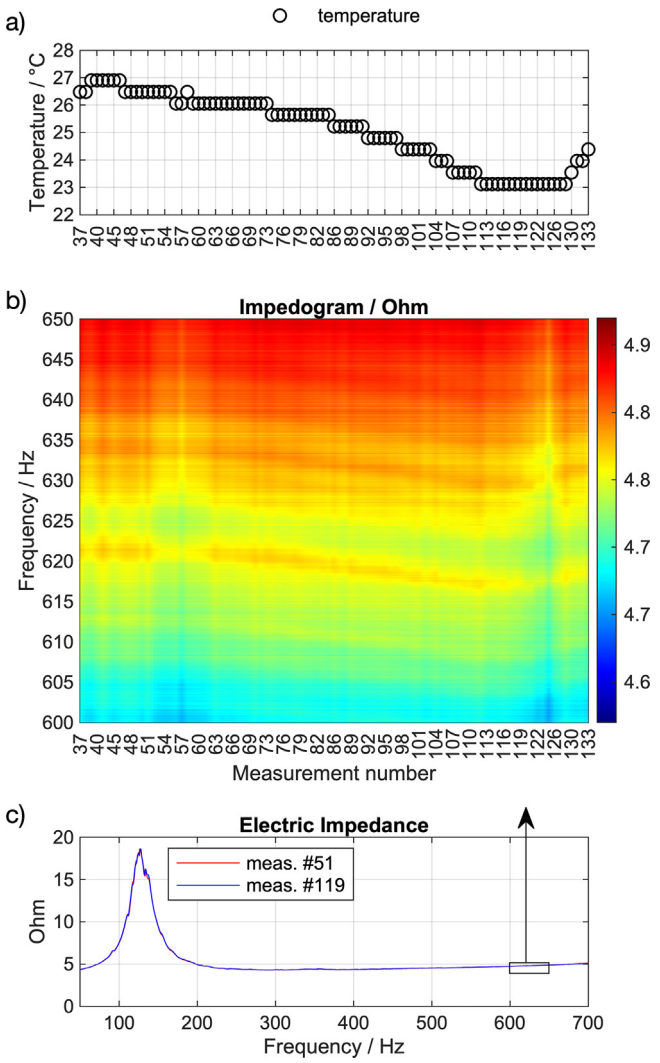

Fig. 10. Same as Fig. 7 but in the electrical domain.

conducted to validate the feasibility of the principle is discussed below.

4.2. Preliminary classification experiment

An experimental setup similar to that described in Section 3.1 was installed in a rectangular, rigid-walled room of 5.97 m $\times$2.885 m $\times$2.615 m. In addition to the acoustic pressure and diaphragm normal velocity, the current and voltage at the input terminals of the loudspeaker were also measured. The four signals were acquired simultaneously at a sampling rate of 4400 Hz using specially developed electronics to drive the loudspeaker by voltage-controlled current source. The underlying idea is to explore which type of impedogram (acoustic p/vor electrical u/i) is best suited for the classification of events using a loudspeaker. As can be seen in Fig. 13, the room is unfurnished and has only one door (from where the photo is taken).

During this experiment, four classes were considered:

- class 1: door closed, room empty (baseline configuration)
- class 2: door closed, with 8 absorbing panels in the room
- class 3: door closed, someone walking in the room
- class 4: door open, room empty

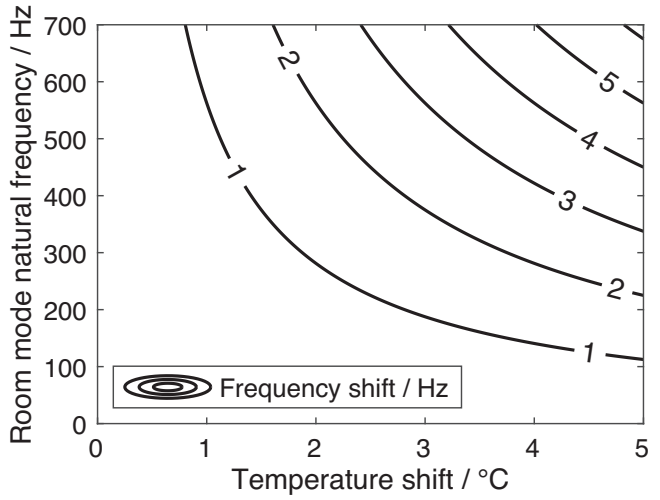

Fig. 11. Contour plot of expected frequency shift of room modes with a change in temperature; y-axis represents the natural frequencies that room modes can take, x-axis represents the temperature change and the contour curves indicate the iso-response values of the corresponding frequency shift Δf_n.

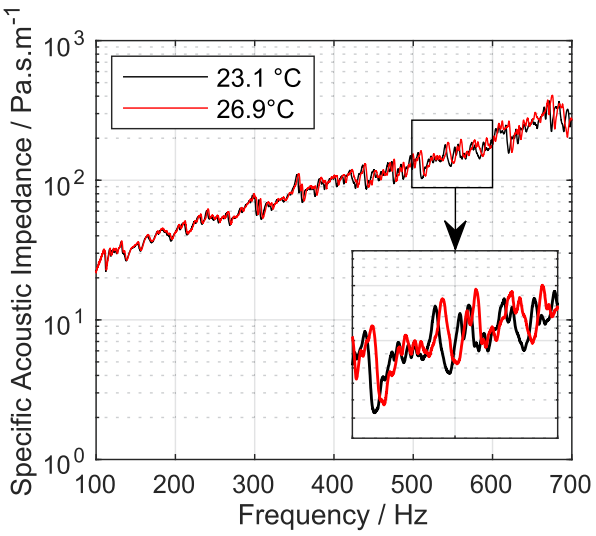

Fig. 12. Specific acoustic impedance measured under the same room configuration for two different temperature values.

In class 2, we used eight panels of Dinaphon B 810 open cell melamine foam with a total effective volume of $1.2 \times 0.6 \times 0.05 \times 8 = 0.288$ m^3. Approximately 45 min of signals were recorded for each class. Data were divided into a training dataset (about 40 min) and a testing dataset (about 5 min). In terms of samples, the training dataset is constituted of 319, 295, 297, and 345 frames for class 1, 2, 3 and 4, respectively, and the testing dataset is constituted of 53, 51, 57, and 56 frames for class 1, 2, 3 and 4, respectively.

The size of each frame was equal to $F \cdot N = 1000 \cdot 10$, where F denotes the frequency bins and N is the number of impedance measurements. The studied bandwidth was set between 70.5 Hz and 205 Hz, with a frequency resolution of 0.13 Hz. Note that for convenience, each impedogram is resized to a square image of 100×100 pixels before feeding the CNN. For the sake of transparency we provide in Table 3 the architecture of the CNN used in this work. Our model is inspired by one of those described in the documentation of the Deep Learning toolbox of MATLAB (The Math Works, Inc). The optimisation of the CNN architecture is beyond the scope of this study, so we have not tried to compare it to others. However, we believe that the classification results obtained with another architecture would have equally proven the feasibility of our approach.

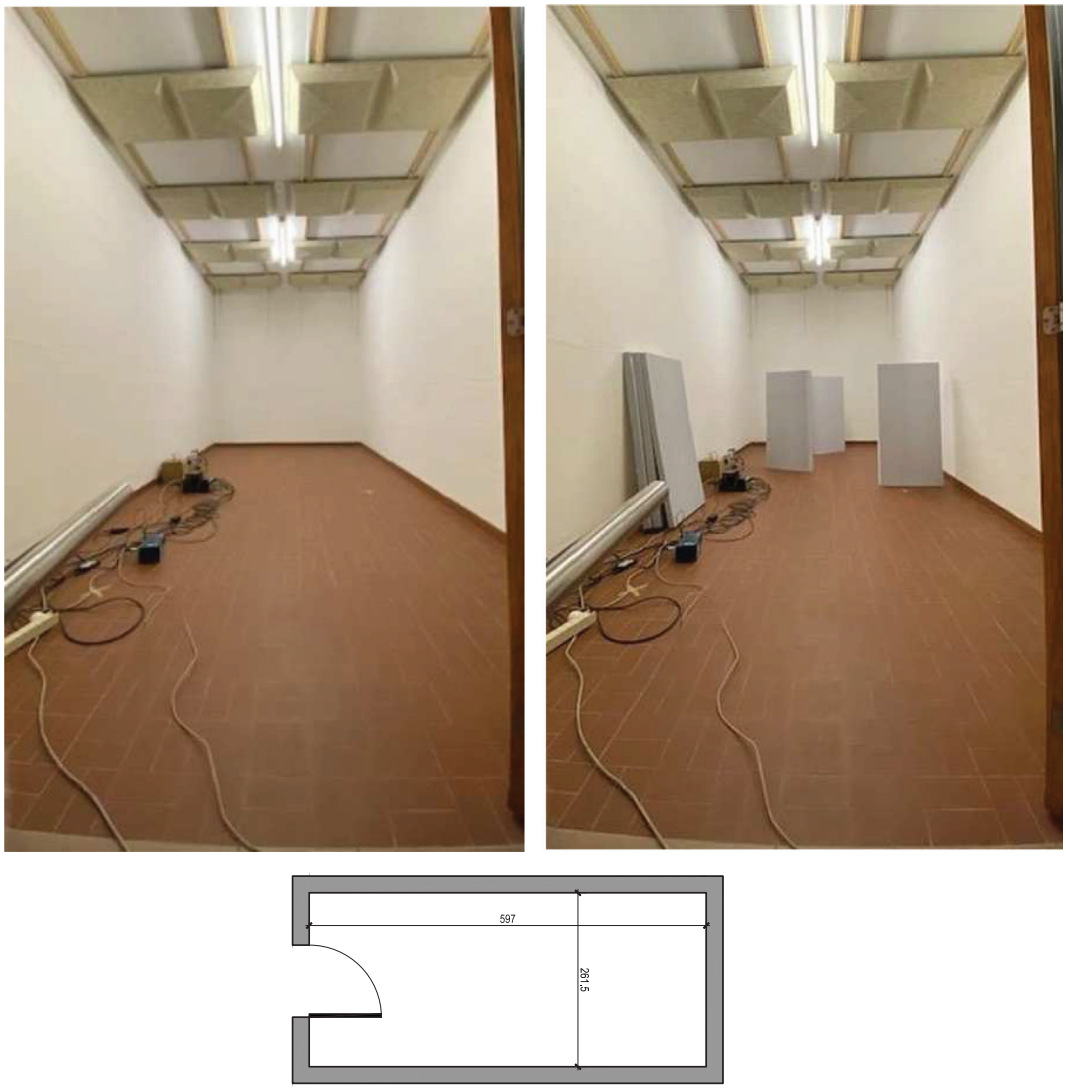

Fig. 13. Test room used for the preliminary classification experiment; a) empty as in class 1 and 4, b) with 8 absorbing panels as in class 2, and c) map of the test room.

Table 3
Architecture of the used 2D-CNN model.

Layer	Name	Type	Activations
1	`imageinput` $100 \times 100 \times 1$ images	Image Input	$100 \times 100 \times 1$
2	`conv_1` 12 3×3 convolutions with stride [1 1]	Convolution	$100 \times 100 \times 12$
3	`batchnorm_1` Batch normalization	Batch Normalization	$100 \times 100 \times 12$
4	`relu_1` ReLu	ReLu	$100 \times 100 \times 12$
5	`maxpool_1` 3×3 max pooling with stride [3 3]	Max Pooling	$34 \times 34 \times 12$
6	`conv_2` 24 3×3 convolutions with stride [1 1]	Convolution	$34 \times 34 \times 24$
7	`batchnorm_2` Batch normalization	Batch Normalization	$34 \times 34 \times 24$
8	`relu_2` ReLu	ReLu	$34 \times 34 \times 24$
9	`maxpool_2` 3×3 max pooling with stride [3 3]	Max Pooling	$12 \times 12 \times 24$
10	`conv_3` 48 3×3 convolutions with stride [1 1]	Convolution	$12 \times 12 \times 48$
11	`batchnorm_3` Batch normalization	Batch Normalization	$12 \times 12 \times 48$
12	`relu_3` ReLu	ReLu	$12 \times 12 \times 48$
13	`maxpool_3` 12×12 max pooling with stride [1 1]	Max Pooling	$1 \times 1 \times 48$
14	`dropout` 20% dropout	Dropout	$1 \times 1 \times 48$
15	`fc` 4 fully connected layer	Fully Connected	$1 \times 1 \times 4$
16	`softmax` softmax	Softmax	$1 \times 1 \times 4$
17	`classoutput` crossentropyex	Classification output	$1 \times 1 \times 4$

4.3. Results

In the following, we evaluate the performance of the classifier in terms of percentage accuracy of the results. Let's M be the number of classes, and let's i and j two integers such that $i \in [1, M]$ and $j \in [1, M], j \neq i$. An impedogram belonging to class i is called a positive sample and an impedogram belonging to class j is called a negative sample. The model accuracy for the i^{th} class can be defined as:

$$Accuracy_i(\%) = \frac{TP_i + TN_j}{TP_i + TN_j + FP_{i,j} + FN_{i,j}} \times 100 \tag{23}$$

where TP_i stands for true positive, meaning the number of positive samples (actually belonging to class i) that are predicted to be positive by the model; TN_j stands for true negative, meaning the number of all negative samples (actually belonging to any class $j, \forall\, j \in [1, N], j \neq i$) that are predicted to be negative by the model; $FP_{i,j}$ stands for false positive, meaning the number of samples that are predicted to be positive by the model but are actually negative; and $FN_{i,j}$ stands for false negative, meaning the number of samples that are predicted to be negative by the model but are actually positive.

Accuracy results after feeding the CNN with the testing dataset are provided in Table 4. Two types of impedograms are considered and compared: those obtained from electrical quantities (current and voltage) at the loudspeaker terminals and those derived from acoustic quantities (output pressure and velocity at the loudspeaker membrane). The accuracy results obtained with the electrical impedograms are slightly lower than those obtained with the acoustic impedograms, but still sufficient to classify the different events of interest. In addition, this sensorless approach has the advantage of being more compact and less expensive. This preliminary result shows how promising impedance-based event classification can be in an enclosed space such as a room.

5. Discussion

Most meeting and office rooms are rectangular in shape with dimensions such that the first room modes appear below a few hundred hertz. This is also the frequency range in which most electrodynamic speakers are used. This matching results in a strong dynamic coupling between the loudspeaker and the modal response of the room. In this way, room modes can be detected at the loudspeaker diaphragm, which forms the acoustic interface. The proposed method therefore uses the impedance of the loudspeaker diaphragm to estimate various changes within the room.

The modal expansion method described in Section 2 is used to show a simple case of coupling between a rectangular room with rigid walls and a square piston. Through this analytical approach, we show how room modes affect the frequency response of the loudspeaker. This highlights that the diaphragm impedance is a relevant quantity to observe changes in a room or at its boundaries. By monitoring the loudspeaker impedance, a range of indoor events can therefore be classified from the sound field. For example, the closed-form solution given in Eq. (18) illustrates how a temperature variation influences room modes. For more realistic room shapes, however, a FEM model may be needed to predict how an indoor event may affect the modal response of the room. In addition, objects such as furniture can also be added to the model to study how room modes are affected in terms of damping and frequency shift. Note that a time-domain parallel wave-based acoustic FEM and a frequency-domain FEM can be used to handle more practical room acoustic designs [32–34]. They allow the modelling of acoustic diffusers or sound absorbers with locally reacting frequency-dependent impedance boundary conditions.

The indoor event detection and classification approach proposed in this article requires a fine frequency resolution to analyse small variations in the modal response. The frequency bins we used experimentally were typically 0.1 Hz resolution or less. In a lightly-damped room with solid walls, the study of a frequency range between 20 Hz and 200 Hz is generally sufficient to extracts

Table 4
Accuracy results in percent of the classifier.

impedance type	door closed + room empty	door closed + abs. panels	door closed + someone	door open
electrical	97.2	100	97.2	100
acoustical	99.5	100	99.5	100

the relevant features from the modal response. The quality factor of the room modes is quite high and modal density is rather low which makes them easily identifiable. In a real room made of lightweight drywall for example, the damping is higher and the normal modes tend to overlap, making them more difficult to distinguish. For the indoor event classes studied in this research, we found that they were nevertheless detectable. This is explained by the fact that each event affects several modes simultaneously and that the algorithm considers the entire frequency response of the room. If this concept were to be applied in an enclosed space of smaller or larger dimensions, the frequency range of interest would need to be adjusted accordingly.

This study demonstrates that it is possible to give new features to a conventional loudspeaker already employed in buildings for audio broadcasting. This widely used and relatively inexpensive transducer can indeed be employ to detect changes in the indoor sound field. An AI algorithm trained to recognise recurring patterns in labelled data sets then extracts features to classify events. One advantage of this technology is no sensitive data is recorded in contrast to security camera. The fact that the device is easy to set up in a room, if not already there, makes it an interesting option in situations where security and privacy must be taken into account. This can be useful in helping companies or building managers to understand how premises and workspaces are used. By detecting the presence of people or the movement of furniture in real time, it is thus possible to use space more efficiently and save money, or to know which offices or meeting rooms are available in near real time.

This study also showed that the interior changes could be detected from the electrical input impedance of the loudspeaker. As can be seen in Table 4, the room modes transfer well to the input impedance of the transducer but the accuracy results are poorer than those obtained using the microphone and laser Doppler vibrometer. This is partly because the response of the electrodynamic transducer dominates over that of the room modes. We are currently working on training a long-term memory neural network (LSTM) [35] to estimate the acoustic impedance directly from the electrical impedance at the loudspeaker input terminals. With this sensorless approach, only easily measurable electrical quantities are processed directly, making the device even more compact and less expensive.

After this preliminary study, many questions remain to be explored and pave the way to future research. It should be examined whether a loudspeaker can recognise events in a different room from the one used for the training data set. It would also be interesting to study the influence of the position of the loudspeaker in the room on the overall performance. This would help to determine whether the classification is improved in an area of high modal density (typically in a corner of the room) or rather in an area of lower modal density. From a loudspeaker perspective, it would be desirable to determine the key electromechanical parameters that should be favoured to increase the accuracy of the classification. Another practical aspect would be to explore whether the speaker can monitor indoor events while playing music at the same time. Regarding feature extraction, the minimum recording time per class in the learning phase of the algorithm would also need to be studied.

6. Conclusion

In this paper, we present an acoustic-based approach to indoor event detection and supervised classification using a loudspeaker and an algorithm trained to recognise recurring patterns from the modal response of the room. The principle relies on an impedance measurement at the diaphragm of the loudspeaker which defines the acoustic interface with the room and is sensitive to room modes. A CNN algorithm is then used to interpret the variations in frequency response induced by changes in or at the boundaries of the room. Experimental evidence in real rooms is provided, showing how changes in occupancy, temperature variations or the opening of doors and windows affect the room modes, and therefore the impedance of the diaphragm. A 25-h experiment conducted in a meeting room shows that changes in acoustic impedance seen through the loudspeaker can be used to identify the cause. This privacy-friendly technology solution is expected to eventually allow monitoring of different categories of indoor events without recording sensitive data, in addition to playing music, notifications or alerts. Future work will focus on estimating acoustic impedance without external sensors to make it more compact, cheaper and simpler, and in improving the automatic event classification method.

CRediT authorship contribution statement

Patrick Marmaroli: Conceptualization, Methodology, Formal analysis, Writing - original draft, Investigation, Visualization. **Mark Allado:** Visualization, Software. **Romain Boulandet:** Formal analysis, Conceptualization, Writing - original draft, Data curation, Funding acquisition, Project administration, Supervision, Writing - review & editing.

Data availability

Data will be made available on request.

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.

Acknowledgments

This work was supported by the interdisciplinary program of the Engineering and Architecture domain of the HES-SO under grant AGP No. 102538. The authors thank Axel Baxarias for developing the automatic click program which made it possible to perform unattended measurements during the test campaign. The authors also thank François Bugnon and Clément Geslin for their help during the experiments, as well as Dr. Samir Ouelha and Dr. Jason Cambridge for their technical advice and pre-review.

References

[1] Z. Chen, C. Jiang, L. Xie, Building occupancy estimation and detection: a review, 2018.
[2] Dong B, Prakash V, Feng F, O'Neill Z. A review of smart building sensing system for better indoor environment control. Energy Build 2019;199:29–46.
[3] Hong SG, Kim N, Kim WW. Reduction of false alarm signals for pir sensor in realistic outdoor surveillance. ETRI J 2013;35.
[4] Mena AR, Ceballos HG, Alvarado-Uribe J. Measuring indoor occupancy through environmental sensors: A systematic review on sensor deployment. Sensors 2022;22.
[5] Q. Huang, Z. Ge, C. Lu, Occupancy estimation in smart buildings using audio-processing techniques, CoRR abs/1602.08507 (2016).
[6] Kim J, Min K, Jung M, Chi S. Occupant behavior monitoring and emergency event detection in single-person households using deep learning-based sound recognition. Build Environ 2020;181:107092.
[7] Kim K, Wang S, Ryu H, Lee SQ. Acoustic-based position estimation of an object and a person using active localization and sound field analysis. Appl Sci 2020;10.
[8] Kuttruff H. Room acoustics. (6th ed.),. CRC Press; 2016.
[9] Bücklein R. The audibility of frequency response irregularities. J Audio Eng Soc 1981;29.

[10] Salava T. Acoustic load and transfer functions in rooms at low frequencies. J Audio Eng Soc 1988;36:763–75.
[11] Olive SE, Schuck PL, Ryan JG, Sally SL, Bonneville ME. The detection thresholds of resonances at low frequencies. J Audio Eng Soc 1997;45:116–28.
[12] Cox TJ, D'Antonio P, Avis MR. Room sizing and optimization at low frequencies. J Audio Eng Soc 2004;52:640–51.
[13] Craggs A, Buma CJ. The effect of an absorbent lining on the natural frequencies and modal damping factors of a small room. Appl Acoust 1989;28:229–39.
[14] Papadopoulos CI. Redistribution of the low frequency acoustic modes of a room: a finite element-based optimization method. Appl Acoust 2001;62:1267–85.
[15] Xiaotian Z, Zhemin Z, Jianchun C. Using optimized surface modifications to improve low frequency response in a room. Appl Acoust 2004;65:841–60.
[16] Cox TJ, D'Antonio P. Acoustic Absorbers and Diffusers: Theory, Design and Application. (2nd ed.),. Taylor and Francis; 2009.
[17] Gibbs MSBM. The effect of construction material, contents and room geometry on the sound field in dwellings at low frequencies. Appl Acoust 2004;65:31–44.
[18] Conti SG, Roux P, Demer DA, de Rosny J. Measurements of the scattering and absorption cross section of the human body. Appl Phys Lett 2004;84.
[19] Morse PM, Ingard KU. Theoretical Acoustics. McGraw-Hill; 1968.
[20] Magalotti R, Cardinal V. Building fem low frequency room model through modal decay time measurements. In: Proc. of 23rd International Congress on Acoustics. Aachen, Germany: Universitätsbibliothek der RWTH Aachen; 2019.
[21] Schroeder MR. Frequency-correlation functions of frequency responses in rooms. J Acoust Soc Am 1962;34.
[22] N. Kanev, Active damping of room eigenmodes, in: Proc. of e-congress Inter-Noise, Seoul, South Korea.
[23] L. Wyse, Audio spectrogram representations for processing with convolutional neural networks, in: Proc. of the First International Workshop on Deep Learning and Music joint with IJCNN, Anchorage, USA, pp. 37–41.
[24] Boashash B, Ouelha S. Designing high-resolution time-frequency and time-scale distributions for the analysis and classification of non-stationary signals: a tutorial review with a comparison of features performance. Digital Signal Process 2018;77:120–52. Digital Signal Processing & SoftwareX - Joint Special Issue on Reproducible Research in Signal Processing.
[25] R. Boulandet, Sensorless measurement of the acoustic impedance of a loudspeaker, in: Proc. of 23rd International Congress on Acoustics, Aachen, Germany.
[26] M. Rossi, Audio, Presses Polytechniques Universitaires Romandes, 2007.
[27] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
[28] Huang J, Chen B, Yao B, He W. Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network. IEEE Access 2019;7:92871–80.
[29] A. Badshah, J. Ahmad, N. Rahim, S. Baik, Speech emotion recognition from spectrograms with deep convolutional neural network, pp. 1–5.
[30] Rawat W, Wang Z. Deep convolutional neural networks for image classification: A comprehensive review. Neural Comput 2017;29:1–98.
[31] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition. Neural Comput 1989;1:541–51.
[32] Yoshida T, Okuzono T, Sakagami K. A parallel dissipation-free and dispersion-optimized explicit time-domain fem for large-scale room acoustics simulation. Buildings 2022;12.
[33] Okuzono T, Sakagami K. A frequency domain finite element solver for acoustic simulations of 3d rooms with microperforated panel absorbers. Appl Acoust 2018;129:1–12.
[34] Hoshi K, Hanyu T, Okuzono T, Sakagami K, Yairi M, Harada S, Takahashi S, Ueda Y. Implementation experiment of a honeycomb-backed mpp sound absorber in a meeting room. Appl Acoust 2020;157:107000.
[35] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735–80.