SCIENCE & TECHNOLOGY

Blood CO Status Classification Using UV-VIS Spectroscopy and PSO-optimized 1D-CNN Model

Audrey Huong[1]*, Kim Gaik Tay[1], Kok Beng Gan[2] and Xavier Ngu[3]

[1]*Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Johor, Malaysia*
[2]*Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor Malaysia*
[3]*Institute of Integrated Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Johor, Malaysia*

ABSTRACT

Rapid and effective blood carbon monoxide (CO) assessment is of great importance, especially in estimating CO-related morbidity and instituting effective preventive measures. The conventional detection methods using CO breath analysis lack sensitivity, while collecting biological fluid samples for CO level measurement is prone to external contamination and expensive for frequent use. This study proposes a one-dimensional convolutional neural network (1D-CNN) consisting of three stacked biconvolutional layers for binary classification of blood CO status using the diffuse reflectance spectroscopy technique. Iterative particle swarm optimization (PSO) has efficiently found the best network parameters to learn important features from the reflectance spectroscopy data. The findings showed good testing accuracy, specificity, and precision of 92.9%, 90%, and 89.7%, respectively, and a high sensitivity of 96.3% in determining abnormal blood CO among smokers using the proposed CNN network. Comparisons with eight existing machine learning and deep learning models revealed the proposed method's effectiveness in classifying blood CO status while reducing computing time by 8–13 folds. The findings of this work provide new insights that are valuable for researchers in neural network design automation, healthcare management, and skin-related research, specifically for application in nondestructive evaluation and clinical decision-making.

Keywords: Carbon monoxide, machine learning, network design, optimization, spectroscopy

E-mail addresses:
audrey@uthm.edu.my (Audrey Huong)
tay@uthm.edu.my (Kim Gaik Tay)
kbgan@ukm.edu.my (Kok Beng Gan)
xavier@uthm.edu.my (Xavier Ngu)
* Corresponding author

INTRODUCTION

Carbon monoxide (CO) is an odorless and colorless gas formed by the incomplete combustion of carbon-based compounds. Once air containing CO gas is inhaled and absorbed into the bloodstream, it binds preferentially to heme irons, forming carboxyhemoglobin (COHb). This condition leads to decreased oxygen-carrying capacity, depriving other cells of the oxygen supply essential for metabolism. It is the key factor for increased oxidative stress and inflammation in cells, causing defects and damage to highly oxygen-dependent organs, including the heart and brain (Carrola et al., 2023).

Smoking tobacco smoke remains a major source of indoor air pollution (Raju et al., 2020). Cigarette smoke produces a CO concentration of 3–6% (Turino, 1981) in the air. This concentration is 2–3 times higher in bidi and cigar smoke (Datta et al., 2022). Indoor combustion sources, current smoking status, and passive smoking are the major contributory factors to the increased CO levels in the blood. Other high-risk groups include workers with significant occupational risk, such as firefighters or charcoal production workers (Idowu et al., 2023). These adverse effects of CO in blood, known as CO toxicity, are factors that increase morbidity and mortality in immunosuppressed patients.

Smoking status and blood CO level have been mentioned in the literature (Hoeng et al., 2019; Nemmar et al., 2022) as reliable biomarkers for assessing outcome measures and predicting future disease development in patients. High-tech medical imaging modalities, such as magnetic resonance imaging (MRI), computed tomography (CT) scan, and retinal imaging, can provide a comprehensive analysis of CO concentration in the body (Vaghefi et al., 2019). They are not widely used for research due to their high cost, unavailability, and lengthy examination time. Biochemical analysis, blood gas analyzer, and ribonucleic acid (RNA) tubes are alternative methods to confirm smoking status, but these devices are invasive. Direct and noninvasive approaches, such as breath analysis and salivary or urinary cotinine tests (Ramani et al., 2023), are more suitable for outpatient management to assess smoking habits. However, the accuracy of expired gas analysis can be compromised in patients suffering from severe airflow obstruction (Papin et al., 2023), while food consumption pattern (i.e., presence of thiocyanate ions in various vegetables, fruits, and milk), lifestyle, and environmental and physical conditions are among the other factors that have a direct effect on the performance of breath analysis (Sharma et al., 2023 & Shreya et al., 2023). The shorter half-life of expired CO also makes it less reliable than cotinine measurement in urine and saliva (Usmani et al., 2008). Nonetheless, external contamination of body fluid samples is common during sample processing, leading to unreliable results. Besides, these test strips can be expensive if large specimens are needed.

Spectroscopy is a simple and noninvasive technique to obtain optical information about a medium across a broad wavelength range. Knowing the properties of light absorption and scattering of the medium enables backward prediction of a substance's concentration

and physical structure based on the measured light signals without destroying its integrity. This nondestructive feature, cost-effectiveness, and flexibility allow its use in endless possibilities. CO oximeter is a state-of-the-art photoelectric device capable of determining the percentage of CO saturation in blood using the optical spectroscopy method and with a fixed calibration curve. A previous study showed an elevation in the percent COHb, or blood CO level measured using a CO oximeter from 4%–6% in the control group to 8%–25% in smoking or acute CO poisoning patients (Bol et al., 2018; Onodera et al., 2016). However, the system can be unreliable for measurements outside the calibration curve defined using the absorbance ratio of isosbestic points 532 nm and 558 nm (Papin et al., 2023).

Related Works

Deep learning is an emerging technique in artificial intelligence that allows rapid and fully automated extraction of important features from a dataset for the classification task. Among them, CNN is the most widely used method for vision-related tasks. A standard CNN consists of convolutional, pooling, and fully connected layers to learn a hierarchical representation of features from raw data. Pretrained CNNs, such as AlexNet, Residual Network (ResNet), GoogleNet, and Visual Geometry Group (VGGNet), are widely used to learn a new complex task by fine-tuning the weights of neurons in the model.

Most models contain alternating convolutional and pooling layers to abstract features layer by layer. While these pre-trained models are generally recommended and proven promising for two-dimensional (2D) image processing tasks to 3D object recognitions, 1D-CNN for signal processing problems has received little attention compared to its extended counterparts. Some works in the past demonstrated using 2D convolutional models for the 1D signal classification task with considerable success (Ahmad et al., 2021), whereas others designed their 1D model from scratch. The process can be tedious and laborious as it involves tuning network hyperparameters, such as learnable weights and biases, as well as the layers' arrangement and size.

The research community used manual, automatic, or hybrid methods in designing the CNN model for a specific classification task (Yang et al., 2021). While manual methods include brute force or a grid search approach, an automatic search for the best model design can be carried out more aggressively using optimization techniques. Various types of optimization algorithms are available for this purpose, including randomized adaptive search procedures, such as the Markov decision process and Monte Carlo (MC) search, particle swarm optimization (PSO), genetic algorithm (GA), Bayesian optimization (BO), pattern search method, and gray wolf optimization (GWO). Each technique adopts different strategies to minimize a given function or cost measure and could be preferred over others in specific applications. Among these, BO is the most popular technique for optimizing custom-built CNN. This method works by incrementally building a probabilistic model

during optimization. Previous works (Kolar et al., 2021; Ling et al., 2022; Shi et al., 2021) reported considerably good classification accuracy using the 1D-CNN architecture and training hyperparameters optimized using BO for fault diagnosis and safety analysis.

However, this technique does not scale well to higher dimensions and has high computational complexity. Pattern search is a direction-based method taking trial steps in each direction for each parameter within the range specified. The PSO works by iteratively searching in a region defined by the best success of the neighborhood particles. Meanwhile, GWO is based on the mathematical models of the social behaviors of gray wolves guided by iterative prey encircling and hunting processes. The GA method works by randomly generating different populations of generations and mutations in its search. These optimization techniques have also been applied in a wide area of decision-support systems. Layek et al. (2021) demonstrated using GA and PSO to determine the elasticity constant for application in thyroid cancer detection. In Goel et al. (2021), comparisons have been made to evaluate the performance of these techniques in optimizing the feature extraction and classification components. GA, pattern search, PSO, and GWO were shown to produce comparable results in their optimized models for automatic diagnosis of coronavirus disease (COVID-19).

Another work by Tan et al. (2019) showed superiority in the performance of the model optimized using the PSO method compared to its competing models, namely GA and BO, in the security of unmanned aerial vehicle (UAV) networks. This finding is further supported by a review by Korani and Mouhoub (2021), who summarized the different advantages of PSO compared to other population-based algorithms as an optimization method.

Unlike the machine learning method, the neural network offers the advantages of superior generalization and representation abilities at the cost of higher computational complexity and expensive hardware requirements (Asgharzadeh et al., 2023). Machine learning classifiers, for example, support vector machines (SVM), k-nearest neighbors (KNN), Naïve-Bayes (NB), and regression methods, require less computational and implementation effort, and they are also highly reputable for solving binary problems. Most machine learning algorithms are simple and easy to use, requiring fewer parameters to tune before training. However, their performance is comparatively inferior, especially when it involves nonlinear solutions or complex decision boundaries. To further enhance the performance of classification, some work in the past preferred using a binary classifier optimized using the optimization methods mentioned above for prediction. Examples of important works in this area are using a hybrid SVM-PSO system to improve landslide susceptibility prediction (Zhao & Zhao, 2021). Thote et al. (2017) used a KNN system optimized using GA to discriminate the different faults in transformer fault diagnosis.

Studies of optical reflectance for the classification of CO in vivo have not received much attention among the research community in the field, which is mainly due to the high sensitivity of the measurement to factors other than COHb, such as diversity in human skin colors, thickness, and time-varying physiological states (Huong et al., 2014). Several

recent studies using the optical method, specifically spectroscopy, in smoking and nicotine-related investigations include a comparison study between a five-wavelength transmissive spectrophotometer and the standard radiography blood analyzer for detecting CO poisoning (Lyon et al., 2022). The study reported high consistency in the performance of the optical spectroscopy technique with the control device. Yi et al. (2023) demonstrated fluorescence spectroscopy for in vitro detection of CO in living tissues. Li et al. (2022) designed a novel optical-acoustic technology using an acoustic resonator to detect acoustic wave signals generated by the human exhaled gas molecules (i.e., CO) absorbing laser energy.

These past studies used calibration curves and fitting models to predict the required value. Additional research using the spectroscopy method combined with the deep learning model for examining tissue COHb would be interesting. This article aims to present an optimized 1D-CNN model for a quick and objective classification of a person's blood CO in skin based on diffuse reflectance spectroscopy. This study chose the PSO method for the optimization problem because of its excellent global convergence speed and effectiveness. The contributions of this work are threefold:

1. Advancement of research in developing the CNN model for blood CO classification by creating a new ultraviolet-visible (UV-VIS) spectroscopy dataset of nonsmoking and smoking subjects.
2. An iterative PSO-based optimization strategy in automatically designing a 1D-CNN model to learn abstract features from the spectroscopy data.
3. An effective detection solution for blood CO status with minimal computing resources.

MATERIALS AND METHODS

Participants and Selection Criteria

There is no public dataset for blood CO detection research using optical spectroscopy. This research built the first smoking-related CO dataset from experiments conducted at the Universiti Tun Hussein Onn Malaysia (UTHM) laboratory between 2015 and 2019. This study was approved by the Ethical Committee of UTHM (approval no. 100-9/39). Figure 1 shows the consort diagram of participant recruitment and analysis. One hundred seven subjects consisting of university students and individuals from the public were randomly approached and invited to participate in the studies; 41 declined. Most cite a lack of time and inconvenient schedules and locations as the main reason for their refusal. The remaining 66 volunteers were required to answer questions about their age and smoking habits, disclose their living and occupational status, health information, and medical history, and repeat prescriptions before the screening. The daily air quality index where the participants lived and worked showed acceptable readings with an average value of 25 ppm during the study period.
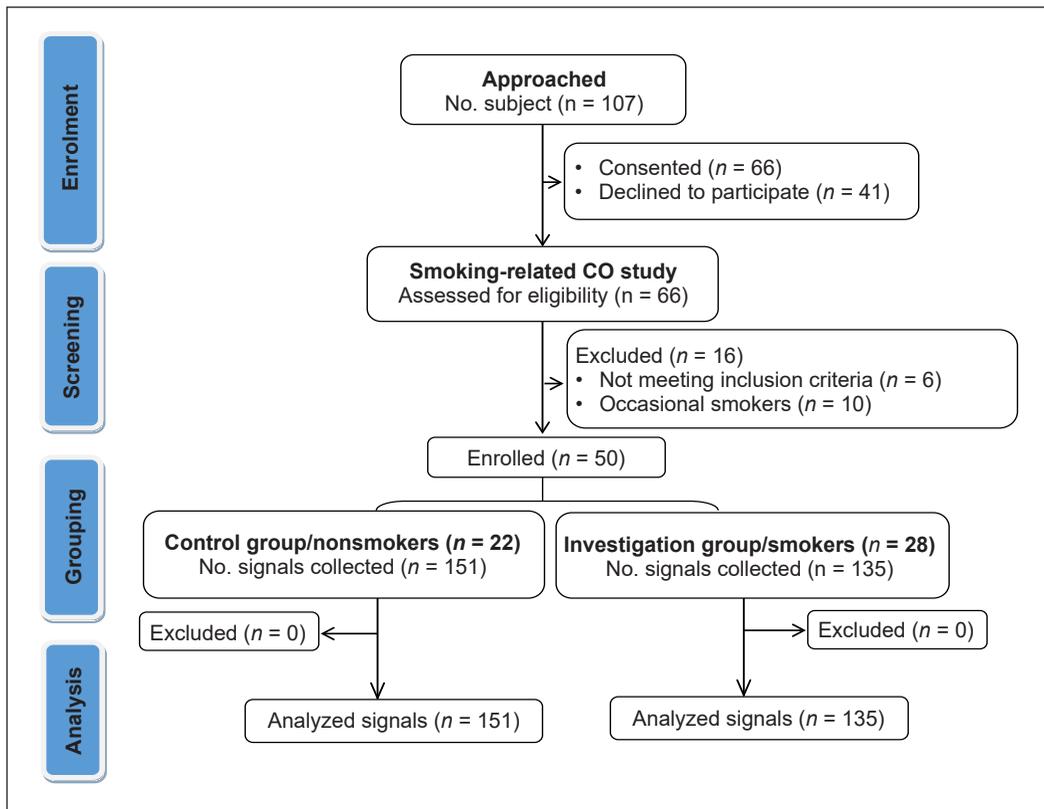
Figure 1. Consort diagram for participant flow

The volunteers were assessed for their eligibility based on inclusion and exclusion criteria. The exclusion criteria included subjects with cardiovascular, endocrine, or respiratory system diseases. Of these respondents, sixteen were excluded because of their inconsistent smoking habits (n = 3), nonsmokers with the possibility of CO exposure in self-report assessment (n = 7), or participants with medical records (n = 6) that could affect the results. Fifty people met the eligibility criteria and were enrolled in the study. These participants are never-smokers with no household cigarette smoke or occupational or indoor combustion exposure (n = 22) or regular smokers (n = 28) aged 18 or older who had smoked at least one cigarette per day for at least two years during the time data was taken. They were aged between 20 and 62 years. They self-declared good general health and were not on long-term medication.

Experimental System and Data Collection

Information about the sampling and data collection processes and experimental setup used for the measurement of diffuse light reflectance from the skin of control (i.e., nonsmoking) and investigation (smoking) groups were provided in the original studies (Huong & Ngu, 2014, 2015). However, for completeness of this paper, they are summarized here.

Figure 2 shows the schematic experimental setup for skin reflectance spectroscopy. During these experiments, a white light-emitting diode (LED) (Model no. SMD 5730 from Aira Technologies), placed at 80 mm from the skin and 20° from the normal axis, illuminated the selected skin area. The detection system consisted of a UV-VIS spectrometer (model no. USB4000 Ocean Optics, Florida) connected to a bifurcated fiber optic bundle. The fiber tip was placed 35° from the incidence plane and 20–40 mm above the skin. Light reflected from the targeted surface was collected into the optical fiber before being diffracted by the grating (1200 groves/mm) in Figure 2, which spreads the light spectrum (in the wavelength range of 178–898 nm) on a charge-coupled detector (CCD)-array inside the spectrometer. The intensity reading was recorded by a laptop installed with SpectraSuite software for further processing and analysis.

On average, between 4 to 8 spectra were recorded from the index finger of the recruits during the resting state. The studies were carried out in a well-ventilated laboratory with an ambient temperature of 24 ± 2ºC. A total of 135 and 151 reflectance spectroscopy signals (N) were recorded from 50 subjects: 28 smoking and 22 nonsmoking individuals, respectively. Data analysis using a paired sample t-test performed with SPSS (version 2, IBM Inc.) with a confidence level of 95% showed statistical significance ($\rho=0.0171$) for inter- and intra-subject variability. Hence, they are treated individually as independent signals for the classification task. These sample sizes are sufficient for clinical research studies, as agree with the minimum sample size requirement of 137 (for smoking) and 101 (nonsmoking group) calculated from the power size formula of Das et al. (2016).

The signals of these volunteers were grouped according to their reported smoking status. These class labels are confirmed with the mean percent carboxyhemoglobin (COHb) values of 3%–7% and 8%–16%, respectively, for smokers and nonsmokers using the MC
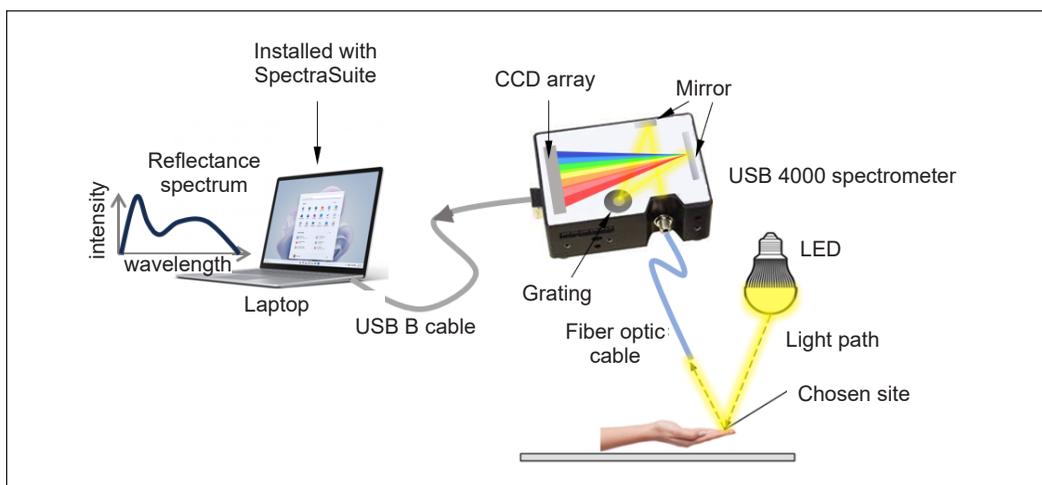
Figure 2. The UV-VIS spectroscopy measurement setup. Diagram also shows the inside of the spectrometer for detection of the reflectance spectrum

approximation (Huong & Ngu, 2014, 2015). Figure 3 shows the participants' stratification based on their smoking status and verification of the class labels.

Based on the original data points of 1×3648 from the spectrum in the range of 178 nm to 898 nm (i.e., with spectral resolution of 0.2 nm), a fixed window of length 1×1415 was applied to truncate signals in the wavelength range of 420–680 nm in Figure 3, consistent with the output wavelength of the employed LED for the data mining process. Most importantly, there is a considerable variation and strong characteristic absorption peaks for hemoglobin variants in this range, which is suitable for detecting the oxygenation state of hemoglobin using the spectroscopy technique (Nitzan et al., 2020). These data handling processes, from data collection to signal processing and database archiving, are summarized in Figure 3. This study does not consider the feature selection method to preserve all information for investigation. Attempts at enriching the dataset using augmented data by adding and swapping uncertainties (i.e., noises) and the sliding window approach proposed by Ullah et al. (2018) have not yielded improved results. Thus, the original dataset was used and randomly split for training, validation, and testing purposes using a ratio of 60%/20%/20% with a random seed value of 1 for results reproducibility.
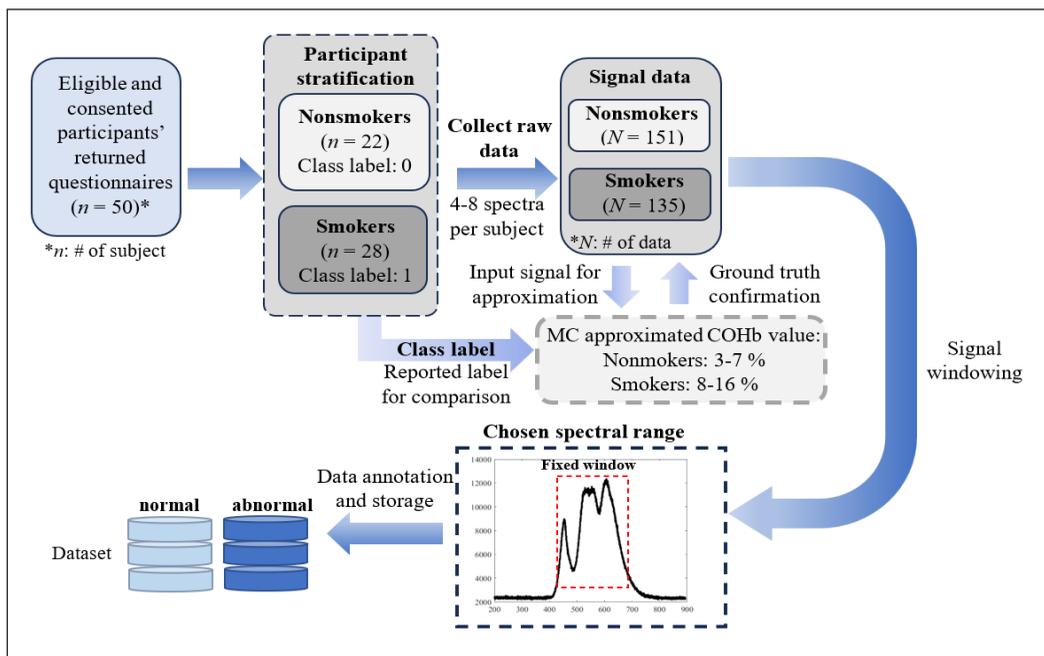
Figure 3. Reflectance spectroscopy data handling process

Network Architecture Design

The design of an end-to-end 1D CNN used in this research is inspired mainly by the structure and organization of the existing pre-trained networks, consisting of convolutional filters

(CONV), rectified linear unit (ReLu), max pool (POOL), fully connected (FC) layer, and a classifier. Specifically, stacked bi-convolutional and one-pooling layers (i.e., $CONV_1$-$ReLu_1$-$CONV_2$-$ReLu_2$-POOL) are adopted as the basic structure of the CNN due to its universality and straightforward design. This architecture of alternating convolutional and pooling layers, like most pre-trained networks such as VGGNet and AlexNet, produces minimal features from the signal. Such an arrangement is referred to as biCONV. Next, two FC layers and a Softmax are followed for probability prediction of the signals for classification. The last layer is the classification output containing abnormal and normal classes.

Several architectures have been attempted in which an additional biCONV block is progressively stacked on top of the preceding layers in each model version to deepen the network. Their classification training and validation accuracies were recorded and compared. The results revealed that the network architecture containing three stacked biCONVs in Figure 4 produced considerably superior classification accuracies (i.e., increased by 5%–10%) than its shallower counterparts and negligible performance differences compared to its deeper counterparts. Therefore, the architecture in Figure 4 is chosen as the final design, and the results from this model are presented in the remainder of this paper.
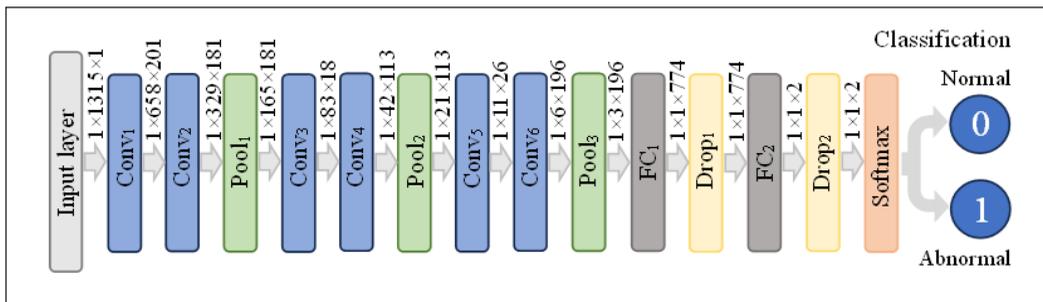
Figure 4. End-to-end 1D-CNN model for blood CO status classification

Network Parameter Optimization

The most common method for determining network hyperparameters (i.e., CNN kernel size, stride and filter numbers, and training parameters) is a grid search or brute-force method, whose decision is guided based on the classification accuracies of the trained model. The procedure is challenging, expensive, time-consuming, and has a limited coverage area. In this study, parameters of hidden nodes (i.e., input weights and biases) and some of the most important training parameters (i.e., optimizer type, mini-batch size, and initial learning rate) are fine-tuned during the optimization.

This study addressed this challenge by proposing a 16-degree-of-freedom problem to generalize the model for the task. The objective function to be minimized is as Equation 1:

$$f\left(T_{acc}, V_{acc}, t_s\right) = \left((100 - T_{acc}) + (100 - V_{acc})\right) \times 1e^3 + \frac{t_s}{1e^3} \tag{1}$$

T_{acc} and V_{acc} are the training and validation accuracies, respectively, while t_s is the run time for each search iteration. The iteration terminates prematurely if the validation accuracy fails to improve for ten iterations during the search process. The upper and lower range of variables that directly affect the objective function result is shown in Table 1. Since the neural network was trained from scratch, the learning weights and biases of the neurons were determined based on the kernel size or filter size of the convolutional layers identified through the optimization method.

Figure 5 shows an overview of network development and optimization processes for smoking status classification. After splitting the dataset (collected and annotated in Figure 3) into training, validation, and testing cohorts, the training and validation sets were used to build the classification model and tune the relevant parameters. The PSO method was used for search purposes. The PSO works by iteratively updating the velocity and position of the particles of swarm size 200 by moving each particle based on the best-known previous positions within the boundary limits of the search space specified in Table 1.

Each candidate solution created based on the random number generator was subjected to the search for a maximum of 500 optimization steps, making the search process stochastic, and it was based on the locations of the previous best solutions. This paper implemented the same padding strategy for each layer. It used a pooling size of (1×4) to overcome the mismatch in the feature map dimensions between the network layers from the chosen parameters. A stride length of (1×2) was used in all convolutional and pooling layers to ensure consistency in the proposed approach. This program attempted to search for the best solution (i.e., minimize the solution) to the objective function given in Equation 1. The search step for all integer variables was set to 1, and the step change of the initial learning rate was given by $1e^{-8}$. Meanwhile, in the efforts to cover the entire search space, this optimization process was repeated 300 times before determining the best solution.

A fixed epoch number was used because the pre-experiment results showed considerably lower significance in its effects on network performance than the mini-batch size and initial learning rate. It could also reduce the complexity of the search and, hence,

Table 1
Parameters range in the fine-tuning process

Description	Parameter	Lower limit	Upper limit
Training hyper-parameter	Optimizer, *Opt*	{1→3; 1: *Adam*, 2: *Sgdm*, 3: *RMSProp*}	
	Mini batch size, β	8	1024
	Initial learning rate, χ	$1e^{-6}$	$1e^{-2}$
Network learnable	Kernel size[*]	1×2	1×10
	Filter no.[*]	1×2	1×256
	FC_1	100	1000

[*]For all convolutional blocks ($Conv_1$-$Conv_6$)
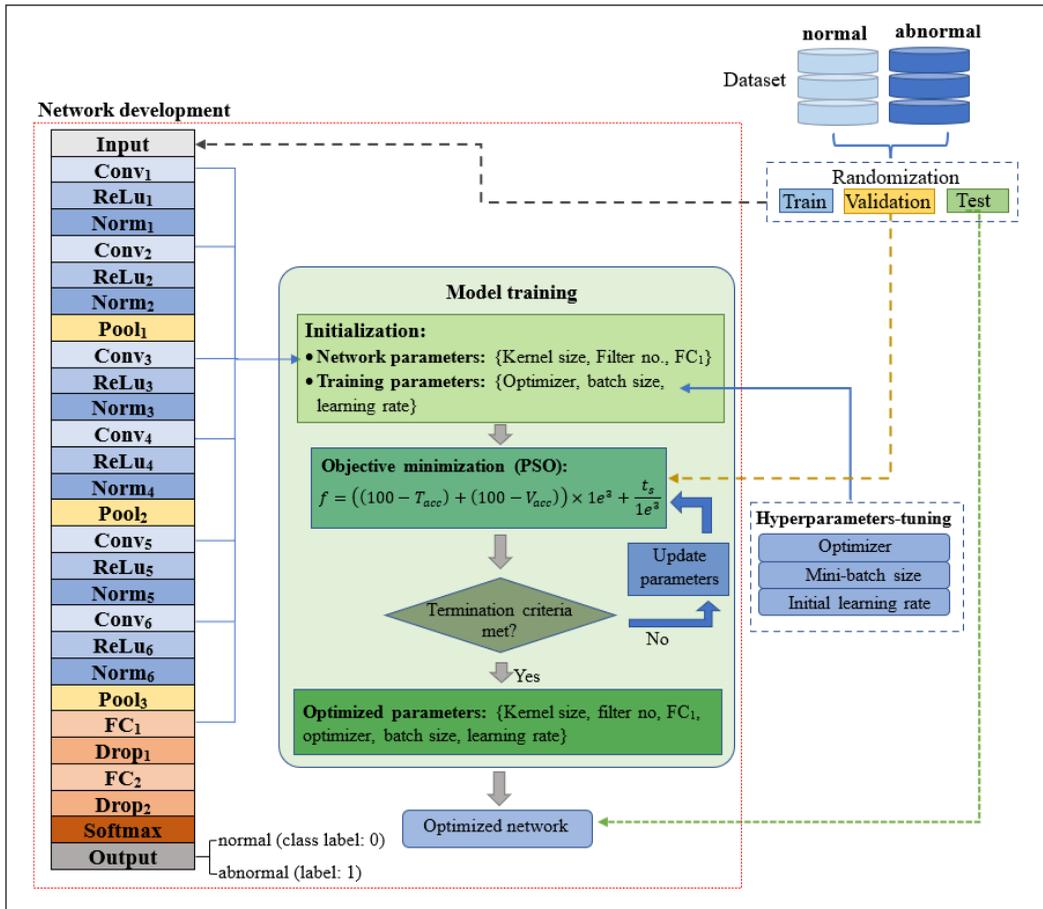
Figure 5. An overview of network development and optimization for smoking status classification

computational resources. Several efforts have been made to improve the generalization of the model, including:

1. incorporating a dropout regularization of factor 0.2 after biCONV and FC layers,
2. using a high epoch number of 500 to increase training time,
3. considering the spectral range that matches the LED output spectrum (i.e., signal windowing shown in Figure 3) to minimize unnecessary features in the analysis and
4. an early stopping function monitors validation accuracy during the training process.

Evaluation Metrics and Comparisons

Different performance metrics have been used to evaluate the model's performance for the classification task. Testing data that do not have a role in the training and validation phase was used to test the ability of the trained model to identify the person's blood CO status. The considered evaluation metrics are accuracy (*Acc*), sensitivity (*Sens*), specificity (*Spec*),

and precision (*Prec*) in Equations 2 to 5, which can be calculated using a confusion matrix. The confusion matrix can be used to detect classification errors with four components: (1) a true negative (*TN*) is when a signal is correctly classified as a normal blood CO (class 0), (2) a true positive (*TP*) correctly classifies the abnormal blood CO case (class 1), (3) a false positive (*FP*) is when a normal blood CO signal is incorrectly classified as abnormal, and (4) a false negative (*FN*) is an error in which an abnormal case is misclassified as normal.

Accuracy (*Acc*) is the proportion of accurate results among the total number of cases examined.

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \text{ \%} \tag{2}$$

Sensitivity (*Sens*) or recall rate is the probability of a positive abnormal CO result among all abnormal data.

$$Sens = \frac{TP}{TP+FN} \times 100 \text{ \%} \tag{3}$$

Specificity (*Spec*) is defined as the ability of a test to exclude normal data.

$$Spec = \frac{TN}{TN+FP} \times 100 \text{ \%} \tag{4}$$

Precision (*Prec*) is the percentage of a correctly classified abnormal case to total positive abnormal results.

$$Prec = \frac{TP}{TP+FP} \times 100 \text{ \%} \tag{5}$$

This research chose some popular classifiers, namely SVM, KNN, Decision Tree (DT), logical regression (LR), Naïve-Bayes (NB), and some of the popular pre-trained networks (i.e., AlexNet, GoogleNet, and ResNet-18) as the benchmark models for comparison purposes, as summarized in Figure 6. These models were trained and tested using the same spectroscopy dataset for a fair comparison of the results. The threshold *k*-value in the KNN classifier varied manually from 1 to 10; the best result chosen based on the classification accuracy is presented in this work. In the
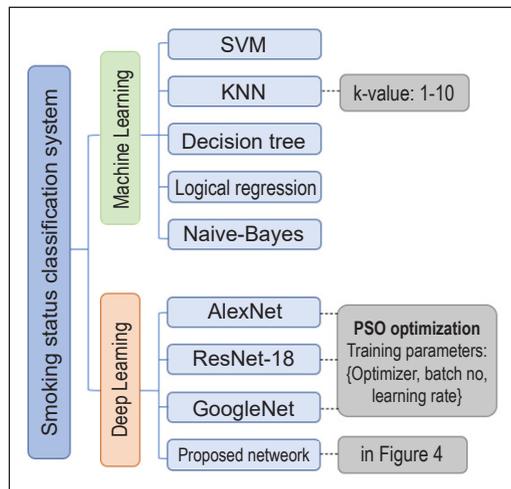
Figure 6. The benchmark methods for comparison with the proposed network

case of pretrained 2D models, the *getframe* function screenshotted each signal's plot before saving it as an image. The image was resized according to the input size allowable for each network. The hyperparameter values, namely optimizer solver, mini-batch size, and initial learning rate, were chosen for training deep models using the PSO method following the minimization process in Figure 5. The weight of all layers was updated during the training process for the spectroscopy dataset.

RESULTS

This study developed the CNN model from scratch, and the model that yielded the lowest objective function value given in Equation 1 is the best for classifying the blood CO status. This research approached this task by systematically stacking additional biCONVs consecutively to deepen the network. The final design of the 28-layer network shown in Figure 4 gives 841,509 learned parameters to extract features from the spectroscopic data for the smoking classification task. This paper presents the results for 100 sets of training parameters (i.e., optimizer (Opt), mini-batch size (β), and initial learning rate (χ)) that produced the lowest f value, plotted against its T_{acc} and V_{acc} in Figure 7. The best training hyperparameter set chosen from this plot is $Opt = 2$ (i.e., $Sgdm$), $\beta = 555$, and $\chi = 0.0035$. Since early termination was adopted to speed up the search process, each search iteration time is recorded as ranging between $t_s = 3$–50 seconds executed on an NVIDIA Tesla K80 GPU with 12 GB of memory and 13 streaming multiprocessors. Based on the identified hyperparameters, the model training was repeated three times to ensure the reproducibility of the results. Figure 8 shows the best confusion matrix of the proposed 1D CNN system evaluated using the testing set.

The same training and testing procedure was performed on the machine mentioned above for the eight competing models in Figure 6. The mean and standard deviation of
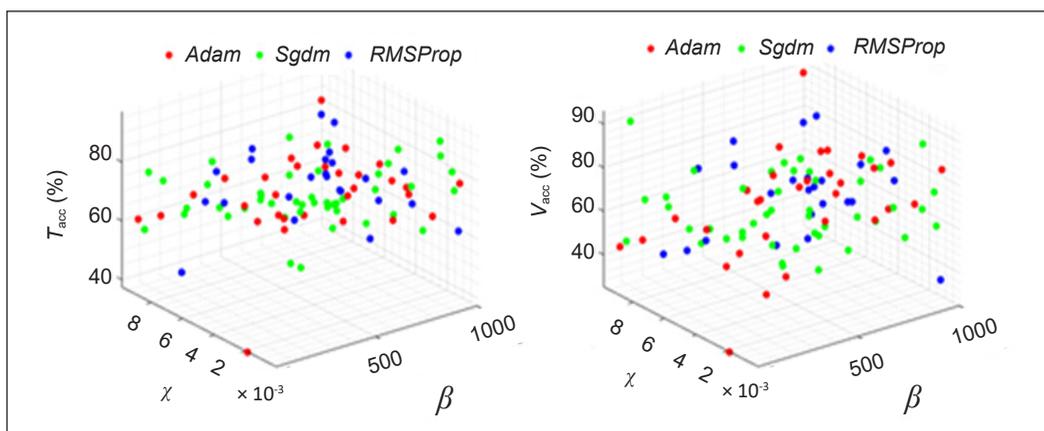
Figure 7. The percent training (T_{acc}) and validation accuracies (V_{acc}) against 100 best training hyperparameter sets (χ: initial learning rate, β: mini-batch size, filled color: optimizer)

the evaluation metrics are shown in Table 2. Also included in this table is the total elapsed time taken to solve the optimal training hyperparameters for the deep learning models (i.e., AlexNet, GoogleNet, ResNet, and the proposed 1D CNN) using the PSO optimization process are shown in Figure 5. The experiments using the KNN method with different k values in Figure 6 produced the best accuracy result with $k = 7$. Thus, it is presented in the Table 2.
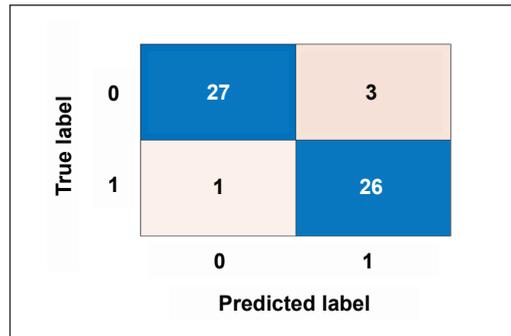
Figure 8. The best confusion matrix of the proposed 1D-CNN model tested on the testing data. Class label 0: normal and 1: abnormal blood CO status

Table 2

A comparison of blood CO status classification performance between the proposed model and state-of-the-arts

Model	Mean ± SD of evaluated metrics (in %)*				Training time (in s)*	Iterative optimization time (in s)†
	Acc	*Sens*	*Spec*	*Prec*		
SVM	67	100	29.6	61.2	0.52 ± 0.15	-
KNN	82.5	93	70.4	78	0.56 ± 0.36	-
DT	65	67	63	67	0.57 ± 0.39	-
LR	92.9	93	92.5	93	0.76 ± 0.11	-
NB	72	86.7	55.6	68.4	1.4 ± 0.07	-
AlexNet	85.9 ± 1.8	93.3 ± 3.3	77.8 ± 6.4	82.5 ± 3.6	54 ± 37	26,542
GoogleNet	81.3 ± 7	88.9 ± 11	72.8 ± 11	78.8 ± 7	210.5 ± 61	38,982
ResNet-18	88.3 ± 5	92.2 ± 5	83.9 ± 5.6	86.5 ± 4.8	166.4 ± 91	22,495
Proposed model	91.8 ± 1	93.3 ± 3.3	90.1 ± 5.6	91.5 ± 4.3	15.8 ± 8.2	4,750

Note. ACC: accuracy, *Sens*: Sensitivity, *Spec*: Specificity, *Prec*: Precision; *Mean and standard deviation (SD) results from three runs for optimized model training and testing; †Total time taken in iterative search of optimum training hyperparameters (deep learning models)

DISCUSSION

This study demonstrated an optimization algorithm to find the appropriate weights and biases, extract important features from the spectroscopy data to classify blood CO and determine the best set of hyperparameters for the improved learning process of the CNN model. This method is used primarily to minimize the problem of classification accuracies and training time. The comparison of different methods in Table 2 shows differences in classification performance depending on the methods. The proposed model achieved a good classification accuracy of 92.9%, confirming the feasibility of the PSO-optimized method for the efficient development of the custom-made 1D-CNN model.

Although classifiers, such as DT, NB, and LR, have a long history of being used for the two-class classification task, the presented results show that the CNN method outperformed

most of these models. The testing accuracy ranged between 85% and 93% using the CNN type of networks compared to 65%–92% using binary classifiers. The spectroscopy data has no visually distinctive features separating normal and abnormal groups. Thus, this may explain why a reasonable decision boundary may not be achieved, rendering most of these linear classifiers fail to differentiate blood CO classes. This problem is not found in the case of pre-trained networks. Even though they were developed for image recognition, they worked acceptably well for the 1D problem, suggesting their sufficiency in extracting important features with deep layers of neurons.

The SVM performs best in many applications through the optimal arrangement of hyperplanes for modeling the decision boundary. However, it fails to detect abnormal blood CO status in smokers with a classification accuracy of 66%. This classifier is overfitting to the nonsmoking class, resulting in perfect precision and specificity rates (i.e., FP rate of 0) but poor sensitivity of 29%. Meanwhile, the KNN method that finds the closest objects in training data to the unknown input produced a reasonably good performance with a classification accuracy of 82% using the k-value of 7, implying that the distance-based strategy works better on small datasets like ours.

An interesting attribute of the proposed model is that it has the least learnable parameters compared to the pre-trained models. Table 2 shows a significant decrease in its optimization and training times by up to 8 and 13 folds, respectively, compared to the deeper models without sacrificing generalization ability. Although some normal data have been misclassified as abnormal, as shown in Figure 8, giving a slightly inferior precision compared to ResNet-18 in Table 2, the designed model eliminates the need for complex architecture and is least likely to suffer from dimensionality problems. The robustness of this model is evident with its superior CO status detection sensitivity of 96.2% in smokers compared to the traditional models in Table 2, suggesting that this shallow and small model can recognize global coarse information with the first two stacked convolutional layers, while final stack layer is sufficient to learn the refining information further.

In addition to optimizing network architecture, this study enhanced the model's generalization through iterative tuning of training hyperparameter sets. Figure 7 shows that Sgdm has a high probability (~50%) of being chosen as the solver, followed by Adam and RMSProp. This outcome is not surprising given the high computing cost of the RMSProp (~1.12 folds longer training time than the others). The experiments found that the most feasible range of initial learning rate is between 0.0043–0.0058 and 522–599 for mini-batch size using the employed spectroscopy dataset. An increase in the initial learning rate produced poor learning efficiency, where the exploding gradients incidence (i.e., training accuracies returned as invalid values) was observed. Meanwhile, the overall lower mean training accuracy of 77.2% compared to 80.98% in validation accuracies implies the possibility of underfitting the network. This issue is understandably due to the small training set. In the earlier experiments, attempts have been made to enlarge the dataset by

incorporating different noise signals into the original data, but no observable changes in the classification performance were noted. This finding can be interpreted as the ability of the proposed model to recognize noise signals (i.e., high frequencies) as unimportant features.

On that note, lower frequency information may be identified as useful features, so extraction of better features through further modifications of the network architecture may be necessary to enhance the performance of the existing model. The final design in Figure 4 is determined after ten tests have been carried out, ranging from one biCONV to five-stack biCONVs, with fully connected layers' dropout rates of 0.2 and 0.5. While the incorporation of dropout 0.2 generally produced better results in terms of classification accuracies than 0.5, the training and validation accuracies remained relatively constant beyond networks comprising three-stack bi-CONVs.

This study is one of the pioneer and preliminary attempts to explore whether the proposed framework and technology can detect blood CO status. Those who chose to participate were self-selected. It limits the generalizability of the study; thus, research needs to be explored in larger samples in the future. Since this is not a strictly controlled experiment, the diet and lifestyles of the participants have not been regulated. Therefore, carefully designed and strict experiment conditions for future investigation are recommended. Studies exploring network capacity for predicting health risks for categories of abnormal CO are also a task for further research.

Classification of blood CO status using light reflectance spectroscopy demonstrated in this study is the first in the field to use optimized deep learning methods. This strategy is an efficient, rapid, and cost-effective approach for the task as compared to conventional breath analysis, which is unreliable in patients with severe airflow obstruction. Unlike the study by Lyon et al. (2023) and Yi et al. (2023) that required daily calibration verification and blood samples to be taken for CO classification using transmissive and fluorescence spectrophotometry systems, the proposed approach performed in-situ and real-time classification based on skin reflectance signals. This system is potentially useful as a tool for pre-hospital triage and evaluation of treatment in CO-poisoning patients.

Future research should include larger data from human subjects of different physical characteristics and health problems to enhance the robustness of the proposed system in assessing a person's CO exposure status and predicting health risks. In terms of classifier design, this study acknowledges the importance of further efforts to improve the developed model for the classification task. A more rigorous design procedure is expected to be further explored to allow more efficient extraction of high-level features for the task.

CONCLUSION

This paper adopts a PSO-optimized 1D-CNN model for blood CO classification using the optical spectroscopy signal. This work is the first to apply a deep learning approach

combined with spectroscopy to the blood CO abnormality classification task. The proposed optimization framework offers a time and effort-efficient approach to customizing the model. This method achieved comparable performance with most existing pre-trained deeper models while outperforming state-of-the-art binary classifiers. Potential directions for further research include deeper investigations and richer evidence about skin CO differences in a larger population with different physical parameters and clinical presentation to improve study generalization. In addition, there is still room for improving the network structure to extract the target features more accurately.

ACKNOWLEDGEMENTS

REFERENCES

Ahmad, Z., Tabassum, A., Guan, L., & Khan, N. M. (2021). ECG heartbeat classification using multimodal fusion. In *IEEE Access* (Vol. 9, pp. 100615-100626). IEEE Publishing https://doi.org/10.1109/ACCESS.2021.3097614

Asgharzadeh, H., Ghaffari, A., Masdari, M., & Gharehchopogh, F. S. (2023). Anomaly-based intrusion detection system in the Internet of Things using a convolutional neural network and multi-objective enhanced Capuchin Search Algorithm. *Journal of Parallel and Distributed Computing*, *175*, 1-21. https://doi.org/10.1016/j.jpdc.2022.12.009

Bol, O., Koyuncu, S., & Gunay, N. (2018). Prevalence of hidden carbon monoxide poisoning in auto service workers: A prospective cohort study. *Journal of Occupational Medicine and Toxicology*, *13*, 13-35. https://doi.org/10.1186/s12995-018-0214-9

Carrola, A., Romão, C. C., & Vieira, H. L. A. (2023). Carboxyhemoglobin (COHb): Unavoidable bystander or protective player? *Antioxidants*, *12*(6), Article 1198. https://doi.org/10.3390/antiox12061198

Das, S., Mitra, K., & Mandal, M. (2016). Sample size calculation: Basic principles. *Indian Journal of Anaesthesia*, *60*(9), 652-656. https://doi.org/10.4103/0019-5049.190621

Datta, R., Singh, S., Joshi, A., & Marwah, V. (2022). Concept of BIDI years: Relevance to the perioperative period. *Lung India*, *39*(4), 337-342. https://doi.org/10.4103/lungindia.lungindia_595_21

Goel, T., Murugan, R., Mirjalili, S., & Chakrabartty, D. K. (2021). OptCoNet: An optimized convolutional neural network for an automatic diagnosis of COVID-19. *Applied Intelligence*, *51*, 1351-1366. https://doi.org/10.1007/s10489-020-01904-z

Hoeng, J., Maeder, S., Vanscheeuwijck, P., & Peitsch, M. C. (2019). Assessing the lung cancer risk reduction potential of candidate modified risk tobacco products. *Internal and Emergency Medicine*, *14*, 821-834. https://doi.org/10.1007/s11739-019-02045-z

Huong, A., & Ngu, X. (2014). Noninvasive diagnosis of carbon monoxide poisoning using Extended Modified Lambert Beer Model. In *2nd International Conference on Electronic Design (ICED)* (pp. 265-269). IEEE Publishing. https://doi.org/10.1109/ICED.2014.7015811.

Huong, A., & Ngu, X. (2015). *In Situ* monitoring of mean blood oxygen saturation using Extended Modified Lambert Beer model. *Biomedical Engineering: Applications, Basis and Communications*, *27*(01), Article 1550004. https://doi.org/10.4015/S1016237215500040

Idowu, O. S., De Azevedo, L. B., Zohoori, F. V., Kanmodi, K., & Pak, T. (2023). Health risks associated with the production and usage of charcoal: A systematic review. *BMJ Open*, *13*(7), Article e065914. https://doi.org/10.1136/bmjopen-2022-065914

Kolar, D., Lisjak, D., Pająk, M., & Gudlin, M. (2021). Intelligent fault diagnosis of rotary machinery by convolutional neural network with automatic hyper-parameters tuning using Bayesian Optimization. *Sensors*, *21*(7), Article 2411. https://doi.org/10.3390/s21072411

Korani, W., & Mouhoub, M. (2021). Review on nature-inspired algorithms. *Operations Research Forum*, *2*, Article 36. https://doi.org/10.1007/s43069-021-00068-x

Layek, K., Basak, B., Samanta, S., Maity, S. P., & Barui, A. (2021). Stiffness prediction on elastography images and neuro-fuzzy based segmentation for thyroid cancer detection. *Applied Optics*, *61*(1), 49-59. https://doi.org/10.1364/ao.445226

Li, B., Feng, C., Wu, H., Jia, S., & Dong, L. (2022). Photoacoustic heterodyne breath sensor for real-time measurement of human exhaled carbon monoxide. *Photoacoustics*, *27*, Article 100388. https://doi.org/10.1016/j.pacs.2022.100388

Ling, Y., Huang, T., Gao, E., Shan, Q., Hei, D., Zhang, X., Shi, C., & Jia, W. (2022). Improving the estimation accuracy of multi-nuclide source term estimation method for severe nuclear accidents using temporal convolutional network optimized by Bayesian optimization and hyperband. *Journal of Environmental Radioactivity*, *242*, Article 106787. https://doi.org/10.1016/j.jenvrad.2021.106787

Lyon, M., Fehlmann, C. A., Augsburger, M., Schaller, T., Zimmermann-Ivol, C., Celi, J., Gartner, B. A., Lorenzon, N., Sarasin, F., & Suppan, L. (2023). Evaluation of a portable blood gas analyzer for prehospital triage in carbon monoxide poisoning: Instrument validation study. *JMIR Formative Research*, *7*, Article e48057. https://doi.org/10.2196/48057

Nemmar, A., Al-Salam, S., Beegam, S., Zaaba, N. E., Elzaki, O., Yasin, J., & Ali, B. H. (2022). Waterpipe smoke-induced hypercoagulability and cardiac injury in mice: Influence of cessation of exposure. *Biomedicine & Pharmacotherapy*, *146*, Article 112493. https://doi.org/10.1016/j.biopha.2021.112493

Nitzan, M., Nitzan, I., & Arieli, Y. (2020). The various oximetric techniques used for the evaluation of blood oxygenation. *Sensors*, *20*(17), Article 4844. https://doi.org/10.3390/s20174844

Onodera, M., Fujino, Y., Kikuchi, S., Sato, M., Mori, K., Beppu, T., & Inoue, Y. (2016). Utility of the measurement of carboxyhemoglobin level at the site of acute carbon monoxide poisoning in rural areas. *Scientifica, 2016*, Article 6192369. https://doi.org/10.1155/2016/6192369

Papin, M., Latour, C., Leclère, B., & Javaudin, F. (2023). Accuracy of pulse CO-oximetry to evaluate blood carboxyhemoglobin level: A systematic review and meta-analysis of diagnostic test accuracy studies. *European journal of emergency medicine: Official Journal of the European Society for Emergency Medicine*, *30*(4), 233-243. https://doi.org/10.1097/MEJ.0000000000001043

Raju, S., Siddharthan, T., & McCormack, M. C. (2020). Indoor air pollution and respiratory health. *Clinics in Chest Medicine*, *41*(4), 825-843. https://doi.org/10.1016/j.ccm.2020.08.014

Ramani, V. K., Mhaske, M., & Naik, R. (2023). Assessment of carbon monoxide in exhaled breath using the smokerlyzer handheld machine: A cross-sectional study. *Tobacco Use Insights, 16*, 1-8. https://doi.org/10.1177/1179173X231184129

Sharma, A., Kumar, R., & Varadwaj, P. (2023). Smelling the disease: Diagnostic potential of breath analysis. *Molecular Diagnosis & Therapy*, *27*, 321-347. https://doi.org/10.1007/s40291-023-00640-7

Shi, D., Ye, Y., Gillwald, M., & Hecht, M. (2021). Designing a lightweight 1D convolutional neural network with Bayesian optimization for wheel flat detection using carbody accelerations. *International Journal of Rail Transportation*, *9*(4), 311-341. https://doi.org/10.1080/23248378.2020.1795942

Shreya, S., Annamalai, M., Jirge, V. L., & Sethi, S. (2023). Utility of salivary biomarkers for diagnosis and monitoring the prognosis of nicotine addiction - A systematic review. *Journal of Oral Biology and Craniofacial Research*, *13*(6), 740-750. https://doi.org/10.1016/j.jobcr.2023.10.003

Tan, X., Su, S., Zuo, Z., Guo, X., & Sun, X. (2019). Intrusion detection of UAVs based on the deep belief network optimized by PSO. *Sensors*, *19*(24), Article 5529. https://doi.org/10.3390/s19245529

Thote, P. B., Daigavane, M. B., Daigavane, P., & Gawande, S. P. (2017). An intelligent hybrid approach Using KNN-GA to enhance the performance of digital protection transformer scheme. *Canadian Journal of Electrical and Computer Engineering*, *40*(3), 151-161. https://doi.org/10.1109/CJECE.2016.2631474

Turino, G. M. (1981). Effect of carbon monoxide on the cardiorespiratory system: carbon monoxide toxicity, physiology, and biochemistry. *Circulation*, *63*(1), 253-259.

Ullah, I., Hussain, M. M., Qazi, E., & Aboalsamh, H. (2018). An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert Systems With Applications*, *107*, 61-71. https://doi.org/10.1016/j.eswa.2018.04.021

Usmani, Z. C., Craig, P., Shipton, D., & Tappin, D. (2008). Comparison of CO breath testing and women's self-reporting of smoking behaviour for identifying smoking during pregnancy. *Substance Abuse Treatment, Prevention, and Policy*, *3*, Article 4. https://doi.org/10.1186/1747-597X-3-4

Vaghefi, E., Yang, S., Hill, S., Humphrey, G., Walker, N., & Squirrell, D. (2019). Detection of smoking status from retinal images: A convolutional neural network study. *Scientific Reports*, *9*, Article 7180. https://doi.org/10.1038/s41598-019-43670-0

Yang, H., Zhang, Y., Yin, C., & Ding, W. (2021). Ultra-lightweight CNN design based on neural architecture search and knowledge distillation: A novel method to build the automatic recognition model of space target ISAR images. *Defence Technology*, *18*(6), 1073-1095. https://doi.org/10.1016/j.dt.2021.04.014

Yi, M., Zhang, N., Liu, X., Liu, J., Zhang, X., Wei, Y., & Shangguan, D. (2023). A mitochondria-targeted fluorescent probe for imaging of endogenous carbon monoxide in living cells. *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*, *291*, Article 122377. https://doi.org/10.1016/j.saa.2023.122377

Zhao, S., & Zhao, Z. (2021). A comparative study of landslide susceptibility mapping using SVM and PSO-SVM models based on Grid and Slope Units. *Mathematical Problems in Engineering*, *2021*, Article 8854606. https://doi.org/10.1155/2021/8854606