IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 June 2024, accepted 22 June 2024, date of publication 25 July 2024, date of current version 11 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3433513

SURVEY

State-of-the-Art in 1D Convolutional Neural Networks: A Survey

AYOKUNLE OLALEKAN IGE AND MALUSI SIBIYA
Department of Computer Science, University of South Africa, Florida Campus, Johannesburg 1709, South Africa
Corresponding author: Malusi Sibiya (sibiym@unisa.ac.za)

ABSTRACT Deep learning architectures have brought about new heights in computer vision, with the most common approach being the Convolutional Neural Network (CNN). Through CNN, tasks previously deemed unattainable, including facial recognition, autonomous driving systems, and sophisticated medical diagnostics, among others can now be achieved. Convolutional layers, non-linear processing units, and subsampling layers are used in conjunction throughout the several learning phases that make up CNN's structure. Generally, 2D and 3D CNNs have been used to achieve impressive results across numerous areas, and several survey papers have been published to review their state-of-the-art applications. However, they are unsuitable in some domain-specific areas where temporal dynamics and dependencies must be captured. Examples of such domains are time series prediction and signal identification, which necessitates the use of one-dimensional signals. Recently, 1D-CNN has evolved and has been used to develop various state-of-the-art models that cut across numerous research fields. However, there has been no survey paper detailing the evolution and advancements in the applications of the 1D-CNN to several computer vision tasks. In addressing this gap, this paper provides the first exhaustive survey to examine the historical development of 1D-CNNs and elucidate their structural intricacies and architectural frameworks. It also highlights recent advancements in their applications across more than twelve distinct domains. Furthermore, this paper provides an overview of the significant challenges impacting the current state-of-the-art 1D-CNN training and deployment while highlighting potential directions for future research exploration. By carrying out this survey, researchers across several fields can have a comprehensive understanding of the evolution, structural intricacies, and recent advancements in the applications of 1D-CNNs across various computer vision tasks. This paper also equip researchers with the knowledge needed to address the significant challenges faced in the current state-of-the-art 1D-CNN hurdles.

INDEX TERMS 1D-CNN, Conv1D, survey, state-of-the-art, challenges.

I. INTRODUCTION

The Convolutional Neural Network (CNN) stands out as the premier algorithm in the field of computer vision. It has emerged as one of the most prominent and widely recognized architectures among deep learning neural networks. CNN has seen applications across various fields, such as medical [1], autonomous vehicles [2], geological mapping [3], art [4], and many others. Several companies, including Google, Microsoft, Meta, and others, have also established research groups to investigate novel CNN designs to address specific problems [5], [6]. The origins of CNN can be traced back to the 1980s when they were initially presented as a neural network model for problems involving image processing [7]. However, the first mathematical representation of neurons, known as the MP model, was first introduced in 1943 by McCulloch and Pitts [8]. During the late 1950s, Rosenblatt [9] introduced a single-layer perceptron model, which enhances the MP model with learning capabilities. However, when LeCun et al. [10] proposed a neural network architecture called LeNet for handwritten character recognition, the advancements in CNN began. CNNs were utilized in the 1990s for increasingly difficult image identification tasks, including face recognition [11], and in the 2000s for object

The associate editor coordinating the review of this manuscript and approving it for publication was Jolanta Mizera-Pietraszko.

recognition and detection in images [12]. The breakthrough for CNNs came in 2012 when Krizhevsky et al. [13] proposed AlexNet, which won the ImageNet Large Scale Visual Recognition Challenge. Since then, CNNs have emerged as the most advanced method for various computer vision problems, such as segmentation, object identification, and image classification.

CNN's capacity to take advantage of spatial correlation in data makes it appealing. Convolutional layers, non-linear processing units, and subsampling layers are used in conjunction throughout the several learning phases that make up CNN's structure [5]. However, it is often domain-specific. For instance, the initial purpose of 2D-CNN was to recognize objects in 2D signals, including pictures and video frames. Also, 3D-CNN extends the 2D CNN by including additional dimensions to process temporal information. The use of 3D convolutional layers allows the network to learn spatial and temporal features from volumetric data, making it useful in applications such as video analysis [14], 3D object recognition [15], and medical imaging [16]. Both systems, however, have high computational complexity, necessitating resource-intensive hardware for model training. Therefore, real-time mobile and low-power/low-memory applications are not a good fit for 2D and 3D CNNs [17].

In some cases where the input signal is one-dimensional, the 1D-CNN is more suitable. The 1D-CNN has recently achieved state-of-the-art across various domains, gaining more interest among researchers. They are particularly useful when dealing with sequential data where the local patterns and dependencies matter, and they have shown competitive performance compared to traditional methods in many domains [18]. For example, Junior et al. [19] used 1D-CNN to detect faults and diagnose electric motors, and Moltra and Mandal [20] leveraged 1D-CNN in classifying non-small cell lung cancer, among many other domains. Due to the recent applications of 1D-CNN, which cut across several research areas, there is a need to carry out comprehensive surveys on the use of 1D-CNN to allow researchers across various domains to have solid background and knowledge of 1D-CNN's architecture and its recent applications.

Recently, several survey papers have been written to review the advancements in CNNs, as seen in Khan et al. [5], where the authors presented a survey on the evolutionary history and the architectural innovations in CNN. Likewise, Li et al. [7] presented a survey on the analysis and applications of CNN. A summary of some existing survey papers on CNN and their focus is presented in Table 1.

As shown in Table 1, various survey papers have been presented in the literature on CNN. However, these survey papers have focused more on 2D-CNN, with less discussion on the recent trend of researchers employing 1D-CNN to achieve state-of-the-art. To the best of our knowledge, the only paper that has carried out a comprehensive survey on 1D-CNN was presented by Kiranyaz et al. [23]. However, their work focused on the application of 1D-CNN in the engineering field, while few details were presented on the recent applications, which cut across other domains. Recently, 1D-CNN has achieved state-of-the-art across several domains, extending further than the engineering field. However, no paper has attempted to review the state-of-the-art in 1D-CNN across these domains. To contribute, our paper presents the first comprehensive survey of the recent advancements in the application of 1D-CNN across several domains. Specifically, our contributions can be summarized in four folds:

- Firstly, we present the general overview of CNN and trace the evolution of these architectures from their inception to their current state.

TABLE 1. Existing survey Papers on CNN.

S/N	Paper	Year	Focus
1.	Bharati & Pramanik [21]	2019	The authors presented Region-based CNN and its recent improvements and further discussed the comparison of the speed, accuracy, and simplicity of Fast-CNN, Faster-CNN, and others.
2.	Bashar [22]	2019	The author focused on the application of CNN in image recognition, audio recognition, and natural language processing.
3.	Khan et al., [5]	2020	The authors discussed the evolutionary history and the architectural innovations in CNN by focusing on exploiting spatial and channel information, depth and width of CNN architectures, and multi-path information processing in Deep CNN.
4.	Kiranyaz et al. [23]	2021	The authors discussed the general architecture and principles of 1D-CNN with respect to its engineering applications.
5.	Sahu & Dash [24]	2021	The authors presented the concept of 2D-CNN with some applications in object classification, face recognition, and automatic handwritten recognition.
6.	Habib & Qureshi [25]	2022	The authors emphasized the exploration of various optimization and acceleration techniques utilized to enhance the performance of CNN models. The authors delve into a comprehensive analysis of these techniques, highlighting their crucial role in achieving optimal CNN performance.
7.	Li et al. [7]	2022	The authors discussed classic and advanced CNN models, emphasizing key elements leading to state-of-the-art results. The review presented findings from experimental analyses, drawing conclusions and offering practical guidelines for function selection.
8.	Mi et al. [26]	2022	The authors presented the methods related to the structural design of efficient CNN in recent years and categorized these methods into model pruning, efficient architecture design, and neural architecture search.
9.	Tombe & Viriri [27]	2022	The authors focused on the advances of CNN with respect to convolutional layer design configurations, pooling layer strategies, activation functions, and loss functions.

- Secondly, we introduce the structure and architecture of 1D-CNNs, elucidating their design principles and components in detail.
- Thirdly, a cohesive narrative of recent advancements in 1D CNN across twelve (12) domains is discussed based on the rationale of the studies and the limitations each model aimed to address.
- Lastly, some challenges and grand research areas that researchers can still explore to improve state-of-the-art are presented.

To ensure a comprehensive review of the state-of-the-art in 1D-CNN applications, we conducted an extensive literature search across multiple academic databases, including Scopus, IEEE Xplore, ScienceDirect, Web of Science, and ACM Digital Library. The search strategy was formulated to capture a wide range of studies and articles that utilized 1D-CNN architectures across various domains. The primary keywords used were "Conv1D" and "1D-CNN," combined with secondary keywords such as "Applications," "Pattern Recognition," "Time Series Analysis," "Signal Processing," "Anomaly Detection," "Biomedical Signal Processing," "Speech Recognition," "Sensor Data Analysis," "Financial Forecasting," "Seismic Data Analysis," "Activity Recognition," "Predictive Maintenance," "Healthcare Diagnostics," "EEG Analysis," and "ECG Analysis." This approach enabled us to gather relevant and high-quality research papers in the domain. The remainder of this paper is organized as follows: Section II presents the General overview of CNNs, Section III discusses the architecture of 1D CNN, Section IV presents recent advancements in 1D-CNN across various domains, Section V presents the future research areas, and Section VI concludes.

II. GENERAL OVERVIEW OF CNN

A convolutional neural network is a deep feed-forward artificial neural network, often composed of numerous layers of distinct neural networks containing multiple neurons [28]. The architecture of the CNN layers is shown in Fig. 1 The complete matrix is moved horizontally and vertically in a 2D CNN filter (height and width). The range of convolution operations for each step is determined by the height and width of the 2D filter, as shown in Fig. 2.

In 2D-CNN, the kernel slides along two dimensions of the data, as shown in Fig. 2 The kernel, also known as a filter, is a smaller 2D matrix used to extract features from the feature map. The value within the kernel represents the weights used to perform the convolution operation. In order to do this, the kernel is placed at the top left corner of the input data, as presented in Fig. 2, and the element-wise multiplication between the kernel and the portion of the input data it covers is computed. Thereafter, the results are summed up to obtain a single value, which is then placed in the output feature map. After the initial convolution, the kernel is moved a certain number of strides along the horizontal and vertical dimensions of the input data to learn and detect different patterns and features at different locations in the input. Several computer vision tasks have been carried out using 2D-CNN. Generally, the success of CNN has been built around 2D-CNN, which cuts across various classification tasks [30], segmentation tasks, object detection, and localization [31], among many others. For instance, Poudyal et al. [32] leveraged 2D-CNN to predict students' academic performance, and Yildirim et al. [33] used pre-trained 2D-CNN to detect diabetic subjects. Several other researchers have leveraged 2D-CNN in health-based research, as seen in [34], [35], [36], [37], and [38], among many others. Also, in 3D-CNN, the kernel slides in 3 dimensions, as shown in Fig. 3. The kernel, which is a 3D tensor, extracts features from the input volume in a way similar to the 2D-CNN.

After the initial 3D convolution, the kernel is moved a certain number of steps along the input volume's three dimensions (width, height, and depth). In a 3D-CNN, the kernel moves through the input volume in a 3D grid, allowing the network to learn and detect spatial and temporal features in three dimensions. The output feature volume represents the spatial and temporal distribution of these features. Most often, 3D data, such as data from Magnetic Resonance Imaging (MRI), is utilized with 3D-CNN. The brain, spinal cords, internal organs, and many more structures may all be examined using MRI [39], [40], [41]. A Computerized Tomography (CT) Scan is also an example of 3D data, which is created by combining a series of X-ray images taken from different angles around the body and has been leveraged in many works, as seen in [42], [43], and [4], among many other works. However, in 1D-CNN, the kernel slides along one dimension of the input data. A more detailed discussion of this is provided in the succeeding section. Recently, several techniques have been used to increase the generalization capacity of deep neural networks, including learning rate tweaking, dropout, batch normalization, and data augmentation. Without concentrating on starting parameters and dropout rate, batch normalization allows for a greater learning rate and reduces the number of training steps required for model convergence [45]. Consequently, a batch normalization layer makes the model's training easier and faster, while dropout [46] reduces overfitting and enhances the generalization ability of a neural network. The dropout algorithm sets the neurons in a certain neural network layer to zero at a certain probability. The effect of dropout in neural networks is shown in Fig. 4.

III. ARCHITECTURE OF 1D-CNN

One-dimensional convolutional neural networks (1D-CNNs) have been around since the late 1980s, but it wasn't until the mid-2000s that they gained popularity in the field of computer vision. The first 1D-CNNs were used in speech recognition tasks, where the input data consisted of audio signals. As discussed earlier, in 1989, LeCun et al. introduced the concept of convolutional neural networks

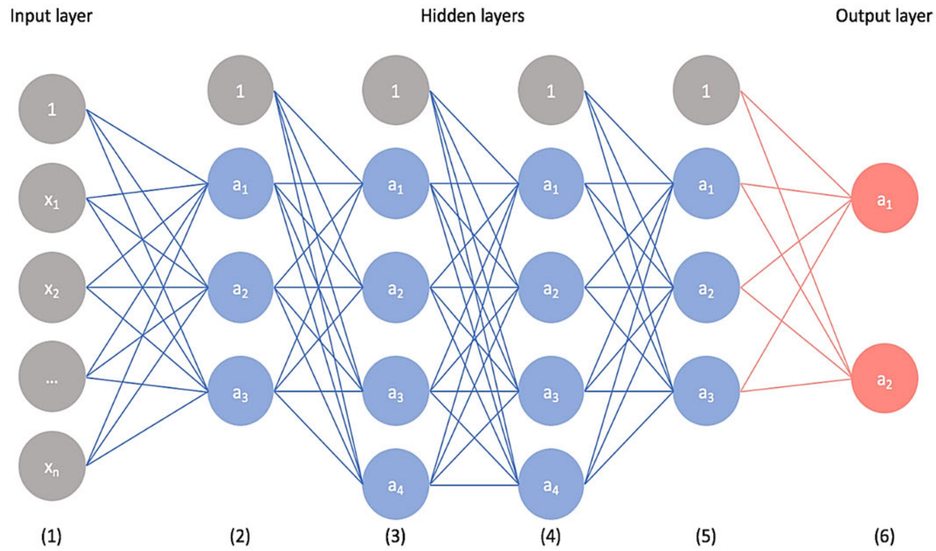

FIGURE 1. Architecture of CNN.

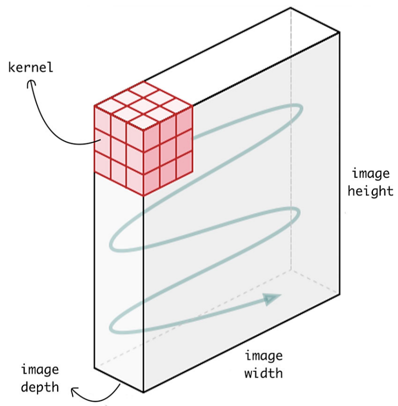

FIGURE 2. Kernel Sliding over an Image in 2D-CNN (Source: [29]).

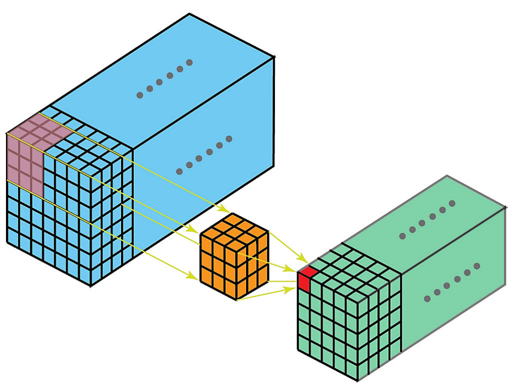

FIGURE 3. Kernel Sliding over an image in 3D-CNN.

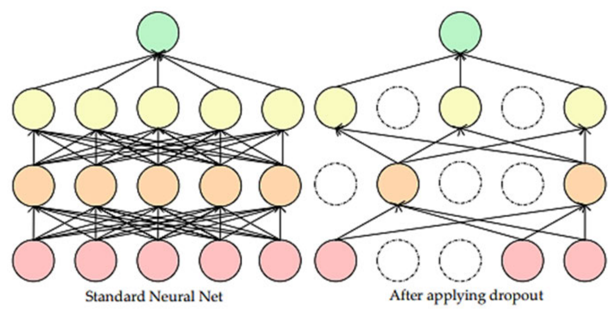

FIGURE 4. Effect of dropout.

in their paper "Backpropagation Applied to Handwritten Zip Code Recognition" [10]. They used CNN to recognize handwritten digits in postal codes, achieving state-of-the-art results at the time. In the early 2000s, 1D-CNNs were used in bioinformatics as seen in [47], and for tasks such as DNA sequence classification and protein structure prediction, as seen in Wang et al. [48] where a deep 1D-CNN, called DeepCNF, for protein secondary structure prediction. Today, 1D-CNNs are widely used in many fields, including finance, healthcare, and manufacturing. They are particularly useful for processing data with a temporal or sequential nature, such as sensor data from IoT devices or signals from medical sensors. Additionally, 1D-CNNs are coupled with other neural network architectures, such as recurrent neural networks (RNNs) and transformers, to create hybrid models that can handle a wider range of data types and tasks. For example, a hybrid model that combines a 1D-CNN with an RNN can be used for speech recognition, where the 1D-CNN is used to extract features from the audio signal, and the RNN is used to model the temporal dependencies between the features. In order to extract the features, the filter moves vertically (height), and the height establishes how many sample points are needed for the convolutional operation in a 1D CNN [19], as shown in Fig. 5.

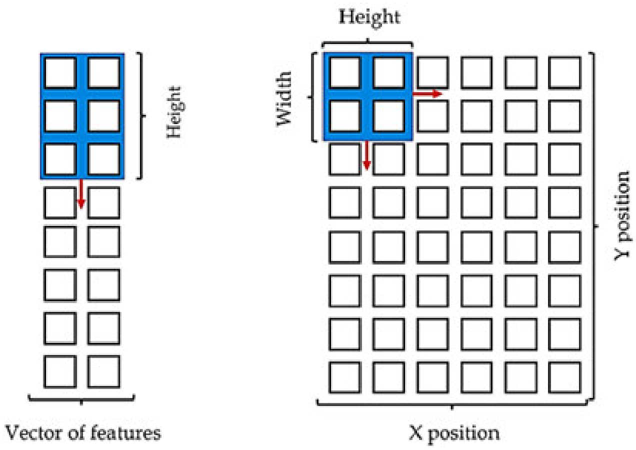

FIGURE 5. Comparison of 1D-CNN to 2D-CNN (Source: [19]).

Generally, 1D-CNNs are designed to handle one-dimensional data, such as time-series data, sequences (e.g., text), or any data where the primary structure is along a single axis. The kernel (or filter) in a 1DCNN moves along one dimension. If the data is represented as a vector $[x_1, x_2, \ldots, x_n]$, the kernel will slide over this vector to detect patterns within the sequence. The shape of the kernel is a 1D array with dimension $(k,)$, where k is the size of the kernel. The kernel slides along the single axis of the input vector. For example, in a time-series application, the kernel moves along the time axis to capture temporal patterns. The receptive fields of a 1DCNN kernel is a contiguous segment of the input sequence. Therefore, as the kernel slides over the input, it aggregates information from k consecutive elements. For an input sequence x and a kernel w, the convolution operation in a 1D-CNN layer is given as:

$$(x * w)(t) = \sum_{i=0}^{k-1} x(t+i) \cdot w(i) \tag{1}$$

where:

- x is the input sequence,
- w is the kernel (or filter),
- $(x * w)(t)$ denotes the convolution of x and w at position t,
- k is the size of the kernel,
- $x(t + i)$ is the element of the input sequence at position $t + i$,
- $w(i)$ is the element of the kernel at position i.

An illustration of three consecutive convolutional layers is presented in Fig. 6, as seen in [49], where x_k^i is used to denote the input, b_k^i is the bias of the neuron at kth position of layer l, and the output of the ith neuron in the incoming layer $l-1$ is given as $s_i^{(l-1)}$, and $w_{ik}^{(l-1)}$ denotes the kernel assigned from the neuron in the ith position of the first convolutional layer $l-1$ to the kth neuron in the second layer l, y_k^i is the intermediate output, SS is the scalar factor used in downsampling, and f is the activation function.

Forward propagation in 1D-CNN involves passing the input through one or more convolutional layers, pooling layers, and fully connected layers, such that feature map Z_c is given as:

$$Z_c = f_c(X * W_c + b_c) \tag{2}$$

where X is the input, W_c denotes the filter weights, b_c is the bias term, and f_c is the activation function of the convolution, and $X * W_c$ is the convolutional operation between filter weights and bias terms. The spatial dimension of Z_c is reduced by aggregating information from nearby values through pooling, which is given as:

$$A_p = P(Z_c) \tag{3}$$

where P is the pooling operation. Then, a fully connected layer Z_f combines the features learned from the convolutional and pooling operations, and the final activation function Y is used to obtain the output of the network. Also, backward propagation in 1D-CNN involves computing gradients of the loss function with respect to the network's parameters, which are used to update the weights and biases. The backward propagation in the fully connected layer is such that:

$$\frac{\partial L}{\partial Z_f} = \frac{\partial L}{\partial Y} \cdot f_f'(Z_f) \tag{4}$$

where L is the loss function and f_f' is the activation function in the fully connected layer. $\frac{\partial L}{\partial W_f}$ is the gradient loss with respect to fully connected weights, and it is given as:

$$\frac{\partial L}{\partial W_f} = \frac{1}{m} \frac{\partial L}{\partial Z_f} \cdot A_p^T \tag{5}$$

and $\frac{\partial L}{\partial b_f}$ is the gradient of the loss with respect to fully connected biases, and it is given as:

$$\frac{\partial L}{\partial b_f} = \frac{1}{m} \sum (\frac{\partial L}{\partial Z_f}) \tag{6}$$

The backpropagation through the pooling layer is calculated as presented in equation (3). The backpropagation through the convolutional layer is given as:

$$\frac{\partial L}{\partial Z_c} = \frac{\partial L}{\partial A_p} \cdot P'(Z_c) \tag{7}$$

$$\frac{\partial L}{\partial W_c} = \frac{1}{m} X * \frac{\partial L}{\partial Z_c} \tag{8}$$

$$\frac{\partial L}{\partial b_c} = \frac{1}{m} \sum (\frac{\partial L}{\partial Z_c}) \tag{9}$$

where m is the batch size, $P'(Z_c)$ is the gradient of the pooling operation, $\frac{\partial L}{\partial W_c}$ is the gradient of the loss with respect to convolutional weights, and $\frac{\partial L}{\partial b_c}$ is that of the biases. The computed gradients are then used to update the convolutional weights W_c and biases b_c using gradient descent.

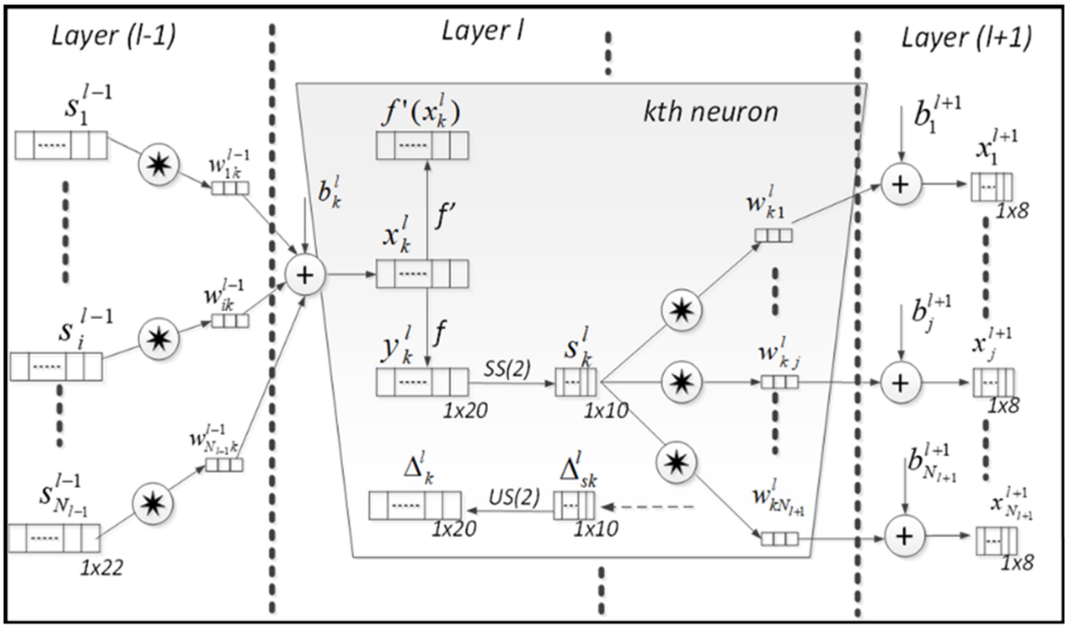

FIGURE 6. Illustration of three consecutive layers in 1D-CNN [49]).

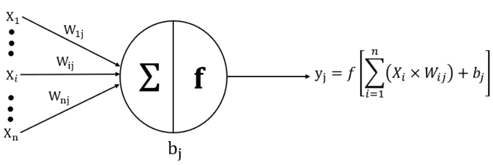

FIGURE 7. Structure of Activation function.

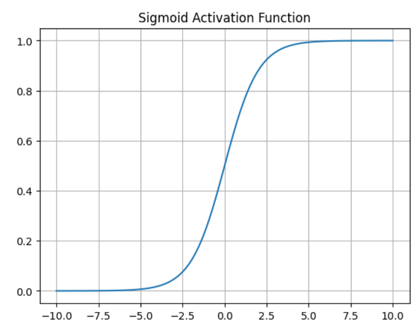

FIGURE 8. Plot of Sigmoid activation function.

A. ACTIVATION FUNCTIONS

The activation function chooses which information should be sent to the succeeding neuron, just like the human brain's neuron model does. Each neuron in the neural network accepts the output value of the neurons from the previous layer as input and passes the processed value to the next layer [7]. Activation functions are used in neural networks to improve non-linear change. A neural network with no activation function is equivalent to a matrix multiplication; this holds true even when several layers are added to the neural network. Neuron operation and input-to-output mapping are carried out via the activation function. It completes the learning of the model and makes it understand complex non-linear functions [50]. The general structure of an activation function is presented in Fig. 7.

Where x_i is the input feature, and n features are the input to neuron j, while W_{ij} is the weight of the connection between x_i and j, which is the value of the bias, and y_j is the output of the neuron j.

The next section presents some common activation functions that researchers have used in learning the inherent features of one-dimensional data. There are several activation functions commonly used in deep learning, among which are ReLU, sigmoid, tanh, softmax, and others.

1) SIGMOID

The sigmoid function transforms the input value to the output between the range from 0 and 1. It is also called a logistic function, and the curve of the function looks S-shaped. It is often used in the output layer of binary classification neural networks. The plot of the sigmoid activation function is shown in Fig. 8.

2) TANH FUNCTION

Tanh (Tangent Hyperbolic) function scales data to the range from -1 to 1 and centres the mean to 0. It is similar to sigmoid, and the curve is S-shaped. The plot is presented in Fig. 9.

The tanh activation function is defined as:

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{10}$$

where e is the base of the natural logarithm.

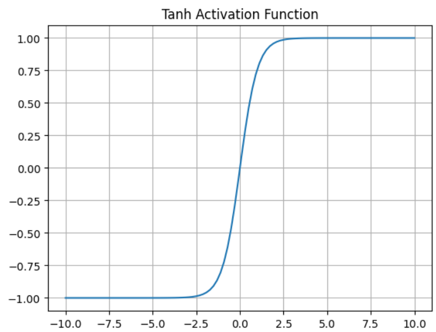

FIGURE 9. Plot of Tanh activation function.

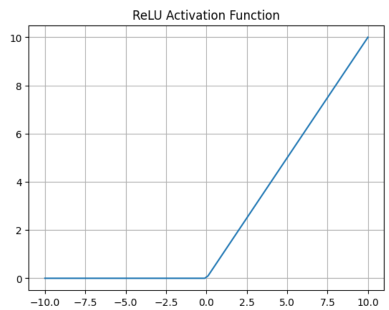

FIGURE 10. Plot of ReLU activation function.

3) RELU

As shown in Fig. 10, Rectified linear units, or ReLUs, are used in functions that convert negative input values to 0 and retain positive values. The output of the ReLU activation is the maximum of 0, and the input x. The ReLU is generally more computationally efficient and less expensive than the sigmoid and tanh functions. Unlike sigmoid and tanh activations that squash input values, ReLU does not saturate positive values, which helps in mitigating the vanishing gradient problem, making it easier for the model to learn and propagate gradients during backpropagation.

4) LEAKY RELU

Since the ReLU function converts all negative inputs to 0, some neurons may become inactive during training because they consistently output zero for all inputs, which results in the dying ReLU problem. The leaky ReLU addresses this by providing a small positive slope for some values less than 0. The Leaky ReLU activation function is defined as:

$$\text{Leaky ReLU}(x) = \begin{cases} x, & \text{if } x > 0 \\ \alpha x, & \text{if } x \leq 0 \end{cases} \tag{11}$$

where α is a positive constant in the range of 0.01. The plot of the Leaky ReLU is shown in Fig. 11.

In Chan and Noor [51], Leaky ReLU was used to design the discriminator of a conditional GAN model, which was designed using 1D-CNN.

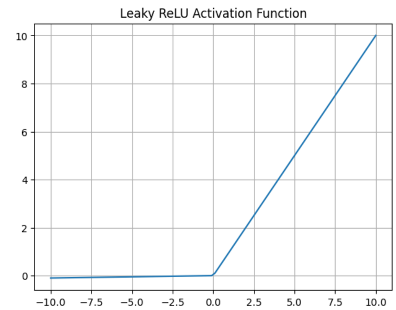

FIGURE 11. Plot of Leaky ReLU activation function.

B. LOSS FUNCTIONS

Gradient descent finds the optimal weights by iteratively updating the weights such that the weights will result in a minimum prediction error over all instances in the training set [52]. A loss function quantifies the prediction error and plays a pivotal role in training 1D Convolutional Neural Networks. Choosing an appropriate loss function is critical for guiding the network's learning process and optimizing its parameters based on the given task. Examples include mean squared error, binary cross-entropy, and categorical cross-entropy.

1) MEAN SQUARED ERROR

Mean Squared Error (MSE) is a commonly used loss function in regression tasks involving 1D-CNNs. The average squared difference between the actual (ground truth) values and the predicted values is measured by the MSE loss [53]. Mathematically, it is defined as: Mathematically, it is defined as:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{12}$$

where n is the number of samples in the dataset, y_i is the actual value of the ith sample, and $\hat{y}_i$ is the predicted value.

2) BINARY CROSS-ENTROPY LOSS

Binary Cross Entropy (BCE), also known as log loss or logistic loss, is a common loss function used in binary classification problems. The BCE loss is particularly suitable when the task involves assigning data points to one of two classes (binary classification), and it is often used in the output layer of a neural network [54]. The BCE is defined as:

$$\text{BCE}(y, \hat{y}) = -\left[y \log(\hat{y}) + (1-y)\log(1-\hat{y})\right] \tag{13}$$

where y is the true label (0 or 1) and $\hat{y}$ is the predicted probability that a sample belongs to class 1.

3) CATEGORICAL CROSS-ENTROPY LOSS

Categorical Cross-Entropy (CCE) is a loss function commonly used in multiclass classification tasks. This loss function is well-suited for problems where each example in the dataset can belong to one of multiple classes. It's frequently used in the output layer of neural networks, especially when the task involves predicting class labels. The CCE is defined as:

$$\mathrm{CCE}(y, \hat{y}) = -\sum_{i} y_i \log(\hat{y}_i) \tag{14}$$

CCE measures the dissimilarity between the predicted probability distribution $\hat{y}$ and the actual probability distribution y for a given sample and takes the sum over all classes. Other loss functions include sparse CCE, Huber loss, Mean Absolute Error, and other custom losses, as seen in [55], [56], and [57], among many others.

IV. RECENT APPLICATIONS OF 1D-CNN

The adoption of 1D-CNN has witnessed significant growth across various domains in recent years, owing to their remarkable ability to capture patterns and features in sequential data. One of the prominent areas where 1D-CNN has made a substantial impact is in natural language processing (NLP). By treating text as a one-dimensional sequence of words or characters, 1D-CNNs have shown impressive results in tasks such as text classification [58], sentiment analysis [59], and named entity recognition [60]. Their ability to capture local and global dependencies within the text has proven to be invaluable for these applications. In healthcare, 1D-CNNs have emerged as a vital tool for time-series data analysis. Medical sensors, such as electrocardiograms (ECG) and electroencephalograms (EEG), generate sequential data that contain critical diagnostic information. 1D-CNNs have been instrumental in detecting anomalies, predicting diseases, and accurately classifying medical conditions. They can automatically learn discriminative features from the temporal patterns, making them invaluable in early disease diagnosis and monitoring [61]. Additionally, the field of finance has seen a surge in the adoption of 1D-CNN for various tasks, such as stock price prediction [62], fraud detection [63], and algorithmic trading [64], among many others. These networks excel at capturing intricate temporal relationships within financial time-series data, enabling more precise predictions and improved decision-making. In the audio domain, 1D-CNNs have been instrumental in speech recognition [65], music genre classification [66], and sound event detection [67]. Their capacity to process sequential audio data efficiently has revolutionized how we interact with voice-controlled devices and enhance our ability to analyze and understand audio content. In other words, 1D-CNN has permeated diverse domains, ranging from NLP and healthcare to finance and audio processing, where sequential data analysis is crucial. Their capacity to capture temporal patterns and dependencies has substantially improved various applications, making them a versatile and powerful tool. The following subsection discusses some recent applications of 1D-CNN across various domains and discussed based on the rationale of the studies and the limitations the models address.

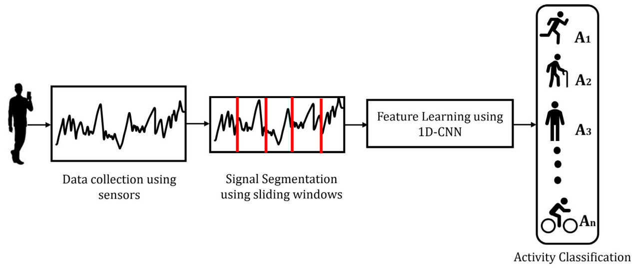

FIGURE 12. Process of 1D-CNN feature learning in HAR.

A. HUMAN ACTIVITY RECOGNITION

Human Activity Recognition (HAR) is a branch of research that defines and tests novel approaches for accurately recognizing human activities using signals [68]. Human activity signals can be obtained using sensors, and these signals are collected in a time series format. Therefore, 1D-CNN has been employed for automatic feature learning, as they perform better on multi-variate time series data [69]. As shown in Fig. 12, the signals from sensors must first be divided into windows, and their features must be extracted before activities can be recognized. In wearable sensor based HAR, numerous benchmark datasets have been established to facilitate research and development. Among these, the Physical Activity Monitoring and Assessment Platform 2 (PAMAP2) dataset stands out, encompassing twelve protocol activities and six optional ones. Similarly, the Wireless Sensor Data Mining (WISDM) dataset offers valuable insights, comprising data from six distinct activities. Another notable dataset is the UCI Human Activity Recognition (UCI-HAR) dataset, which has a collection of six activities, both dynamic and static. Furthermore, the Human Activity Prediction and Recognition of Activities of Daily Living using Smartphone Sensors (HAPT) dataset expands upon the UCI-HAR dataset, incorporating six transitional activities alongside its dynamic and static counterparts. These datasets and several others serve as pivotal resources for researchers in the domain of activity recognition, enabling the development and evaluation of machine and deep learning models.

Several state-of-the-art architectures have been developed for human activity recognition from wearable sensors using 1D-CNN and benchmarked on these datasets. For instance, in order to find the best network connections and hyper-parameters to enhance model performance, Ragab et al. [70] proposed a random search 1D-CNN model. They evaluated the model on the UCI-HAR dataset, and the result showed that the model achieved improved recognition accuracy when classifying the six activities in the dataset. However, the model exhibited extended training times due to the dynamic nature of some activities within the dataset. In order to minimize model training time when dealing with dynamic activities, Banjarey et al. [71] used 1D-CNN to recognize activities such as reading, tilting, walking, sleeping, and others, and experiments showed that 1D-CNN had faster training time compared to 2D-CNN on the same dataset.

Even though 1D-CNN had minimal training time, the quality of features learned using 1D-CNN was not optimal. For this reason, Khan and Ahmad [72] used 1D-CNN to design three lightweight convolutional heads, as shown in Fig. 13, with each head designed to extract features from wearable sensor data. One dimensional CNN layers were stacked across three heads, and an attention mechanism was included in each head to boost feature learning. The results showed that the fusion of multiple 1D-CNN heads can improve feature learning in HAR. However, 1D-CNN has limitations when learning features of activities with inter-class similarity, a situation whereby some activities have similar patterns.

To address the limitations of 1D-CNNs in dealing with activities with inter-class similarity, the authors in [69] combined 1D-CNN with squeeze and excitation block and used an adaptive reduction ratio to increase the precision and recall of human activities with similar characteristics. The model used 1D-CNN with progressively increasing kernel sizes for feature learning before flattening the feature maps to ensure features in the channel dimension were considered. The input of the SE block was then scaled with the fully connected layer with a sigmoid in the SE block before adding another flattened layer. The model was evaluated on PAMAP2, WISDM, and UCI-HAR datasets, and experiments showed that the model could learn more intricate features of similar activities. However, the model had bulky model parameters. In recent times, HAR researchers have focused on how to improve feature learning with minimal model parameters. To achieve this, Goh et al. [73] proposed the integration of 1D-CNN with LSTM for the identification of human activities. The LSTM was utilized to encode the temporal dependencies of the features, and the 1D-CNN was employed to extract high-level features from the sensor data. A similar work by Luwe et al. [74] proposed a hybrid deep learning model that combined 1D-CNN with Bidirectional-LSTM. Dominant features were converted into high-level representative features using the one-dimensional CNN. Then, the Bidirectional-LSTM encodes the long-range relationships in the features via gating mechanisms. The model was evaluated on PAMAP2, WISDM and UCI-HAR datasets, and the results showed that improved features were learned with minimal model parameters. However, evaluation of the model on dynamic and transitional activities revealed that an improvement is necessary. To contribute, Abdel-Basset [75] proposed the ST-DeepHAR model, presented in Fig. 14, which takes advantage of the attention mechanism with LSTMs for sequential feature learning before concatenating the feature maps with a 1D-CNN layer. Experimental results showed increased model performance compared to existing architectures, but low feature learning in transitional activities was not discussed.

In the work of Mohd Noor et al. [76], a Deep Temporal Conv-LSTM, which used the relationship of the sliding window to improve feature learning, was proposed. The authors combined three 1D-CNN pipelines with fixed kernel sizes and evaluated the model on a dataset that consists of transitional activities (HAPT dataset). Experiments showed that the model achieved an improved recognition accuracy but with high model training time. The architecture developed by Mohd Noor et al. [76] is presented in Fig. 15.

Also, Mohd Noor [77] proposed a Convolutional denoising autoencoder model, which used 1D-CNN in an autoencoder. The 1D convolution operation leverages the local temporal structure of unlabeled human activity signals in the encoder before upsampling. The model was evaluated using transitional activities and achieved a recognition accuracy of 93.40%. Similar research in [78], which extends the work of Mohd Noor [77], proposed a denoising autoencoder model with an attention mechanism. The model was designed using 1D-CNN to offer smaller and improved features for activity classification based on the value of the latent layer. Experiments showed that the model improved feature learning in transitional activities by achieving a recognition accuracy of 94.40% on the HAPT dataset. However, the model had high training time and was relatively bulky. To improve feature learning with minimal model training time, Bhattacharya et al. [79] proposed the Ensem-HAR, which is an ensemble of four deep learning-based classification models, namely, CNN-net, CNNLSTM-net, ConvLSTM-net, and StackedLSTM-net. The models leveraged 1D-CNN, and a meta-learner was trained on the stacked prediction to output the final prediction on test data. The model was evaluated on three datasets and showed improved feature learning capabilities but at a high computational cost. Also, experiments showed that the size of the sliding window used for activity signal segmentation directly affects the quality of learned features and model parameters of HAR models. To address this, WSense, a plug-and-play neural network module, was proposed in [80], which uses 1D-CNN with ELU activation function to minimize the difference between model size and learned features, regardless of the sliding window size. In the work of Ige and Noor [18], the WSense module was placed at the top of an architecture developed for improved feature learning. The model combined 1D-CNN and LSTM feature learning pipelines across six heads and used the WSense to reduce the number of model parameters. Experiments showed that the model improved feature learning with minimal model size, achieving state-of-the-art on the evaluated datasets. Recently, the application of 1D-CNN in HAR has been extended to synthetic data generation, as seen in [81], where synthetic human activity data was generated using the Wasserstein Generative Adversarial Network. The GAN model was built using 1D-CNN in the generator and discriminator. The quality of the generated data was also evaluated using a feature learning pipeline designed using 1D-CNN. In another research, AFAR [82] was proposed for activity monitoring and real-time fall detection, and 1D-CNN was employed in designing the AFAR model. Based on the discussed literature, the importance of 1D-CNN in activity

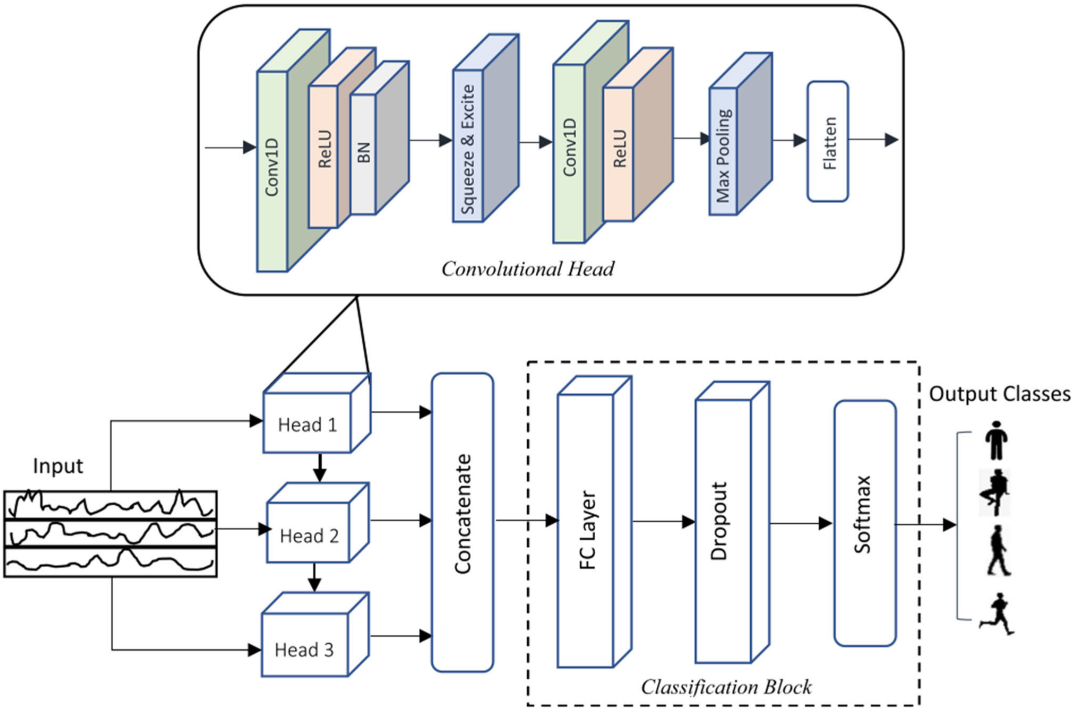

FIGURE 13. **Multi-head Attention induced 1D-CNN proposed by Khan and Ahmad [72]).**

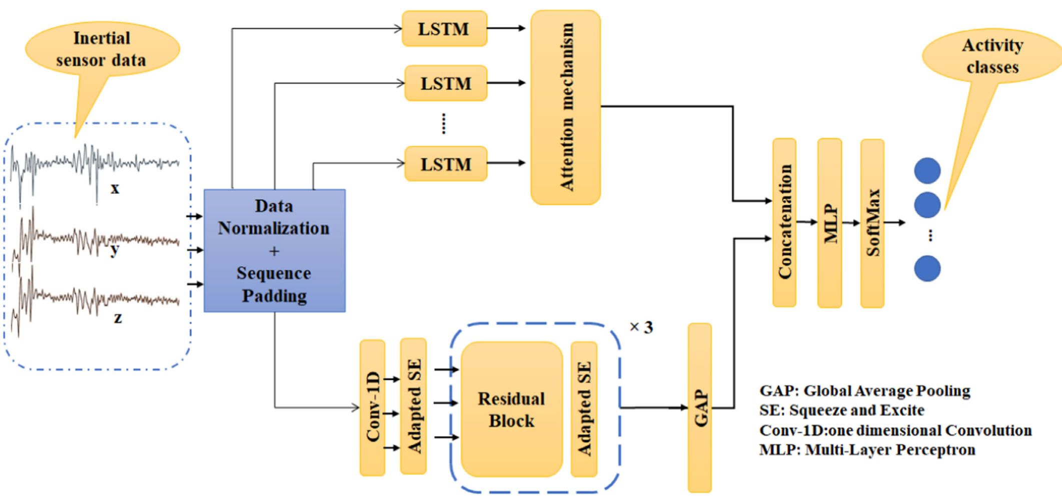

FIGURE 14. **The model proposed by Abdel-Basset [75]).**

recognition cannot be over-emphasized, as several state-of-the-art supervised, unsupervised, and generative models have been developed using 1D-CNN in the domain.

B. SPEECH EMOTION

Speech Emotion Recognition (SER) is crucial to many successful services, including call center support and monitoring of consumer emotions to improve services [83]. Similar to this, in the sphere of medicine, speech-based diagnostic tools are created to assess the degree of depression and distress [84]. SER extracts human emotion from brief voice inputs to facilitate realistic human-computer interactions. To develop interactive real-time applications, precise emotion recognition is necessary. In real-time scenarios, inaccurate or bad forecasts may lead to awkward and annoying situations [85]. Several researchers have developed sophisticated SER systems using 1D-CNN in order to address specific challenges. For instance, Chourasia et al. [86] developed a model based on the features extracted using Mel-frequency cepstral coefficient spectrograms with 1D-CNN to capture long-term temporal dependency in speech signals. The

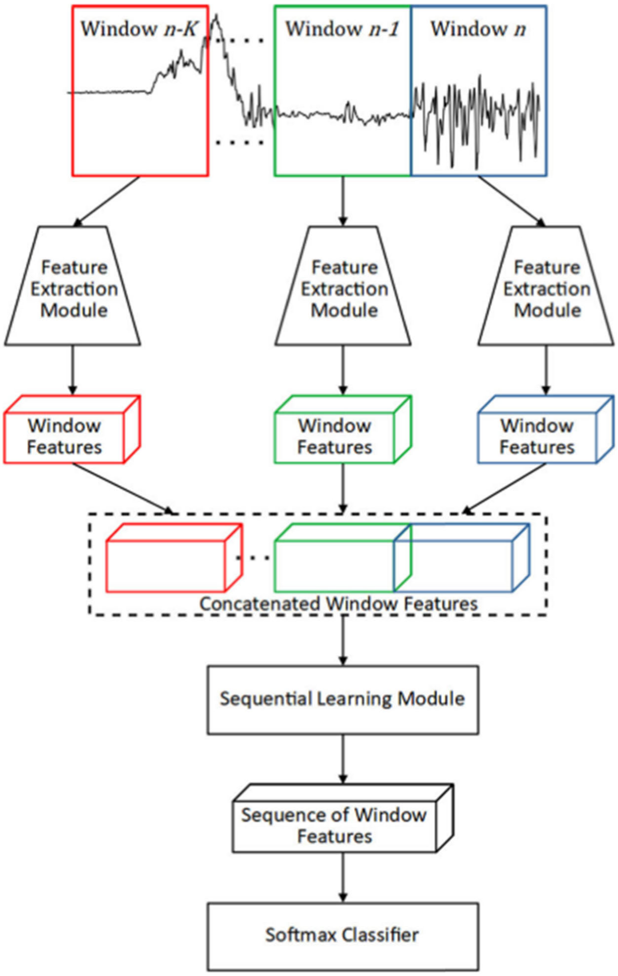

FIGURE 15. Deep Temporal Conv-LSTM model proposed by Mohd Noor et al. [76].

model was evaluated on the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) dataset that consists of happy, calm, sad, angry, and nervous emotions based on gender. However, because emotion recognition systems often rely on understanding the dynamics of speech over time, the architecture could not learn improved features due to the limited receptive field of 1D-CNN.

To address the challenges of feature learning caused by the limited receptive field of 1D-CNN, the authors in [85] proposed a Neural Network-based Blended Ensemble Learning model developed using 1D-CNN, LSTM, and Capsule networks. The model was evaluated on two benchmark datasets (RAVDESS and IEMOCAP). Even though the model achieved improved feature learning, it ignored the long-term dependency of the features. To address this, Mustaqeem and Kwon [67] proposed MLT-DNet, shown in Fig. 16. The MLT-DNet architecture worked as an end-to-end real-time speech emotion recognition system developed using dilated 1D-CNN. The MLT-DNet model extracted spatial and long-term contextual dependency using a multi-learning

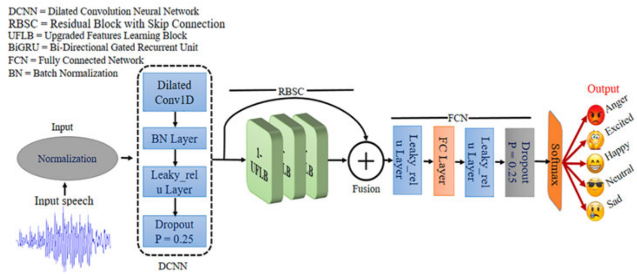

FIGURE 16. The MLT-DNet model proposed in Mustaqeem and Kwon [67].

strategy and evaluated it on IEMOCAP and EMO - DB datasets.

Also, an end-to-end learning strategy for speech recognition based on multiscale convolutions that acquire the representation straight from audio waveforms developed was proposed in Zhu et al. [87]. Three 1D convolutional layers with various kernel sizes were utilized for feature extraction, and a pooling layer concatenated the features to guarantee a constant sample rate for the remainder of the network. A dataset consisting of read, conversational, accented, and noisy speech was used to calculate the word mistake rate, which was reported to be 23.28%. The results showed that concatenating multiple 1D-CNN heads for feature learning in speech emotion recognition can allow 1D-CNN to learn rich hierarchical representations automatically for accurate classification. However, selecting the most discriminative features for emotion recognition remains an ongoing challenge.

C. MUSIC GENRE AND SOUND CLASSIFICATION

Classification of musical genres is an extensively researched topic in the Music Information Research (MIR) field. Generally, researchers have shown interest in speaker identification [88] and environmental sound classification [89] due to their applications in crime detection, environmental context-aware processing, and many others [90]. There are primarily two components in music classification problems. One is preprocessing the unprocessed audio data and the other is building the classification model. According to Bian et al. [91], acoustic feature extraction and spectrogram transformation are the two main ways to preprocess raw audio. However, the steps of preprocessing raw audio for machine learning classification are presented in Fig. 17.

As shown in Fig. 17, the sampling rate of the raw audio is first converted to ensure that all audio samples have a consistent sampling rate. After that, frame segmentation, which divides the audio into smaller segments or frames, is used to create a sequence of fixed-length samples before handcrafted features can be extracted. Examples of such handcrafted features include Mel-frequency cepstral coefficients (MFCCs) [92], which are widely used in speech and music processing due to their ability to capture spectral information, Chroma features [93], which represent the

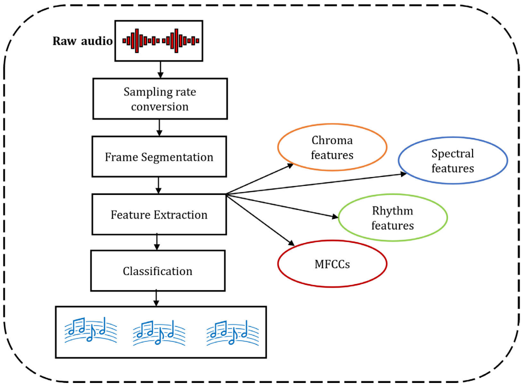

FIGURE 17. Raw Audio Preprocessing steps in Machine Learning.

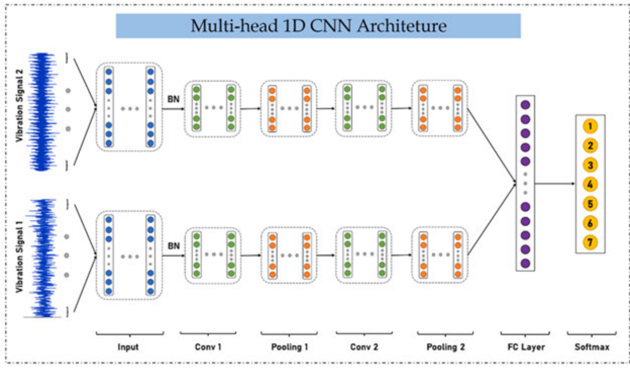

FIGURE 18. Multi-head 1D-CNN Architecture proposed in Junior et al. [19].

distribution of musical pitch classes in an audio segment. They are useful for capturing tonal information in music, Spectral contrast [94], which measures the difference in amplitude between peaks and valleys in the audio spectrum and can be indicative of the texture and timbre of the music, and Rhythm Features [95], which relates to tempo, beat, and rhythm, such as beat histograms or tempo estimation, which can be informative for distinguishing music genres. Others include Zero-Crossing Rate [96], Statistical Moments [97], and others. However, manually extracting features in music genre classification often limits expressiveness since they are extracted based on intuition and domain knowledge. Similarly, manual feature extraction is tedious in music genre classification since it has limited adaptability to new genres. For instance, if a music genre that was not considered during the feature design process is to be classified, revisiting and extending the feature set is important to achieve improved classification, which can be a challenging task. To address these limitations, recent studies have adopted 2D and 1D-CNN for music genre classification, as seen in [91], [98], and [99] and most recently in Farajzadeh et al. [100], where the authors used 2D-CNN for feature learning from the PMG-dataset which consist of pop, rap, monody, rock and traditional genres of Persian music, and achieved an accuracy of 86%. However, the limitations of converting audio signals into 2D representations, which often add to computational overhead and cause information loss, made 1D-CNN important in automatic music genre classification. The existing research works that have leveraged 1D-CNN for music genre classification is presented in Table 6.

D. FAULT DETECTION IN ELECTRIC MOTORS

Electric motors are the backbone of many industries, serving as vital energy sources. Maintaining their operational integrity is crucial for sustaining productivity and job security. In the early stages of research, Ince et al. [106] pioneered using 1D-CNN for fault detection in electric motors, demonstrating its ability to streamline feature extraction and classification processes into a single cohesive framework. Subsequent studies, such as the comparative analysis conducted by [107], underscored the superiority of 1D-CNN over 2D-CNN in terms of accuracy and efficiency for fault detection tasks. Building upon these findings, researchers have explored various avenues to enhance fault detection techniques leveraging 1D-CNN. For instance, Junior et al. [19] adopted a multi-head 1D-CNN approach, harnessing vibration signals from accelerometers, as seen in Fig. 18, to achieve high detection accuracy.

In a related study, Chen et al. [108] proposed a fault diagnosis model for rolling bearings utilizing 1D-CNN and employing the tanh activation function. The model achieved a detection accuracy of 98.30% when multiple loads were used and 99.20% when a single load was placed on the equipment, outperforming other methods, including LSTM. Furthermore, the development of hybrid models, as seen in [109], for analyzing voltage and current signals in three-phase induction motors, has contributed to more accurate fault diagnosis, ultimately improving maintenance efficiency and reducing downtime in industrial processes. These advancements highlight the evolution of fault detection methodologies within the electric motor domain, driven by the continuous refinement of 1D-CNN architectures. Other works that have leveraged 1D-CNN for fault diagnosis are presented in Table 4.

E. STROKE

Stroke, the second leading cause of death and the third leading contributor to serious disability globally, has seen a significant rise in mortality and disability-adjusted life-years over recent decades [114]. According to the World Health Organization (WHO), between 1990 and 2010, stroke-related deaths increased by 26%, with disability-adjusted life-years rising by 19%. Annually, 15 million people worldwide suffer from stroke, resulting in 5 million deaths and leaving another 5 million permanently disabled [115]. This trend underscores stroke as a major public health challenge with substantial social and economic consequences [116]. In response to the urgent need for more effective diagnostic and predictive tools, researchers have turned to innovative approaches such as leveraging 1D-CNN models. These models offer the

TABLE 2. Existing works in music genre classification.

Author	Year	Method	Dataset/Classified genres	Results (Acc./F1)
Dieleman & Schrauwen [101]	2014	1D-CNN for improved frequency decomposition extraction and phase- and translation-invariant feature representations	MagnaTagATune (MTAT) dataset (50 classes of the total 188 in the dataset were considered)	83.87%
Ghosal & Kolekar [99]	2018	1DCNN-LSTM on Mel Spectrogram features	GTzan dataset (rock, reggae, blues, classical, disco, country, hip- hop, metal, jazz, and pop)	94.20%
Abdoli et al. [89]	2019	1D-CNN for feature learning from raw audio frame. The first layer was initialized as a Gammatone filter bank.	UrbanSound8k (Air conditioner, Car horn, Children playing, Dog bark, Drilling, Engine, idling, Gun shot, Jackhammer, Siren, Street music)	89.00%
Noman et al. [104]	2019	1D-CNN for heart sound classification	PhysioNet CinC challenge 2016 database (normal and abnormal)	86.34%
Allamy & Koerich [102]	2021	Residual 1D-CNN for raw feature learning	GTzan dataset (rock, reggae, blues, classical, disco, country, hip-hop, metal, jazz, and pop)	80.93%
Falola & Akinola [103]	2021	Three 1D-CNN layers	Custom Traditional Nigerian Music dataset (afro, apala, fuji and juju)	92.50%
Almazaydeh et al. [66]	2022	1DCNN with LSTM	Arabic Music Genre (eastern takht, rai, mawwal, muwashshah, the poem)	86.00%
Adesuyi et al. [105]	2022	1D-CNN for snoring sound classification	Custom snoring dataset (snoring and non-snoring (clock ticks, sleep-talk, yawning, pet snoring, door sound, bed spring sound))	99.70%

potential for lightweight yet powerful solutions to enhance stroke diagnosis and prediction, thereby improving patient care and outcomes. However, the complexity of stroke diagnosis and prediction demands continuous refinement and innovation in modeling techniques. Earlier studies have laid the groundwork for advancements in this field. For instance, the research in [117] proposed a basic 1D-CNN model for stroke prediction, which, although promising, faced limitations in terms of accuracy and generalizability due to its simplistic architecture. Building upon this foundation, Wang et al. [118] introduced a more sophisticated 1D-CNN architecture that addressed the shortcomings of the previous model by increasing the number of kernels in the convolutional layers, resulting in improved diagnostic accuracy and predictive performance. Despite these advancements, challenges remain in achieving optimal stroke diagnosis and prediction outcomes. Further research, such as that conducted by Elbagoury et al. [119], focused on integrating multimodal data sources and incorporating advanced machine learning algorithms to enhance the capabilities of 1D-CNN models in stroke prediction, which represent significant strides towards more accurate and reliable stroke diagnosis and prediction methods. Some existing works that have leveraged 1D-CNN for stroke detection are presented in Table 5.

F. SEIZURE DETECTION

Seizures, or abnormal electrical discharges, are the cause of epilepsy, a common chronic brain disease. Seizures of this nature may cause aberrant brain activity, knockout, severe wounds, and sometimes death. About 50 million people worldwide are diagnosed with epilepsy, with the biggest impact on children and adults aged 65–70 years [125]. Electroencephalography (EEG) is a widely used technique for the detection of epileptic seizure, and recently, researchers have leveraged 1D-CNN to detect seizures from highly complex EEG signals [126]. For instance, Hassan et al. [127] introduced a novel approach by integrating 1D-CNN with machine learning classifiers. Their methodology involved preprocessing EEG data using the Butterworth filter before employing 1D-CNN to extract pertinent features. However, challenges in effectively discerning relevant features from the extracted set hindered the initial model's performance,

TABLE 3. Existing works on electric motor fault detection.

Author	Year	Method	Dataset/Classified	Results (Acc/F1)
Khan et al. [112]	2018	1D-CNN and LSTM	Custom dataset (healthy and inter-turn fault in the stator)	92.00%
Mukhopadhyay et al. [107]	2018	Quasi 1D-CNN	Custom dataset (bearing, broker rotor bar, healthy, stator)	80.90%
Eren et al. [110]	2019	Compact Adaptive 1D-CNN	Intelligent Maintenance System (IMS) Bearing Dataset (healthy, inner race, outer race, roller element defects), Case Western Reserve University (CWRU) Bearing Dataset (inner race fault, ball fault, and three different outer race faults)	IMS - 93.90% CWRU – 93.20%
Chih-Cheng et al. [108]	2021	1D-CNN	(CWRU) Bearing Dataset (inner race fault, ball fault, and three different outer race faults)	99.20%
Ozcan et al. [111]	2022	Multilevel 1D-CNN	University of Cincinnati Intelligent Maintenance System dataset (healthy, inner race fault, rolling element fault, and early/advanced fault levels)	84.64%
Chuya-Samba et al. [113]	2022	1D-CNN	CWRU (inner race, outer race, rolling element), T-Y Wu (inner race, outer race, rolling element), NSF-IMS (inner race, outer race, rolling element)	99.52%, 99.31%, 99.64%

TABLE 4. Existing works on stroke detection.

Author	Year	Method	Dataset/Classified	Results (Acc/F1)
Hsieh et al. [120]	2020	Fixed kernel 1D-CNN	PhysioNet Challenge 2017 (atrial fibrillation, normal sinus rhythm, noisy, other)	78.20%
Dewi et al. [124]	2020	1D-CNN	National Brain Center Hospital Indonesia (normal, mild stroke, moderate stroke, severe stroke)	97.30%
Mamun [121]	2021	1D-CNN	PhysioNet (atrial fibrillation, normal sinus rhythm, noisy, other)	96.00%
Ho & Ding [122]	2022	1D-CNN with Gradient-weighted class activation mapping	PhysioNet (atrial fibrillation, normal sinus rhythm, noisy, other)	90.00%
Naufal et al. [123]	2022	1D-CNN	BSD dataset	98.00%
Wang et al. [118]	2023	1D-CNN	Custom dataset	90.53%

thereby impacting classification accuracy. In addressing the limitations observed in earlier endeavors and building upon the work, Salafian et al. [128] proposed an enhanced feature selection technique based on mutual information-based estimators. By incorporating this refined feature selection process into the 1D-CNN framework, the model demonstrated improved accuracy in seizure detection, addressing the shortcomings identified in previous studies. Emerging research efforts are exploring innovative approaches to further enhance the capabilities of 1D-CNN in seizure detection. This is exemplified by the work of Shanmugam and Dharmar [129], which presented an end-to-end automated seizure detection method based on deep learning, specifically utilizing a one-dimensional convolutional neural network-long short-term memory (1D-CNN-LSTM) model. The approach aimed to differentiate normal, ictal, and interictal EEG data without requiring extensive EEG data preprocessing or feature extraction. The results demonstrated impressive accuracy values, with the best model achieving 99–100% accuracy. Other works that have leveraged 1D-CNN for seizure detection are presented in Table 3.

G. WEATHER AND WIND SPEED PREDICTION

Weather prediction is a pivotal field that frequently leverages the advantages of big data and deep learning to enhance prediction accuracy. This collaboration between big data and deep learning enables more precise and timely forecasts, providing substantial benefits to various sectors heavily relying on weather information. Notably, the application of 1D CNN has gained prominence in recent years in weather forecasting due to its speed, efficiency, and accuracy. In data stream classification, Alex et al. [138] addressed the specific challenges of concept drift and class imbalance, with a particular focus on real-world weather data forecasting. The proposed framework utilized a self-organizing Auto-Encoder-based 1D-CNN model to enhance classification accuracy and stability while assessing concept drift handling through the Population Stability Index (PSI) on a Weather dataset, thereby demonstrating improved prediction performance and classifier stability. The hybrid model proposed by researchers in [139], combining 1D CNN and Bi-LSTM techniques, demonstrated enhanced prediction accuracy, laying the groundwork for more precise forecasts. In [140], various deep learning and traditional machine learning models were evaluated for multi-step ahead daily reference evapotranspiration (ETo) forecasting, with the MIMO approach using 1D CNN-LSTM, as shown in Fig. 19, demonstrating improved accuracy over baseline methods, particularly benefiting regions relying on historical monthly mean ETo data for forecasting. Additionally, studies such as that by Lawal et al. [141], which investigated short-term wind speed prediction using a hybrid approach, showcased the practical applicability of 1D CNN in real-world scenarios.. The study was conducted to predict short-term wind speed at different heights above ground level (AGL) using a hybrid approach combining 1D CNN and bidirectional LSTM (Bi-LSTM) networks. It demonstrated improved prediction accuracy with increasing AGL height and outperformed benchmark models, as seen in a real-world case study in Saudi Arabia. Transitioning to precipitation forecasting, research efforts have delved into the intricacies of rainfall prediction using 1D CNN models. Studies by [142] and [143] highlighted the significance of deep learning techniques in achieving commendable accuracy in forecasting rainfall and addressing complex time series challenges, respectively. The study in [142] aimed to forecast Indonesian rainfall over 14 days using 1D CNN models trained on data from BMKG weather observation stations. The approach involved data preprocessing steps such as interpolation, segmentation with overlap, and normalization, achieving a commendable testing accuracy of 81.46% with a minimal loss of 0.0018. On the other hand, the research by [143] employed a bidirectional LSTM model to capture temporal complexities and compare its efficacy with a 1D CNN model for addressing complex time series challenges. A distinct research [144] inquiry focused on conducting spatial and temporal analysis to forecast weekly rainfall patterns by leveraging climate data from three proximate BMKG stations spanning twelve years. An evaluation was carried out, contrasting the performance of a hybrid 2D CNN and RNN model against 1D CNN and 3D CNN models. The results demonstrated the superior predictive capability of the 2D CNN-RNN model, attributed to its enhanced modeling of spatial correlations in the dataset. 1D CNN has found applications in various fields beyond those mentioned above, including environmental forecasting and transportation management. Studies such as that by researchers [145] and [146] focused on air quality forecasting, leveraging 1D CNN and hybrid models to predict air pollution levels and hourly PM2.5 levels, respectively. These efforts demonstrate the versatility of 1D CNN in addressing diverse forecasting challenges and improving prediction accuracy in various domains. The study by [147] introduced a framework that employed two deep-learning models to predict traffic speed under varying conditions, including adverse weather. Notably, the 1D CNN demonstrated superior performance among the different models, emphasizing its potential in predicting spatiotemporal congestion due to adverse weather conditions. A summary of some existing research papers on windspeed prediction is presented in Table 6.

H. SPORTS

Integrating deep learning methodologies within sports analytics presents many opportunities for substantial advancements. These possibilities encompass a spectrum ranging from fundamental applications like predictive modeling of match outcomes to the more intricate and nuanced terrain of video analysis, fostering enhanced strategies for game planning. Deep learning techniques enable intricate tasks, including real-time identification and classification of in-game events, the nuanced recognition of human motion patterns, predictive models for injury occurrence [148], and fine-grained action classification. In sports analytics, recognizing human activities or events is a challenging field for computer vision due to the complex and coarse nature of the data. Researchers strive to develop algorithms that can extract meaningful insights from dynamic and unstructured visual data in sports events, requiring innovative solutions for accurate recognition.

Sports activity recognition relies heavily on data acquisition from sensors and RGB video sources. Initially, research focused on utilizing data from smart wearable devices to recognize activities. Zhang [149] pioneered using a

TABLE 5. Summary of some research studies on Seizure detection.

Author	Year	Method	Dataset/Classified	Results (Acc/F1)
Ullah et al. [134]	2018	Pyramidal 1D-CNN	University of Bonn dataset (normal, interictal, ictal)	96.10%
Chowdhury et al. [136]	2019	1D-CNN	University of Bonn dataset (normal, interictal, ictal)	98.50%
Xu et al. [131]	2020	1D-CNN-LSTM	UCI Epileptic Seizure (epileptic seizure, first normal, second normal, third normal, and fourth normal)	82.00%
Sharan & Berkovsky [135]	2020	Multichannel wavelet power spectra and 1D-CNN	CHB-MIT	97.25%
Zhu et al. [132]	2021	1D-CNN	CHB-MIT (clonic seizure, atonic seizure, tonic seizure)	97.35%
Wang et al. [133]	2021	Stacked 1D-CNN with random selection and data augmentation	CHB-MIT and SWEC-ETHZ	CHB-MIT - 90.00% SWEC-ETHZ - 97.52%
Prawiti et al. [130]	2022	1D-CNN	Custom dataset	97.01%
Ra et al. [137]	2023	Synchro-extracting transform and 1D-CNN.	CHB-MIT (clonic seizure, atonic seizure, tonic seizure) and University of Bonn dataset (normal, interictal, ictal)	CHB-MIT -99.71% Bonn - 100.00%

hybrid model combining 1D CNN with LSTM, achieving a significant accuracy of 89.8% in recognizing various activities. Building upon this, Waghchaware and Joshi [150] further improved the accuracy by employing a 1D CNN-LSTM architecture with sensor data from smartwatches and smartphones. The developed model had an impressive accuracy of 97% in recognizing six specific activities. However, the reliance on wearable devices limited the scope of data collection and presented challenges in real-world implementation. To address these limitations, Srimath et al. [151] introduced a novel approach using 2D skeleton key points extracted from RGB video frames. Their 1D-CNN model surpassed previous methods, offering enhanced accuracy in human activity recognition. This shift from wearable devices to video-based data acquisition expanded the potential for real-world applications and improved the robustness of activity recognition systems. Furthermore, Vats et al. [152] introduced a novel multi-tower temporal 1D-CNN architecture tailored for event detection within the context of ice hockey, with each tower representing a temporal 1D CNN possessing distinct receptive fields.

Deep learning techniques have also significantly improved athlete health monitoring and injury prediction. Guo et al. [153] demonstrated the potential of wearable devices and data analysis in assessing physical fitness levels among teenagers. By employing a 1D CNN-LSTM model, their model achieved remarkable accuracy in predicting fitness levels, highlighting the effectiveness of deep learning in evaluating adolescent health (98.27% accuracy for boys and 99.26% accuracy for girls). Expanding on this research, Hui et al. [154] introduced a wearable device for real-time heart rhythm monitoring coupled with a 1D CNN-based classification approach. Their work not only improved accuracy but also reduced network complexity compared to existing algorithms, offering promising prospects for enhanced arrhythmia detection in athletes. These advancements underscore the transformative potential of deep learning in promoting athlete well-being and injury prevention.

Additionally, in the context of multiple object tracking, a 1D CNN-based solution effectively mitigated gaps in tracking data, enhancing tracking accuracy compared to linear interpolation techniques [155]. These studies underscore the substantial potential of deep learning within sports analytics, encompassing a spectrum of applications from activity recognition and health monitoring to action classification and object tracking, thus expanding the horizons of our understanding and performance enhancement in sports.

In addition to activity recognition and health monitoring, 1D CNNs have been found to be useful across various dimensions of the sports domain. Highlight detection, fine-grained action classification, and object tracking represent some of the diverse applications where 1D CNNs have demonstrated

TABLE 6. Existing works on windspeed prediction.

Author	Year	Method	Dataset	Result (Acc/F1)
Patel et al. [143]	2018	Bidirectional LSTM model to capture temporal complexities and comparing its efficacy with a 1D CNN model for addressing complex time series challenges	Indian weather pattern: available for Ahmedabad airport specifically, from 2010 to 2017 Historical Hourly Weather Data 2012-2017	Indian data 1D CNN Avg. Accuracy: 84.60 NA data 1D CNN Avg. Accuracy: 89.18%
Fu et al. [139]	2019	Hybrid 1D CNN and Bi-LSTM model	OBS, RMAPS	Avg. RMSE: 0.4541
Shabarek et al. [147]	2020	1D CNN with other deep learning models for comparison	Probe-vehicle database (INRIX) New Jersey Congestion Management System (NJCMS) New Jersey Straight Line Diagram (NJSLD) DarkSky	Avg RMSE Rain: 5.4 Fog: 5.5 Snow: 5.8
Ferreira & da Cunha [140]	2020	1D CNN-LSTM	Brazilian National Institute of Meteorology (INMET)	AVG. RMSE: 0.67
Sari et al. [142]	2020	1D CNN	BMKG	Accuracy: 81.46%
Ragab et al. [145]	2020	1D CNN and Exponential Adaptive Gradients (EAG) optimization to predict the Air Pollution Index (API)	Collected from the Department of the Environment, Malaysia	RMSE: 2.354
Lawal et al. [141]	2021	Hybrid 1D CNN and Bi-LSTM model	Wind data collection station in Saudi Arabia	RMSE: 0.4295
Lestari & Djamal [144]	2022	Comparing the performance of a hybrid 2D CNN and RNN model against 1D CNN and 3D CNN models	BMKG	Accuracy: 80.21%
Alex et al. [138]	2022	A self-organizing Auto-Encoder-based 1D-CNN model	Collected from U.S.National Oceanic and Atmospheric Administration (NOAA)	Accuracy: 83.60%
Zhu & Xie [146]	2023	Parallel multi-input 1D-CNN-BiLSTM model	Collected from the Chinese Ministry of Environmental Protection	RMSE: 3.88

efficacy. For instance, in highlight detection [156], 1D-CNN models accurately estimated blink rates, closely aligning with human perception. Furthermore, in fine-grained action classification for table tennis, a three-stream network incorporating RGB data, optical flow, and player pose information yielded notable improvements in convergence speed and performance, achieved through combining 3D and 1D CNNs [156]. Padel tennis benefited from a pioneering approach that created the inaugural Padel shot dataset and employed a 1D CNN, achieving an impressive 93% shot classification accuracy [157]. A summary of some existing research papers and their details is presented in Table 7.

I. OTHERS

Other domains that have leveraged 1D-CNN to achieve state-of-the-art are discussed in this section. For example, Kim et al. [159] classified multiclass missiles using 1D-CNN. Several missile shapes with the same trajectories and the same missile shapes with various trajectories were experimented with, and the researchers combined 1D-CNN with GRU for classification. The results were compared with a standalone 1D-CNN model and a 1D-CNN-LSTM model, with the 1D-CNN-GRU model achieving a classification accuracy of 99.40%. In depression detection, several researchers such as Lin et al. [160], Devika et al. [161], and others have achieved state-of-the-art using 1D-CNN. In Abo-Tabik et al. [162], 1D-CNN was also used for smoke detection, where the authors combined a Control Theory model of smoking with a 1D-CNN classifier to adapt to individual differences between smokers and predict smoking events based on motion and geolocation values collected using mobile devices. Olmedilla et al. [163] used 1D-CNN to assess the early prediction and the modeling of online review helpfulness to account for online reviews' textual characteristics. Also, Wang et al. [164] designed ECA-Net, a neural network architecture designed to improve

TABLE 7. Some existing research papers in sport domain.

Author	Year	Method	Dataset	Results (Acc/F1)
Vats et al. [152]	2020	Novel multi-tower temporal 1D-CNN architecture tailored for event detection	NHL dataset, Soccer-Net Dataset	55.00%
Rana et al. [155]	2020	1D CNN to mitigate the complexity of multiple object tracking	Netball & hockey dataset (self-collected)	84.60%
Nakano et al. [156]	2020	1D CNN to extract spatio-temporal pose features from video frames	Olympics or formal international competition (self-collected)	94%
Srimath et al. [151]	2021	1D CNN model to extract features from 2D skeleton points generated from RGB videos using OpenPose API	UTD-MHAD, KTH and UCF-Sports	94.32%, 97.23%, 98.02%
Hui et al. [154]	2021	1D CNN used for low complexity and high accuracy as classification	ECG (self-collected)	99.83%
Zhang [149]	2021	Hybrid 1D CNN-LSTM model	Smart wearable device data (self-collected)	89.80%
Martin et al. [158]	2021	Three-stream network, incorporating RGB data, optical flow, and player pose information with attention blocks and bilinear layers, resulting in improved convergence speed and performance achieved through a combination of 3D and 1D CNN	TTStroke-21 Dataset	86.50%
Guo et al. [153]	2022	1D-CNN with LSTM model for classification	PPG recordings (self-collected)	98.27%
Waghchaware and Josi [150]	2022	1D CNN-LSTM architecture	WISDM	97.00%
Cartes [157]	2023	1D CNN used in comparison with other ML algorithms for evaluation	First padel shot dataset (self-collected)	93.00%

the efficiency of channel attention mechanisms using 1D-CNN. The ECA-Net was proposed to address the limitations of traditional channel attention mechanisms, such as the Squeeze-and-Excitation (SE) block [165]. The primary focus of ECA-Net is to enhance the modeling of interdependencies between channels in a computationally efficient manner. In foot type classification, Mei et al. [166] collected data through sensor-enabled insoles and processed used 1D-CNN to distinguish normal, cavus, and planus feet. To determine which sensor modalities best represented different foot types, the authors experimented with a number of combinations, and results showed that combining angular velocity and force sensing produced a maximum classification accuracy of 99.26%. Several authors have also leveraged 1D-CNN for moving vehicle identification and driver behaviour detection, as seen in [167], [168], and [169], among many others. However, there is still room for improvement, as most of these architectures have several limitations with respect to their applicable domains.

V. CHALLENGES AND FUTURE DIRECTIONS

Even though 1D-CNN has achieved state-of-the-art across several fields, some challenges still limit the advancements of 1D-CNN across applicable domains. This section presents a brief discussion of these challenges, and the future directions researchers can focus on.

A. ENHANCING GLOBAL CONTEXT UNDERSTANDING

1D-CNNs, while powerful for processing sequential data, face specific challenges in capturing long-range dependencies and global contextual information. Unlike 2D-CNNs used in image processing, where the spatial relationships between features are more easily encoded, 1D-CNNs can struggle with understanding the broader context in time-series or sequential data. This limitation arises because 1D-CNNs typically apply convolutional filters uniformly across the input sequence, potentially overlooking important dependencies that span larger segments of the data. To address this limitation, future research endeavors should leverage attention mechanisms and integrate "plug-and-play" modules. These additions can dynamically assign varying levels of importance to different parts of the sequence, thereby enhancing the model's ability to capture long-range dependencies and global context while emphasizing the significant features. This approach holds promise in advancing the efficacy of 1D-CNNs across various tasks.

B. INTERPRETABILITY AND EXPLAINABILITY

Understanding the decision-making process of 1D-CNNs can be complex, which can be a major drawback, especially in critical applications like healthcare, finance, and security, where understanding the rationale behind a model's decision is essential. Future research should prioritize the development of methodologies and frameworks to enhance the interpretability and explainability of 1D-CNN models. Techniques such as saliency maps, layer-wise relevance propagation, and Grad-CAM (Gradient-weighted Class Activation Mapping) can be adapted and refined for 1D-CNNs to provide visual or numerical explanations of the features and segments of the input data that most influence the model's decisions. Additionally, integrating these explainability techniques with attention mechanisms can offer more detailed insights into the model's focus and reasoning process. By improving interpretability, researchers can not only build trust and transparency in 1D-CNN applications but also gain deeper insights into model behavior, leading to more robust and reliable systems across various domains. This approach will facilitate better validation and debugging of models, ultimately advancing the practical applicability and acceptance of 1D-CNNs in real-world scenarios.

C. SENSITIVITY TO INPUT VARIABILITY

1D-CNNs are often sensitive to variations in input sequences, such as noise, signal distortions, or varying sequence lengths. This sensitivity can lead to decreased performance and reliability in real-world applications, where data often comes with inherent variability and imperfections. Ensuring robustness against such variations is crucial for deploying 1D-CNNs in practical settings.

Future research should focus on developing more robust architectures and training strategies that can effectively handle input signal variations. One promising approach is to incorporate uncertainty modeling techniques, such as Bayesian neural networks or dropout-based uncertainty estimation. By treating the weights and biases of the network as probability distributions instead of fixed parameters, these techniques can help the model to account for and adapt to the uncertainty and variability in the input data. Additionally, advancements in data augmentation techniques specific to 1D data, such as time warping, adding noise, or generating synthetic data, can help in training models that are more resilient to input variations. Regularization techniques and ensemble methods can also be explored to enhance robustness. By focusing on these strategies, researchers can develop 1D-CNN models that are not only more resilient to noise and variability but also more reliable and effective in real-world applications across various domains, including healthcare, finance, and communications.

D. DATA EFFICIENCY CHALLENGES

Generally, deep learning models require substantial data, which are sometimes tedious to label. For instance, an accelerometer or gyroscope signal collected for human activity recognition will require an expert to be present during data collection and be fully aware of the time a particular subject is carrying out a specific activity, including transitions. Also, sensors are sometimes prone to errors and malfunctions, affecting the data collection quality.

To address these challenges, future research should explore the development and integration of state-of-the-art generative and unsupervised models within 1D-CNN frameworks. Techniques such as Generative Adversarial Networks or Variational Autoencoders can be employed to create high-quality synthetic data that mimics the characteristics of real one-dimensional data. This synthetic data can then be used to augment existing datasets, alleviating the burden of manual data collection and labeling. By focusing on these approaches, researchers can significantly reduce the data requirements for training 1D-CNNs and improve the robustness and generalization capabilities of the models. This will facilitate the application of 1D-CNNs in diverse fields where obtaining large labeled datasets is challenging.

E. LIGHTWEIGHT ARCHITECTURES

While 1D-CNNs generally have less model complexity compared to their 2D and 3D counterparts, they can become parameter-heavy when multiple layers are stacked, and with increasing input data size. This can be particularly problematic for deployment on end devices with limited resources, such as CPUs, battery life, and memory capacity. In many practical applications, such as wearable devices for health monitoring or mobile phones for real-time signal processing, resource constraints are a critical consideration. Moreover, in scenarios where real-time processing and low-latency responses are essential, such as real-time health monitoring or anomaly detection in IoT systems, lightweight 1D-CNN architectures can significantly improve inference times, enabling faster and more responsive systems. By focusing on these approaches, researchers can develop sophisticated 1D-CNN models that are both lightweight and powerful, facilitating their deployment on resource-constrained end devices and ensuring efficient real-time processing. This will expand the applicability of 1D-CNNs to a wider range of practical and real-world applications where computational resources are limited.

VI. CONCLUSION

Despite the success of 2D and 3D CNNs in achieving state-of-the-art performance across diverse areas, their limitations in capturing one-dimensional signals led to the evolution of 1D-CNNs. These one-dimensional architectures have addressed the specific requirements of time series prediction and signal identification and paved the way for sophisticated models applicable across a wide spectrum of research fields. Generally, 1D-CNN has evolved and aided the development of various state-of-the-art models that cut across numerous research fields, and it is gaining more interest among researchers. They are particularly useful when dealing with sequential data where local patterns and dependencies matter. However, there has been no survey paper detailing the evolution and advancements in the applications of the 1D-CNN to several computer vision tasks. This survey paper has comprehensively explored the evolution, methodologies, and applications of 1D-CNN in domains where capturing temporal dynamics and dependencies is paramount. Through an in-depth exploration of the general overview and approaches employed in 1D-CNNs, this survey has illuminated the versatility and adaptability of these networks. Furthermore, it has delved into recent advancements in the applications of 1D-CNNs, spanning over twelve distinct domains. Also, by presenting the grand research areas that remain unexplored, this survey will inspire further investigations on the application of 1D-CNN by researchers across several fields.

REFERENCES

[1] S. Haggenmüller et al., "Skin cancer classification via convolutional neural networks: Systematic review of studies involving human experts," *Eur. J. Cancer*, vol. 156, pp. 202–216, Oct. 2021, doi: 10.1016/j.ejca.2021.06.049.

[2] G. Rjoub, O. A. Wahab, J. Bentahar, and A. S. Bataineh, "Improving autonomous vehicles safety in snow weather using federated YOLO CNN learning," in *Proc. 17th Int. Conf. Mobile Web Intell. Inf. Syst. (MobiWIS)*, Aug. 2021, pp. 121–134.

[3] X. Sang, L. Xue, X. Ran, X. Li, J. Liu, and Z. Liu, "Intelligent high-resolution geological mapping based on SLIC-CNN," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 2, p. 99, Feb. 2020.

[4] J. Delanoy, M. Aubry, P. Isola, A. A. Efros, and A. Bousseau, "3D sketching using multi-view deep volumetric prediction," *Proc. ACM Comput. Graph. Interact. Techn.*, vol. 1, no. 1, pp. 1–22, Jul. 2018.

[5] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.

[6] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," 2023, *arXiv:2304.02643*.

[7] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[8] W. Mcculloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biol.*, vol. 52, nos. 1–2, pp. 99–115, 1990.

[9] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958.

[10] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989, doi: 10.1162/neco.1989.1.4.541.

[11] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 1, pp. 23–38, Jan. 1998, doi: 10.1109/34.655647.

[12] P. Juell and R. Marsh, "A hierarchical neural network for human face detection," *Pattern Recognit.*, vol. 29, no. 5, pp. 781–787, May 1996, doi: 10.1016/0031-3203(95)00129-8.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, pp. 1097–1105, 2012, doi: 10.1201/9781420010749.

[14] T. Uchiyama, N. Sogi, K. Niinuma, and K. Fukui, "Visually explaining 3D-CNN predictions for video classification with an adaptive occlusion sensitivity analysis," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2023, pp. 1513–1522.

[15] D. Maturana and S. Scherer, "VoxNet: A 3D convolutional neural network for real-time object recognition," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 922–928, doi: 10.1109/IROS.2015.7353481.

[16] S. P. Singh, L. Wang, S. Gupta, B. Gulyás, and P. Padmanabhan, "Shallow 3D CNN for detecting acute brain hemorrhage from medical imaging sensors," *IEEE Sensors J.*, vol. 21, no. 13, pp. 14290–14299, Jul. 2021, doi: 10.1109/JSEN.2020.3023471.

[17] S. Kiranyaz, T. Ince, O. Abdeljaber, O. Avci, and M. Gabbouj, "1-D convolutional neural networks for signal processing applications," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 8360–8364, doi: 10.1109/ICASSP.2019.8682194.

[18] A. O. Ige and M. H. M. Noor, "A deep local-temporal architecture with attention for lightweight human activity recognition," *Appl. Soft Comput.*, vol. 149, Dec. 2023, Art. no. 110954, doi: 10.1016/j.asoc.2023.110954.

[19] R. F. R. Junior, I. A. D. S. Areias, M. M. Campos, C. E. Teixeira, L. E. B. da Silva, and G. F. Gomes, "Fault detection and diagnosis in electric motors using 1D convolutional neural networks with multi-channel vibration signals," *Measurement*, vol. 190, Feb. 2022, Art. no. 110759, doi: 10.1016/j.measurement.2022.110759.

[20] D. Moitra and R. K. Mandal, "Classification of non-small cell lung cancer using one-dimensional convolutional neural network," *Expert Syst. Appl.*, vol. 159, Nov. 2020, Art. no. 113564, doi: 10.1016/j.eswa.2020.113564.

[21] P. Bharati and A. Pramanik, "Deep learning techniques—R-CNN to mask R-CNN: A survey," in *Computational Intelligence in Pattern Recognition*, A. K. Das, J. Nayak, B. Naik, S. K. Pati, and D. Pelusi, Eds., Singapore: Springer, 2020, pp. 657–668.

[22] D. A. Bashar, "Survey on evolving deep learning neural network architectures," *J. Artif. Intell. Capsule Networks*, vol. 2019, no. 2, pp. 73–82, Dec. 2019, doi: 10.36548/jaicn.2019.2.003.

[23] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, Art. no. 107398, doi: 10.1016/j.ymssp.2020.107398.

[24] M. Sahu and R. Dash, "A survey on deep learning: Convolution neural network (CNN)," in *Smart Innovation, Systems and Technologies*, D. Mishra, R. Buyya, P. Mohapatra, and S. Patnaik, Eds., Singapore: Springer, 2021, pp. 317–325.

[25] G. Habib and S. Qureshi, "Optimization and acceleration of convolutional neural networks: A survey," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 7, pp. 4244–4268, Jul. 2022, doi: 10.1016/j.jksuci.2020.10.004.

[26] J.-X. Mi, J. Feng, and K.-Y. Huang, "Designing efficient convolutional neural network structure: A survey," *Neurocomputing*, vol. 489, pp. 139–156, Jun. 2022, doi: 10.1016/j.neucom.2021.08.158.

[27] R. Tombe and S. Viriri, "Effective processing of convolutional neural networks for computer vision: A tutorial and survey," *IETE Tech. Rev.*, vol. 39, no. 1, pp. 49–62, Jan. 2022, doi: 10.1080/02564602.2020.1823252.

[28] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," *Mobile Netw. Appl.*, vol. 25, no. 2, pp. 743–755, Apr. 2020, doi: 10.1007/s11036-019-01445-x.

[29] S. Verma. (2019). *Understanding 1D and 3D Convolution Neural Network*. Keras. Accessed: Mar. 22, 2023. [Online]. Available: https://towardsdatascience.com/understanding-1d-and-3d-convolution-neural-network-keras-9d8f76e29610

[30] A. O. Akinwumi, A. O. Akingbesote, O. O. Ajayi, and F. O. Aranuwa, "Detection of distributed denial of service (DDoS) attacks using convolutional neural networks," *Nigerian J. Technol.*, vol. 41, no. 6, pp. 1017–1024, Mar. 2023.

[31] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[32] S. Poudyal, M. J. Mohammadi-Aragh, and J. E. Ball, "Prediction of student academic performance using a hybrid 2D CNN model," *Electronics*, vol. 11, no. 7, p. 1005, Mar. 2022.

[33] O. Yildirim, M. Talo, B. Ay, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals," *Comput. Biol. Med.*, vol. 113, Oct. 2019, Art. no. 103387, doi: 10.1016/j.compbiomed.2019.103387.

[34] A. O. Ige, N. K. Tomar, F. O. Aranuwa, O. Oriola, A. O. Akingbesote, M. H. M. Noor, M. Mazzara, and B. S. Aribisala, "ConvSegNet: Automated polyp segmentation from colonoscopy using context feature refinement with multiple convolutional kernel sizes," *IEEE Access*, vol. 11, pp. 16142–16155, 2023, doi: 10.1109/ACCESS.2023.3244789.

[35] M. R. Goni, A. O. Ige, and N. I. R. Ruhaiyem, "TL-AttSharpNet: Automated lung image segmentation using transfer learning with depthwise convolution and attention," in *Proc. IEEE 2nd Nat. Biomed. Eng. Conf. (NBEC)*, Sep. 2023, pp. 133–137, doi: 10.1109/nbec58134.2023.10352617.

[36] M. R. Goni and N. I. R. Ruhaiyem, "Salient feature extraction using attention for brain tumor segmentation," in *Proc. IEEE Int. Conf. Comput. (ICOCO)*, Nov. 2022, pp. 305–309, doi: 10.1109/ICOCO56118.2022.10031677.

[37] M. R. Goni, N. I. R. Ruhaiyem, M. Mustapha, A. Achuthan, and C. M. N. C. M. Nassir, "Brain vessel segmentation using deep learning—A review," *IEEE Access*, vol. 10, pp. 111322–111336, 2022, doi: 10.1109/ACCESS.2022.3214987.

[38] P. A. Prabha, M. D. Priya, A. C. J. Malar, S. Karthik, G. Dakshin, and S. D. Kumar, "Improved ResNet-based image classification technique for malaria detection," in *Proc. 6th Int. Conf. Recent Trends Comput.*, in Lecture Notes in Networks and Systems. Singapore: Springer, 2021, pp. 795–803, doi: 10.1007/978-981-33-4501-0_73.

[39] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017, doi: 10.1016/j.media.2016.10.004.

[40] A. Rehman, M. A. Khan, T. Saba, Z. Mehmood, U. Tariq, and N. Ayesha, "Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture," *Microsc. Res. Technique*, vol. 84, no. 1, pp. 133–149, Jan. 2021, doi: 10.1002/jemt.23597.

[41] L. Yuan, X. Wei, H. Shen, L.-L. Zeng, and D. Hu, "Multi-center brain imaging classification using a novel 3D CNN approach," *IEEE Access*, vol. 6, pp. 49925–49934, 2018, doi: 10.1109/ACCESS.2018.2868813.

[42] A. S. Neethi, S. Niyas, S. K. Kannath, J. Mathew, A. M. Anzar, and J. Rajan, "Stroke classification from computed tomography scans using 3D convolutional neural network," *Biomed. Signal Process. Control*, vol. 76, Jul. 2022, Art. no. 103720, doi: 10.1016/j.bspc.2022.103720.

[43] C. K. Vishnudas and G. R. Gnana King, "Computer-aided diagnosis for early detection and staging of human pancreatic tumors using an optimized 3D CNN on computed tomography," *Multimedia Syst.*, vol. 29, no. 5, pp. 2689–2703, Oct. 2023, doi: 10.1007/s00530-023-01146-2.

[44] Q. Zhang, H. Wang, S. W. Yoon, D. Won, and K. Srihari, "Lung nodule diagnosis on 3D computed tomography images using deep convolutional neural networks," *Proc. Manuf.*, vol. 39, pp. 363–370, Jan. 2019, doi: 10.1016/j.promfg.2020.01.375.

[45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.

[46] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[47] CR Formatted R. Dogaru and L. O. Chua, "CNN genes for one-dimensional cellular automata: A multi-nested piecewise-linear approach," *Int. J. Bifurcation Chaos*, vol. 8, no. 10, pp. 1987–2001, Oct. 1998, doi: 10.1142/s0218127498001649.

[48] S. Wang, J. Peng, J. Ma, and J. Xu, "Protein secondary structure prediction using deep convolutional neural fields," *Sci. Rep.*, vol. 6, no. 1, p. 18962, Jan. 2016, doi: 10.1038/srep18962.

[49] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, "Convolutional neural networks for patient-specific ECG classification," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2608–2611, doi: 10.1109/EMBC.2015.7318926.

[50] Y.-H. Cheng and C.-N. Kuo, "Machine learning for music genre classification using visual mel spectrum," *Mathematics*, vol. 10, no. 23, p. 4427, Nov. 2022, doi: 10.3390/math10234427.

[51] M. H. Chan and M. H. M. Noor, "A unified generative model using generative adversarial network for activity recognition," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 7, pp. 8119–8128, Jul. 2021, doi: 10.1007/s12652-020-02548-0.

[52] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for neural networks for image processing," 2015, *arXiv:1511.08861*.

[53] D. S. K. Karunasingha, "Root mean square error or mean absolute error? Use their ratio as well," *Inf. Sci.*, vol. 585, pp. 609–629, Mar. 2022.

[54] . U. R. Dr. A, "Binary cross entropy with deep learning technique for image classification," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 5393–5397, Aug. 2020.

[55] S. Ullah, J. Ahmad, M. A. Khan, E. H. Alkhammash, M. Hadjouni, Y. Y. Ghadi, F. Saeed, and N. Pitropakis, "A new intrusion detection system for the Internet of Things via deep convolutional neural network and feature engineering," *Sensors*, vol. 22, no. 10, p. 3607, May 2022, doi: 10.3390/s22103607.

[56] W. Hong, M. Chen, P. Gao, and D. Hong, "Medium- and long-term orbit prediction of satellite based on LSTNet," in *Proc. 15th Int. Conf. Comput. Autom. Eng. (ICCAE)*, Mar. 2023, pp. 534–540, doi: 10.1109/ICCAE56788.2023.10111400.

[57] R. Orjesek, R. Jarina, and M. Chmulik, "End-to-end music emotion variation detection using iteratively reconstructed deep features," *Multimedia Tools Appl.*, vol. 81, no. 4, pp. 5017–5031, Feb. 2022, doi: 10.1007/s11042-021-11584-7.

[58] S. Soni, S. S. Chouhan, and S. S. Rathore, "TextConvoNet: A convolutional neural network based architecture for text classification," *Int. J. Speech Technol.*, vol. 53, no. 11, pp. 14249–14268, Jun. 2023, doi: 10.1007/s10489-022-04221-9.

[59] K. Dashtipour, C. Ieracitano, F. C. Morabito, A. Raza, and A. Hussain, "An ensemble based classification approach for Persian sentiment analysis," in *Progresses in Artificial Intelligence and Neural Systems*, A. Esposito, M. Faundez-Zanuy, F. C. Morabito, and E. Pasero, Eds., Paser, Singapore: Springer, 2021, pp. 207–215, doi: 10.1007/978-981-15-5093-5_20.

[60] V. Kocaman and D. Talby, "Biomedical named entity recognition at scale," in *Proc. Int. Conf. Pattern Recognit.*, Jan. 2021, pp. 635–646.

[61] J. A. Jacob, J. P. Cherian, J. George, C. D. Reji, and D. Mohan, "Heart diseases classification using 1D CNN," in *Emerging Technologies in Data Mining and Information Security*, A. E. Hassanien, S. Bhattacharyya, S. Chakrabati, A. Bhattacharya, and S. Dutta, Eds., Singapore: Springer, 2021, pp. 755–765.

[62] S. Cavalli and M. Amoretti, "CNN-based multivariate data analysis for Bitcoin trend prediction," *Appl. Soft Comput.*, vol. 101, Mar. 2021, Art. no. 107065, doi: 10.1016/j.asoc.2020.107065.

[63] X. Liu, K. Yan, L. B. Kara, and Z. Nie, "CCFD-Net: A novel deep learning model for credit card fraud detection," in *Proc. IEEE 22nd Int. Conf. Inf. Reuse Integr. Data Sci. (IRI)*, Aug. 2021, pp. 9–16, doi: 10.1109/IRI51335.2021.00008.

[64] C. C. Hung, Y. J. Chen, S. J. Guo, and F. C. Hsu, "Predicting the price movement from candlestick charts: A CNN-based approach," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 34, no. 2, pp. 111–120, 2020, doi: 10.1504/ijahuc.2020.107821.

[65] A. Yadav and D. K. Vishwakarma, "A multilingual framework of CNN and bi-LSTM for emotion classification," in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2020, pp. 1–6, doi: 10.1109/ICCCNT49239.2020.9225614.

[66] L. Almazaydeh, S. Atiewi, A. Al Tawil, and K. Elleithy, "Arabic music genre classification using deep convolutional neural networks (CNNs)," *Comput., Mater. Continua*, vol. 72, no. 3, pp. 5443–5458, 2022.

[67] Mustaqeem and S. Kwon, "MLT-DNet: Speech emotion recognition using 1D dilated CNN based on multi-learning trick approach," *Expert Syst. Appl.*, vol. 167, Apr. 2021, Art. no. 114177, doi: 10.1016/j.eswa.2020.114177.

[68] A. O. Ige and M. H. M. Noor, "A survey on unsupervised learning for wearable sensor-based activity recognition," *Appl. Soft Comput.*, vol. 127, Sep. 2022, Art. no. 109363, doi: 10.1016/j.asoc.2022.109363.

[69] A. O. Ige and M. H. M. Noor, "A lightweight deep learning with feature weighting for activity recognition," *Comput. Intell.*, vol. 39, no. 2, pp. 315–343, Apr. 2023, doi: 10.1111/coin.12565.

[70] M. G. Ragab, S. J. Abdulkadir, and N. Aziz, "Random search one dimensional CNN for human activity recognition," in *Proc. Int. Conf. Comput. Intell. (ICCI)*, Oct. 2020, pp. 86–91, doi: 10.1109/ICCI51257.2020.9247810.

[71] K. Banjarey, S. P. Sahu, and D. K. Dewangan, "Human activity recognition using 1D convolutional neural network," in *Advances in Intelligent Systems and Computing*, S. Shakya, V. E. Balas, S. Kamolphiwong, and K.-L. Du, Eds., Singapore: Springer, 2022, pp. 691–702.

[72] Z. N. Khan and J. Ahmad, "Attention induced multi-head convolutional neural network for human activity recognition," *Appl. Soft Comput.*, vol. 110, Oct. 2021, Art. no. 107671, doi: 10.1016/j.asoc.2021.107671.

[73] J. X. Goh, K. M. Lim, and C. P. Lee, "1D convolutional neural network with long short-term memory for human activity recognition," in *Proc. IEEE Int. Conf. Artif. Intell. Eng. Technol. (IICAIET)*, Sep. 2021, pp. 1–6, doi: 10.1109/IICAIET51634.2021.9573979.

[74] Y. J. Luwe, C. P. Lee, and K. M. Lim, "Wearable sensor-based human activity recognition with hybrid deep learning model," *Informatics*, vol. 9, no. 3, p. 56, Jul. 2022, doi: 10.3390/informatics9030056.

[75] M. Abdel-Basset, H. Hawash, R. K. Chakrabortty, M. Ryan, M. Elhoseny, and H. Song, "ST-DeepHAR: Deep learning model for human activity recognition in IoHT applications," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4969–4979, Mar. 2021, doi: 10.1109/JIOT.2020.3033430.

[76] M. H. M. Noor, S. Y. Tan, and M. N. A. Wahab, "Deep temporal conv-LSTM for activity recognition," *Neural Process. Lett.*, vol. 54, no. 5, pp. 4027–4049, Mar. 2022, doi: 10.1007/s11063-022-10799-5.

[77] M. H. M. Noor, "Feature learning using convolutional denoising autoencoder for activity recognition," *Neural Comput. Appl.*, vol. 33, no. 17, pp. 10909–10922, Jan. 2021, doi: 10.1007/s00521-020-05638-4.

[78] A. O. Ige and M. H. M. Noor, "Unsupervised feature learning in activity recognition using convolutional denoising autoencoders with squeeze and excitation networks," in *Proc. 5th Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Aug. 2022, pp. 435–440, doi: 10.1109/ICOIACT55506.2022.9972095.

[79] D. Bhattacharya, D. Sharma, W. Kim, M. F. Ijaz, and P. K. Singh, "Ensem-HAR: An ensemble deep learning model for smartphone sensor-based human activity recognition for measurement of elderly health monitoring," *Biosensors*, vol. 12, no. 6, p. 393, Jun. 2022, doi: 10.3390/bios12060393.

[80] A. O. Ige and M. H. M. Noor, "WSense: A robust feature learning module for lightweight human activity recognition," 2023, *arXiv:2303.17845*.

[81] F. Alharbi, L. Ouarbya, and J. A. Ward, "Synthetic sensor data for human activity recognition," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2020, pp. 1–9, doi: 10.1109/IJCNN48605.2020.9206624.

[82] J. J. P. Suarez, N. Orillaza, and P. Naval, "AFAR: A real-time vision-based activity monitoring and fall detection framework using 1D convolutional neural networks," in *Proc. ACM Int. Conf. Ser.*, Jun. 2022, pp. 555–559, doi: 10.1145/3529836.3529862.

[83] P. Gupta and N. Rajput, "Two-stream emotion recognition for call center monitoring," in *Proc. 8th Annu. Conf. Int. Speech Commun. Assoc.*, Aug. 2007, pp. 2241–2244.

[84] R. Rana, S. Latif, R. Gururajan, A. Gray, G. Mackenzie, G. Humphris, and J. Dunn, "Automated screening for distress: A perspective for the future," *Eur. J. Cancer Care*, vol. 28, no. 4, Jul. 2019, Art. no. e13033.

[85] B. Yalamanchili, S. K. Samayamantula, and K. R. Anne, "Neural network-based blended ensemble learning for speech emotion recognition," *Multidimensional Syst. Signal Process.*, vol. 33, no. 4, pp. 1323–1348, Dec. 2022, doi: 10.1007/s11045-022-00845-9.

[86] M. Chourasia, S. Haral, S. Bhatkar, and S. Kulkarni, "Emotion recognition from speech signal using deep learning," in *Intelligent Data Communication Technologies and Internet of Things*, J. Hemanth, R. Bestak, and J. I.-Z. Chen, Eds., Singapore: Springer, 2021, pp. 471–481.

[87] Z. Zhu, J. H. Engel, and A. Hannun, "Learning multiscale features directly from waveforms," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, Sep. 2016, pp. 1305–1309, doi: 10.21437/interspeech.2016-256.

[88] Z. K. Abdul, "Kurdish speaker identification based on one dimensional convolutional neural network," *Comput. Methods Differ. Equ.*, vol. 7, no. 4, pp. 566–572, 2019.

[89] S. Abdoli, P. Cardinal, and A. L. Koerich, "End-to-end environmental sound classification using a 1D convolutional neural network," *Expert Syst. Appl.*, vol. 136, pp. 252–263, Dec. 2019, doi: 10.1016/j.eswa.2019.06.040.

[90] S. Chu, S. Narayanan, and C.-C.-J. Kuo, "Environmental sound recognition with time–frequency audio features," *IEEE Trans. Audio, Speech, Language Process.*, vol. 17, no. 6, pp. 1142–1158, Aug. 2009, doi: 10.1109/TASL.2009.2017438.

[91] W. Bian, J. Wang, B. Zhuang, J. Yang, S. Wang, and J. Xiao, "Audio-based music classification with DenseNet and data augmentation," in *PRICAI 2019: Trends in Artificial Intelligence* (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)), vol. 11672. Cham, Switzerland: Springer, 2019, pp. 56–65, doi: 10.1007/978-3-030-29894-4_5.

[92] B. Logan, "Mel frequency cepstral coefficients for music modeling," in *Proc. ISMIR*, 2000, pp. 1–11.

[93] D. P. W. Ellis, "Classifying music audio with timbral and chroma features," in *Proc. ISMIR*. New York, NY, USA: Columbia Univ., 2007, pp. 339–340.

[94] D.-N. Jiang, L. Lu, H.-J. Zhang, J.-H. Tao, and L.-H. Cai, "Music type classification by spectral contrast feature," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jan. 2002, pp. 113–116, doi: 10.1109/ICME.2002.1035731.

[95] W.-H. Tsai and H.-C. Lee, "Automatic evaluation of karaoke singing based on pitch, volume, and rhythm features," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 4, pp. 1233–1243, May 2012, doi: 10.1109/TASL.2011.2174224.

[96] C. Xu, N. C. Maddage, X. Shao, F. Cao, and Q. Tian, "Musical genre classification using support vector machines," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2003, p. V-429, doi: 10.1109/ICASSP.2003.1199998.

[97] M. Ramona, G. Richard, and B. David, "Vocal detection in music with support vector machines," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 1885–1888, doi: 10.1109/ICASSP.2008.4518002.

[98] J. Pons and X. Serra, "Randomly weighted CNNs for (music) audio classification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 336–340, doi: 10.1109/ICASSP.2019.8682912.

[99] D. Ghosal and M. H. Kolekar, "Music genre recognition using deep neural networks and transfer learning," in *Proc. Interspeech*, Sep. 2018, pp. 2087–2091, doi: 10.21437/interspeech.2018-2045.

[100] N. Farajzadeh, N. Sadeghzadeh, and M. Hashemzadeh, "PMG-Net: Persian music genre classification using deep neural networks," *Entertainment Comput.*, vol. 44, Jan. 2023, Art. no. 100518, doi: 10.1016/j.entcom.2022.100518.

[101] S. Dieleman and B. Schrauwen, "End-to-end learning for music audio," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 6964–6968, doi: 10.1109/ICASSP.2014.6854950.

[102] S. Allamy and A. L. Koerich, "1D CNN architectures for music genre classification," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2021, pp. 1–7, doi: 10.1109/SSCI50451.2021.9659979.

[103] T. Shaikh and A. Jadhav, "Music genre classification using neural network," *ITM Web Conf.*, vol. 44, p. 03016, May 2022, doi: 10.1051/itm-conf/20224403016.

[104] F. Noman, C.-M. Ting, S.-H. Salleh, and H. Ombao, "Short-segment heart sound classification using an ensemble of deep convolutional neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1318–1322.

[105] T. Adesuyi, B.-M. Kim, and J. Kim, "Snoring sound classification using 1D-CNN model based on multi-feature extraction," *Int. J. Fuzzy Log. Intell. Syst.*, vol. 22, no. 1, pp. 1–10, Mar. 2022.

[106] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016, doi: 10.1109/TIE.2016.2582729.

[107] R. Mukhopadhyay, P. S. Panigrahy, G. Misra, and P. Chattopadhyay, "Quasi 1D CNN-based fault diagnosis of induction motor drives," in *Proc. 5th Int. Conf. Electric Power Energy Convers. Syst. (EPECS)*, Apr. 2018, pp. 1–5, doi: 10.1109/EPECS.2018.8443552.

[108] C.-C. Chen, Z. Liu, G. Yang, C.-C. Wu, and Q. Ye, "An improved fault diagnosis using 1D-convolutional neural network model," *Electronics*, vol. 10, no. 1, p. 59, Dec. 2020, doi: 10.3390/electronics10010059.

[109] T.-T. Vo, M.-K. Liu, and C.-L. Hsieh, ''Hybrid 1D CNN-RNN network for fault diagnosis in induction motors using electrical signals,'' in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, Jul. 2023, pp. 212–217, doi: 10.1109/icsse58758.2023.10227168.

[110] L. Eren, T. Ince, and S. Kiranyaz, ''A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier,'' *J. Signal Process. Syst.*, vol. 91, no. 2, pp. 179–189, Feb. 2019, doi: 10.1007/s11265-018-1378-3.

[111] I. H. Ozcan, O. C. Devecioglu, T. Ince, L. Eren, and M. Askar, ''Enhanced bearing fault detection using multichannel, multilevel 1D CNN classifier,'' *Electr. Eng.*, vol. 104, no. 2, pp. 435–447, Apr. 2022, doi: 10.1007/s00202-021-01309-2.

[112] T. Khan, P. Alekhya, and J. Seshadrinath, ''Incipient inter-turn fault diagnosis in induction motors using CNN and LSTM based methods,'' in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting (IAS)*, Apr. 2018, pp. 1–6, doi: 10.1109/IAS.2018.8544707.

[113] J. Chuya-Sumba, L. M. Alonso-Valerdi, and D. I. Ibarra-Zarate, ''Deep-learning method based on 1D convolutional neural network for intelligent fault diagnosis of rotating machines,'' *Appl. Sci.*, vol. 12, no. 4, p. 2158, Feb. 2022, doi: 10.3390/app12042158.

[114] G. J. Hankey, ''Secondary stroke prevention,'' *Lancet Neurol.*, vol. 13, no. 2, pp. 178–194, Feb. 2014, doi: 10.1016/s1474-4422(13)70255-2.

[115] WHO. (2023). *WHO Stroke Satistics*. Accessed: Oct. 17, 2023. [Online]. Available: https://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html

[116] A. Z. Sherzai and M. S. V Elkind, ''Advances in stroke prevention,'' *Ann. New York Acad. Sci.*, vol. 1338, no. 1, pp. 1–15, 2015.

[117] M. Ashrafuzzaman, S. Saha, and K. Nur, ''Prediction of stroke disease using deep CNN based approach,'' *J. Adv. Inf. Technol.*, vol. 13, no. 6, pp. 604–613, 2022, doi: 10.12720/jait.13.6.604-613.

[118] T. Wang, F. Li, X. Zhang, L. Huang, and W. Jia, ''A 1D-CNN prediction model for stroke classification based on EEG signal,'' in *Proc. 8th Int. Conf. Commun. Inf. Process.* New York, NY, USA: Association for Computing Machinery, Nov. 2022, pp. 191–196, doi: 10.1145/3571662.3571695.

[119] B. M. Elbagoury, L. Vladareanu, V. Vlădăreanu, A. B. Salem, A.-M. Travediu, and M. I. Roushdy, ''A hybrid stacked CNN and residual feedback GMDH-LSTM deep learning model for stroke prediction applied on mobile AI smart hospital platform,'' *Sensors*, vol. 23, no. 7, p. 3500, Mar. 2023.

[120] C.-H. Hsieh, Y.-S. Li, B.-J. Hwang, and C.-H. Hsiao, ''Detection of atrial fibrillation using 1D convolutional neural network,'' *Sensors*, vol. 20, no. 7, p. 2136, Apr. 2020, doi: 10.3390/s20072136.

[121] M. M. R. K. Mamun, ''Assessment of potential primary and recurrent ischemic stroke by detecting atrial fibrillation using 1D-CNN and CHA2DS2-VA score,'' in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2021, pp. 1963–1968, doi: 10.1109/BIBM52615.2021.9669885.

[122] E. S. Ho and Z. Ding, ''Electrocardiogram analysis of post-stroke elderly people using one-dimensional convolutional neural network model with gradient-weighted class activation mapping,'' *Artif. Intell. Med.*, vol. 130, Aug. 2022, Art. no. 102342, doi: 10.1016/j.artmed.2022.102342.

[123] N. R. Kifli, H. Hidayat, Rahmawati, F. P. Sukoco, A. R. Yuniarti, and S. Rizal, ''Brain stroke classification using one dimensional convolutional neural network,'' in *Proc. IEEE Asia–Pacific Conf. Wireless Mobile (APWiMob)*, Dec. 2022, pp. 1–6, doi: 10.1109/APWiMob56856.2022.10014207.

[124] F. Y. Dewi, A. Faza, P. Prajitno, and S. K. Wijaya, ''Stroke severity classification based on EEG signals using 1D convolutional neural network,'' *J. Phys., Conf. Ser.*, vol. 1528, no. 1, Apr. 2020, Art. no. 012006, doi: 10.1088/1742-6596/1528/1/012006.

[125] Z. Lasefr, S. S. V. N. R. Ayyalasomayajula, and K. Elleithy, ''Epilepsy seizure detection using EEG signals,'' in *Proc. IEEE 8th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2017, pp. 162–167, doi: 10.1109/UEMCON.2017.8249018.

[126] M. Sameer and B. Gupta, ''CNN based framework for detection of epileptic seizures,'' *Multimedia Tools Appl.*, vol. 81, no. 12, pp. 17057–17070, May 2022, doi: 10.1007/s11042-022-12702-9.

[127] F. Hassan, S. F. Hussain, and S. M. Qaisar, ''Epileptic seizure detection using a hybrid 1D CNN-machine learning approach from EEG data,'' *J. Healthcare Eng.*, vol. 2022, pp. 1–16, Nov. 2022, doi: 10.1155/2022/9579422.

[128] B. Salafian, E. F. Ben-Knaan, N. Shlezinger, S. De Ribaupierre, and N. Farsad, ''MICAL: Mutual information-based CNN-aided learned factor graphs for seizure detection from EEG signals,'' *IEEE Access*, vol. 11, pp. 23085–23096, 2023, doi: 10.1109/ACCESS.2023.3252897.

[129] S. Shanmugam and S. Dharmar, ''A CNN-LSTM hybrid network for automatic seizure detection in EEG signals,'' *Neural Comput. Appl.*, vol. 35, no. 28, pp. 20605–20617, Oct. 2023, doi: 10.1007/s00521-023-08832-2.

[130] N. K. C. Pratiwi, I. Wijayanto, and Y. N. Fu'adah, ''Performance analysis of an automated epilepsy seizure detection using EEG signals based on 1D-CNN approach,'' in *Proc. 2nd Int. Conf. Electron., Biomed. Eng., Health Informat.*, T. Triwiyanto, A. Rizal, and W. Caesarendra, Eds. Singapore: Springer, 2022, pp. 265–277.

[131] G. Xu, T. Ren, Y. Chen, and W. Che, ''A one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signal analysis,'' *Frontiers Neurosci.*, vol. 14, pp. 1–9, Dec. 2020, doi: 10.3389/fnins.2020.578126.

[132] L. Zhu, D. Liu, X. Li, J. Lu, L. Wei, and X. Cheng, ''An efficient hardware architecture for epileptic seizure detection using EEG signals based on 1D-CNN,'' in *Proc. IEEE 14th Int. Conf. ASIC (ASICON)*, Oct. 2021, pp. 1–4, doi: 10.1109/ASICON52560.2021.9620467.

[133] X. Wang, X. Wang, W. Liu, Z. Chang, T. Kärkkäinen, and F. Cong, ''One dimensional convolutional neural networks for seizure onset detection using long-term scalp and intracranial EEG,'' *Neurocomputing*, vol. 459, pp. 212–222, Oct. 2021, doi: 10.1016/j.neucom.2021.06.048.

[134] I. Ullah, M. Hussain, E.-U.-H. Qazi, and H. Aboalsamh, ''An automated system for epilepsy detection using EEG brain signals based on deep learning approach,'' *Expert Syst. Appl.*, vol. 107, pp. 61–71, Oct. 2018, doi: 10.1016/j.eswa.2018.04.021.

[135] R. V. Sharan and S. Berkovsky, ''Epileptic seizure detection using multi-channel EEG wavelet power spectra and 1-D convolutional neural networks,'' in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 545–548, doi: 10.1109/EMBC44109.2020.9176243.

[136] T. T. Chowdhury, A. Hossain, S. A. Fattah, and C. Shahnaz, ''Seizure and non-seizure EEG signals detection using 1-D convolutional neural network architecture of deep learning algorithm,'' in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–4, doi: 10.1109/ICASERT.2019.8934564.

[137] J. S. Ra, T. Li, and YanLi, ''A novel epileptic seizure prediction method based on synchroextracting transform and 1-dimensional convolutional neural network,'' *Comput. Methods Programs Biomed.*, vol. 240, Oct. 2023, Art. no. 107678, doi: 10.1016/j.cmpb.2023.107678.

[138] S. A. Alex, U. Ghosh, and N. Mohammad, ''Weather prediction from imbalanced data stream using 1D-convolutional neural network,'' in *Proc. 10th Int. Conf. Emerg. Trends Eng. Technol. Signal Inf. Process. (ICETET-SIP-22)*, Apr. 2022, pp. 1–6, doi: 10.1109/ICETET-SIP-2254415.2022.9791638.

[139] Q. Fu, D. Niu, Z. Zang, J. Huang, and L. Diao, ''Multi-stations' weather prediction based on hybrid model using 1D CNN and bi-LSTM,'' in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 3771–3775, doi: 10.23919/CHICC.2019.8866496.

[140] L. B. Ferreira and F. F. da Cunha, ''Multi-step ahead forecasting of daily reference evapotranspiration using deep learning,'' *Comput. Electron. Agricult.*, vol. 178, Nov. 2020, Art. no. 105728, doi: 10.1016/j.compag.2020.105728.

[141] A. Lawal, S. Rehman, L. M. Alhems, and M. M. Alam, ''Wind speed prediction using hybrid 1D CNN and BLSTM network,'' *IEEE Access*, vol. 9, pp. 156672–156679, 2021, doi: 10.1109/ACCESS.2021.3129883.

[142] Y. R. Sari, E. C. Djamal, and F. Nugraha, ''Daily rainfall prediction using one dimensional convolutional neural networks,'' in *Proc. 3rd Int. Conf. Comput. Informat. Eng. (IC2IE)*, Sep. 2020, pp. 90–95, doi: 10.1109/IC2IE50715.2020.9274572.

[143] M. Patel, A. Patel, and D. R. Ghosh, ''Precipitation nowcasting: Leveraging bidirectional LSTM and 1D CNN,'' 2018, *arXiv:1810.10485*.

[144] N. D. P. Lestari and E. C. Djamal, ''Rainfall prediction using spatial convolutional neural networks and recurrent neural networks,'' in *Proc. Int. Conf. Data Sci. Its Appl. (ICoDSA)*, Jul. 2022, pp. 12–17, doi: 10.1109/ICoDSA55874.2022.9862821.

[145] M. G. Ragab, S. J. Abdulkadir, N. Aziz, Q. Al-Tashi, Y. Alyousifi, H. Alhussian, and A. Alqushaibi, ''A novel one-dimensional CNN with exponential adaptive gradients for air pollution index prediction,'' *Sustainability*, vol. 12, no. 23, p. 10090, Dec. 2020, doi: 10.3390/su122310090.

[146] M. Zhu and J. Xie, ''Investigation of nearby monitoring station for hourly $PM_{2.5}$ forecasting using parallel multi-input 1D-CNN-biLSTM,'' *Expert Syst. Appl.*, vol. 211, Jan. 2023, Art. no. 118707, doi: 10.1016/j.eswa.2022.118707.

[147] A. Shabarek, S. Chien, and S. Hadri, ''Deep learning framework for freeway speed prediction in adverse weather,'' *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2674, no. 10, pp. 28–41, Oct. 2020, doi: 10.1177/0361198120947421.

[148] X. Chen and G. Yuan, ''Sports injury rehabilitation intervention algorithm based on visual analysis technology,'' *Mobile Inf. Syst.*, vol. 2021, pp. 1–8, May 2021, doi: 10.1155/2021/9993677.

[149] X. Zhang, ''Application of human motion recognition utilizing deep learning and smart wearable device in sports,'' *Int. J. Syst. Assurance Eng. Manage.*, vol. 12, no. 4, pp. 835–843, Aug. 2021, doi: 10.1007/s13198-021-01118-7.

[150] S. Waghchaware and R. D. Joshi, ''Human activity recognition using 1D-CNN and stacked LSTM,'' in *Proc. Int. Conf. Signal Inf. Process. (IConSIP)*, Aug. 2022, pp. 1–5, doi: 10.1109/ICoNSIP49665.2022.10007450.

[151] S. Srimath, Y. Ye, K. Sarker, R. Sunderraman, and S. Ji, ''Human activity recognition from RGB video streams using 1D-CNNs,'' in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI)*, Oct. 2021, pp. 295–302, doi: 10.1109/SWC50871.2021.00048.

[152] K. Vats, M. Fani, P. Walters, D. A. Clausi, and J. Zelek, ''Event detection in coarsely annotated sports videos via parallel multi receptive field 1D convolutions,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 3856–3865, doi: 10.1109/cvprw50498.2020.00449.

[153] J. Guo, B. Wan, S. Zheng, A. Song, and W. Huang, ''A teenager physical fitness evaluation model based on 1D-CNN with LSTM and wearable running PPG recordings,'' *Biosensors*, vol. 12, no. 4, p. 202, Mar. 2022, doi: 10.3390/bios12040202.

[154] Y. Hui, Z. Yin, M. Wu, and D. Li, ''Wearable devices acquired ECG signals detection method using 1D convolutional neural network,'' in *Proc. 15th Int. Symp. Med. Inf. Commun. Technol. (ISMICT)*, Apr. 2021, pp. 81–85, doi: 10.1109/ISMICT51748.2021.9434935.

[155] M. S. Rana, A. Nibali, Z. He, and S. Morgan, ''Automated repair of fragmented tracks with 1D CNNs,'' *Image Vis. Comput.*, vol. 101, Sep. 2020, Art. no. 103982, doi: 10.1016/j.imavis.2020.103982.

[156] T. Nakano, A. Sakata, and A. Kishimoto, ''Estimating blink probability for highlight detection in figure skating videos,'' 2020, *arXiv:2007.01089*.

[157] G. C. Domínguez, E. F. Álvarez, A. T. Córdoba, and D. G. Reina, ''A comparative study of machine learning and deep learning algorithms for padel tennis shot classification,'' *Soft Comput.*, vol. 27, no. 17, pp. 12367–12385, Sep. 2023, doi: 10.1007/s00500-023-07874-x.

[158] P.-E. Martin, J. Benois-Pineau, R. Péteri, and J. Morlier, ''Three-stream 3D/1D CNN for fine-grained action classification and segmentation in table tennis,'' in *Proc. 4th Int. Workshop Multimedia Content Anal. Sports*, Oct. 2021, pp. 35–41, doi: 10.1145/3475722.3482793.

[159] A. R. Kim, H. S. Kim, C. H. Kang, and S. Y. Kim, ''The design of the 1D CNN–GRU network based on the RCS for classification of multiclass missiles,'' *Remote Sens.*, vol. 15, no. 3, p. 577, Jan. 2023, doi: 10.3390/rs15030577.

[160] L. Lin, X. Chen, Y. Shen, and L. Zhang, ''Towards automatic depression detection: A BiLSTM/1D CNN-based model,'' *Appl. Sci.*, vol. 10, no. 23, p. 8701, Dec. 2020, doi: 10.3390/app10238701.

[161] S. P. Devika, M. R. Pooja, and V. Ravi, ''1D CNN model: BERT-based approach for depression and suicide identification,'' in *Lecture Notes in Electrical Engineering*, S. Jain, N. Marriwala, C. C. Tripathi, and D. Kumar, Eds., Singapore: Springer, 2023, pp. 627–633.

[162] M. Abo-Tabik, N. Costen, J. Darby, and Y. Benn, ''Towards a smart smoking cessation app: A 1D-CNN model predicting smoking events,'' *Sensors*, vol. 20, no. 4, p. 1099, Feb. 2020, doi: 10.3390/s20041099.

[163] M. Olmedilla, M. R. Martínez-Torres, and S. Toral, ''Prediction and modelling online reviews helpfulness using 1D convolutional neural networks,'' *Expert Syst. Appl.*, vol. 198, Jul. 2022, Art. no. 116787, doi: 10.1016/j.eswa.2022.116787.

[164] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, ''ECA-Net: Efficient channel attention for deep convolutional neural networks,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539, doi: 10.1109/CVPR42600.2020.01155.

[165] J. Hu, L. Shen, and G. Sun, ''Squeeze-and-excitation networks,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.

[166] Z. Mei, K. Ivanov, G. Zhao, Y. Wu, M. Liu, and L. Wang, ''Foot type classification using sensor-enabled footwear and 1D-CNN,'' *Measurement*, vol. 165, Dec. 2020, Art. no. 108184, doi: 10.1016/j.measurement.2020.108184.

[167] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, ''Driver behaviour detection using 1D convolutional neural networks,'' *Electron. Lett.*, vol. 57, no. 3, pp. 119–122, Feb. 2021, doi: 10.1049/ell2.12076.

[168] S. Mohine, B. S. Bansod, R. Bhalla, and A. Basra, ''Acoustic modality based hybrid deep 1D CNN-BiLSTM algorithm for moving vehicle classification,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 16206–16216, Sep. 2022, doi: 10.1109/TITS.2022.3148783.

[169] H. Parineh, M. Sarvi, and S. A. Bagloee, ''Detecting emergency vehicles with 1D-CNN using Fourier processed audio signals,'' *Measurement*, vol. 223, Dec. 2023, Art. no. 113784, doi: 10.1016/j.measurement.2023.113784.

AYOKUNLE OLALEKAN IGE received the B.Sc. (Hons.) and M.Sc. degrees in computer science from Adekunle Ajasin University, Akungba Akoko, Ondo, Nigeria, in 2014 and 2019, respectively, and the Ph.D. degree in intelligent system techniques from the School of Computer Sciences, Universiti Sains Malaysia, in 2023, under the mentorship of Dr. Mohd Halim Mohd Noor.

He is a Faculty Member with the Department of Computer Science, Adekunle Ajasin University. He currently holds a postdoctoral research fellowship position with the University of South Africa, Florida Campus, Johannesburg. He is dedicated to advancing the field of pervasive computing, machine learning, and deep learning through both innovative research and active participation in the academic community. He has authored numerous research papers and presented his findings at prestigious conferences. His contributions to the field have been recognized globally, most notably when he received the Best Paper Award at the 2nd IEEE NBEC Conference in Melaka, Malaysia, underscoring his dedication in the field.

MALUSI SIBIYA received the bachelor's degree in electronics engineering and the National Diploma degree from the Central University of Technology, Free State, and the B.Tech. degree in electrical engineering with a specialization in control engineering from the University of South Africa, in 2015. He is currently pursuing the Ph.D. degree in science, engineering, and technology, specializing in machine learning and natural language processing.

His passion for programming, self-taught, led him to focus on computer science-related research during both the master's and Ph.D. studies. He has accumulated 19 years of experience in academia, with 14 years spent at the college level as a Lecturer of programming and electronics subjects, and five years at the university level teaching electronics and computer science.