2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS) | 979-8-3503-1071-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/PCEMS58491.2023.10136076

Capsule Network for 1-D Biomedical signals: A Review

M Krishna Chaitanya
School of Electronics Engineering
VIT-AP University
Amaravati, Andhra Pradesh, 522237, India
kchaitanya277@gmail.com

Lakhan Dev Sharma
School of Electronics Engineering
VIT-AP University
Amaravati, Andhra Pradesh, 522237, India
devsharmalakhan@gmail.com

***Abstract*—The heartbeat, muscle contractions, and other physiological functions are examples of biomedical signal sources. Electrocardiograms (ECG), electroencephalograms (EEG), and electromyograms (EMG) are examples of the signals that can be non-invasively recorded and used for diagnosis and as health indicators. Hence, timely and accurate diagnosis of the biomedical signals plays a prominent role. Professional healthcare workers assess the signal in search of a clear pattern that would indicate a normal or abnormal heartbeat is a tedious job. Manual interpretation of the signals may lead to misdiagnosis. The automated computer-aided diagnosis (CAD) method is one way to support decision-making for the eradication of these deficiencies. The CAD tool should operate as a real-time system for early diagnosis, requiring little time investment, data dependence, and device-specific measurement variances. Deep learning-based methods are becoming more and more common in CAD techniques. Convolutional neural network (CNN), one of the well-known deep learning network, fail of recognise position, texture, and genetic anomalies in the image. A capsule network is one of the newest and most promising deep learning algorithms that tackles CNN's shortcomings. In this study, we present a thorough analysis of the cutting-edge methodology, tools, and topologies used in current capsule network implementations. The key contribution with this review study is its explanation and summary of major existing Capsule Network implementations and architectures.**

***Index Terms*—Electrocardiogram, Electroencephalogram, Electromyogram, Computer-aided diagnosis, Deep learning, Capsule network**

I. Introduction

Biomedical signal processing is collecting as well as preprocessing physiological signals in order to extract useful data that can be used to spot patterns and trends [1], [2]. The basic steps involved in the biomedical signal processing are signal acquisition, preprocessing, signal analysis. The process of signal acquisition entails employing sensors to record pertinent biomedical data and transforming it into a format that can be processed. Preprocessing is the process of eliminating unwanted inferences that arise during the signal acquisition process. Decision-making and feature extraction are involved in signal analysis. Feature extraction is utilised to communicate a biological system's state [3], [4]. Making decisions is crucial in clinical settings where a doctor must take action based on the results.

A. 1-D Biomedical signals

Numerous different systems make up living creatures. Biosignals are gathered from the human body and transmit details about the condition or behaviour of numerous physical systems in the human body [1]. In order to make the crucial information easily accessible for computer analysis or for human observers, signal processing is applied to the raw biological signal by removing specific unnecessary information from the signal [2]. Examples of the 1-D biomedical signals are Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyogram (EMG), etc.,

*1) **Electrocardiogram (ECG)**:* The ECG is a tool that is the graphical depiction used to examine the heart's electrical activity [5]. A conventional ECG signal consists of a P wave, a QRS complex, a T wave, and a U wave [6]. Atrial depolarization is represented by the P wave, ventricular depolarization and repolarization are represented by the QRS complex [1], and ventricular repolarization of the heart is represented by the T wave [7], The repolarization of the Purkinje fibres is hypothesised to be represented by U waves . Visual illustration of ECG signal can be viewed in [8]. Any variation in the normal morphology of ECG signal might be attributed as cardiac abnormalities.

*2) **Electroencephalogram (EEG)**:* A tool called an EEG is used in the method of electroencephalography to interpret electrical potential from the brain [9]. Based on signal frequencies that span between 0.1 Hz and 100 Hz, brain signals are typically classed as delta, theta, alpha, beta, and gamma. The assessment of EEG signals is frequently used to identify the cognitive strain [10], brain activity, emotions, and neurological abnormalities [11].

*3) **Electromyogram (EMG)**:* The EMG signal, a biological signal which identifies electrical currents generated in muscles while their contractions and denotes neuromuscular actions, is created when muscles contract. Muscle contraction and relaxation are always under the direction of the neurological system. In light of this, the EMG signal is a complex signal that is managed by the nervous system and depending on the physiological and anatomical characteristics of muscles [12].

To achieve better results, these biomedical signals must be preprocessed before being applied to the deep neural network. Recently, several preprocessing methods have been developed. A novel design approach is put forth in [13] for digital bandpass as well as band-stop IIR filters with almost linear phase response. Swarm intelligence-based filters were used in conjunction with fractional derivatives (FDs) in [14]–[17]. Particle swarm optimization and the artificial bee colony algorithm are used to improve the performance of digital IIR filters, and this is examined in the frequency domain in [13], [18].

B. *Artificial Intelligence and deep learning*

A famous area of computer science called artificial intelligence (AI) is entrenched on neuronal network of brain, which is what gives that organ its intelligence. The weighted synapses that connect the layers and neurons in the Network of Artificial Neurons are connected to one another. By using back-propagation, the weights are changed [19]. An activation function governs how strongly a neuron fires. They are required for non-linearity as well [20]. Rectified Linear Unit (ReLU), Sigmoid, Hyperbolic Tangent function (tanh), and the SoftMax Activation Function are a few examples of activation functions. Readers are advised to study the article [21] to learn how to choose activation functions. Calculating the cost function is a key element of neural network theory. Deep neural networks frequently apply these ideas to do tasks like disease diagnosis, language translation, image processing, speech recognition, and facial (expression) identification.

The remainder of this article is structured as follows: section II consists of Convolutional neural networks, section III is about introduction to Capsule network, section IV briefs about one dimensional capsule network, brief discussion is presented in the section V and concluded in the section VI.

II. Convolutional Neural Networks (CNNs)

Convolutional neural networks feature flattening, pooling, and fully connected layers in addition to convolutional layers [22]. A $n \times n$ kernel ($n > 0$ and $m > 0$) examines the image input during convolutions to instinctively perform feature extraction. ReLU [23], [24] is used immediately following the convolutional process to introduce nonlinearity to the model and decrease computational complexity. To guarantee that CNN detects the identical item in images of various formats and to lower model's memory exigencies, pooling is done on the feature map. There are various types of pooling, including maximum, minimum, average (also known as sub-sampling), and sum [25]. Regarding the fully linked layers, CNNs differ from artificial neural networks in that hidden layers in CNNs necessarily fully connected. However, a rigid necessity in artificial neural networks. For improved performance, back-propagation adjusts both the weights as well as the feature detectors.

CNNs face the following significant challenges:

- Unable to distinguish between the position, texture, and genetic abnormalities of image or specific image elements [26].
- CNN's pooling process causes some visual details to be lost. As a result, they need a lot of training data to make up for this loss.
- CNNs are more prone to adversarial attacks such as pixel perturbations resulting in wrong classifications [27].
- Image reconstruction is particularly labor-intensive as a consequence of max pooling in CNN..

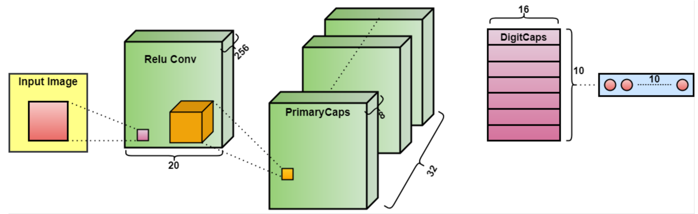

Fig. 1: An illustration of the architecture of CapsNet as depicted in [26]

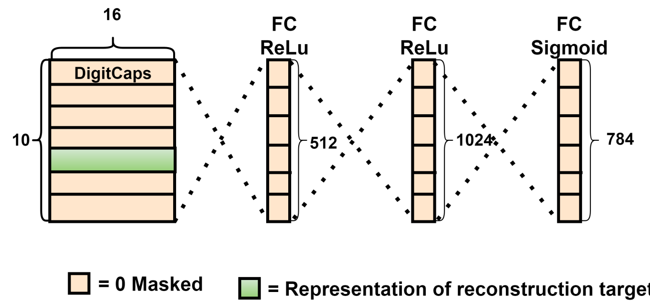

Fig. 2: The architecture of Decoder network used on top of DigitCaps as presented in [26]

III. Capsule Networks (CapsNet)

Capsule Networks (CapsNet), a replacement for CNNs, were suggested by Hinton and his associates [28]. Unlike CNNs, which only receive and output scalar information, capsules are equivariant neural networks made up of neurons. Because of this characteristic, a capsule can learn not only the deformations as well as viewing conditions of picture, but also its characteristics. The output (or characteristics) from a CNN are the feed to a capsule. Depending on the kind of capsule used, these properties are processed differently. The chance that the feature represented by the capsule is existing and a group of vector values generally referred to as instantiation parameters make up a capsule's output. The instantiation properties are used to illustrate the network's equivariance, which shows that it can distinguish between deformations, texture, and posture. Ability of the prototype to stay consistent despite changes to the inputs is known as invariance. The architecture of the CapsNet is illustrated in the Figure 1 and the decoder used above the DigitCaps layer is depicted in the Figure 2 .

There are three standard ways to implement capsules in literature. They are transforming dynamic routing-based vector capsules [26], auto-encoders [28], and expectation-maximization-based matrix capsules [29].

1) Transforming Auto-encoders: It was designed to highlight a network's capacity to recognise pose. The objective was to accept a picture as well as its stance as the input and produce the very identical image in the initial stance rather than to identify items in the pictures. In this initial approach, a capsule's output vector consisted of outputs, of which one reflected the probability of the attribute existing and others the basis factors. The levels in which capsules can be organised are lower stage *k*, called as primary capsules, as well as upper stage *k* + 1, termed as secondary capsules. To start a part-whole hierarchy, lower level capsules extract posture information from pixel vividness. This part-to-whole arrangement is advantageous for capsule networks as it enables the identification of the complete system by first identifying its key elements.

2) Dynamic routing between capsules: The following improvement to capsule networks defines a capsule as a collection of neurons with activity vectors serving as the instantiation parameters as well as duration of the vector serving as the possibility that feature will sustain. Primary, Convolutional, as well as class capsule layers make up the network. The initial capsule layer is the Primary capsule (PC) layer , and subsequent layers can come in any order up until the Class capsule (CC) layer , which is the final layer. The PC layer receives the output from the convolutional layer, which retrieves features from image. The amount that the PC i serves to the CC j is shown by a PC's prediction vector, $\widehat{u}_{j|i}$.

To determine a sole PC's forecasting the CC, the coupling coefficient indicating the congruence between these capsules is multiplied by the prediction vector. The two capsules are pertinent to one another if the concordance is high. If this happens, the coupling coefficient would rise; if not, it will fall. A weighted sum (s_j) is created from each of these distinct PC forecasts for the class capsule in order to identify the candidates for the squashing function (v_j). The squashing function makes sure that, like a probability, the duration of output first from capsule lies within 0 and 1. the equation for the squashing function is given by the equation 1.

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \tag{1}$$

3) Matrix capsules with expectation- maximization routing: The expectation maximisation algorithm also took the place of dynamic routing via agreement [30]. A cosine between two posture vectors was used in dynamic routing, although it wasn't completely successful. Additionally, a parameter rather than the length of a vector was used to replace the chance of an object being described by capsule. This made it easier to stay away from squashing function, that was regarded as "not objective and rational." In order to function efficiently, the EM routing algorithm makes use of capsule networks with multiple layers of capsules. Let Ω_L stand for the group of capsules in the primary layer, with M standing for the posture matrix and a for the activation probability of each capsule. In order for expectation maximization to function, the stance matrix of capsule i must be modified by transformation weight matrix W_{ij} in order to be elected in the stance matrix belonging to capsule j at level $L+1$. Then the resultant matrix M_i as well as transformation matrix W_{ij} are components of the vote.i.e.,

$$V_{ij} = M_i W_{ij} \tag{2}$$

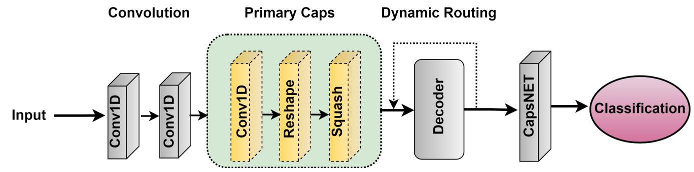

Fig. 3: An illustration of the 1D-CapsNet model as depicted in [31].

A. Analysis of the CapsNets architecture

Two convolutional layers made up the fundamental framework of Sabour and Hinton's in the year 2017 initially successful capsule network [26]. Conv1 utilised ReLU activation function on a MNIST picture with dimensions of $28 \times 28 \times 1$ and comprised 256 channels, each of which was composed of $times9$ filters with just a stride of 1. A convolutional capsule layer with $6 \times 6 \times 32$ capsules that each output an 8D vector was the second layer's design. Each PC operates with 8 convolutional units and a 9×9 kernel with a stride of 2. Non-linearity is produced when a squashing function is used to create 10 and 16D capsules. Additionally, this layer serves as the main capsule layer that receives features as scalar outputs from the Conv1 layer, following which all subsequent layers must deal with 8D vector values. Each of the layer's 32 channels has a 6x6 grid of main capsules. The third layer (DigitCaps) is a completely connected layer with 10 16D capsules that execute classification based on ten classes by receiving input from every capsule in the layer below.

IV. 1D-CapsNet

For automatic 1-D signal (ECG, EEG, EMG, etc.) detection, the 1D-CapsNet framework of a one-dimensional deep capsule network was created [31]. The original capsule network framework [29], created for handling picture information, has been modified as function for 1D signals. In addition, to distinguish 1-D signals from the original prototype, different layer specifications were established and further layers were created. In addition, to distinguish 1-D signals from the original model, different layer settings were established and further layers were appended. The schematic of 1D-CapsNet model intended for classification of abnormalities as well as normal can be seen in Figure 3. In the convolution section, two 1-D convolution (Conv1D) layers were applied to the input signals in order to extract feature representations from them. The Primary Caps layer's entry was created from a variety of characteristic maps derived from convolution layers. Information about the orientation, magnitude, and other factors is contained in the activation vectors. Activation vector lengths display the predicted confidence interval of a vector relying upon signal characteristics. The orientation portion that belongs to the activation vector contains the abstract depictions of signals. If an activation vector is long enough, it indicates that the capsules found key characteristics in the Primary Caps block that they were searching for. On the other hand, if an activation vector's length is quite modest, it means that the capsules failed to identify any noteworthy features of the data. The set of feature maps was transformed into the appropriate vectors using reshape layer, which was put ahead of convolutional layer. In the last stage of the Primary Caps layer, the squash function was then utilised to ensure that all vector durations were within the range of 0 to 1. By using the reconstruction loss, a deep decoder structure with four layers was employed. A total of 3 masking layers, 3 dense layers, each with dimensions of 128, 256, and 2 , were present in the decoder network. The output layer of the 1D CapsNet was its final component. Here, the sigmoid function was used to identify the classifications of ECG signals. By calculating squared difference between input signal and reconstituted signal, the decoder network rebuilt the input signal.

Due to dynamic routing, the 1D-CapsNet model has the drawback of having a larger computing cost than CNN models.

V. Discussion

Capsule network is the new sensation in the AI and produced incredible results in comparison with traditional neural networks and CNNs. The comparision of various researches carried out using CapsNet is shown in the Table I. In [32], the ECG signals are converted to two dimensional image (spectrogram) and then fed to the CapsNet for classification of cardiac arrhythmias. In [33], utilized GPU-based computing capabilities, a CapsNet architecture is applied for the categorization of RAS family of proteins structures. In [34], the efficiency and sensitivity of CapsNets were examined in relation to changes that were applied to them gradually as well as under control and found that on datasets that were just a little bit harder than MNIST, CapsNets performed better than AlexNet, but they weren't as excellent at handling nonlinear behavior as one might anticipate. A deep learning oriented 1D-CapsNet network was recommended by researchers of [31] for the automatic investigation of coronary artery disorders from two- and five-second ECG data.

Across various datasets, the capsule's performance varies. CapsNets continue to face difficulties with datasets like CIFAR10 [34], ImageNet. Unfortunately, these images and ones of a similar nature can provide the information needed as entries to models needed to carry out crucial activities. Researchers have suggested a number of modifications since it was first developed, demonstrating that we have not yet found the best form. Unfortunately, current research efforts are mostly focused on the research of Sabour et al. [26] rather than its enhanced counterpart Hinton et al. [29], which now has undergone relatively modest changes. The approach might not be suited for several additional applications that really are unrelated to machine vision due to additional issues like the tight definition of things for capsules. Existing loss functions is challenged by vector/matrix outcome of capsules, and CapsNets find it challenging to discriminate between several instances of element at particular given point in the source image. Despite the fact that their training duration is superior to CNN's, it is still unacceptable for time-sensitive activities and terribly inappropriate for virtual training. Promising investigation is now being conducted in this field [35]. The science community can attempt to tackle these restrictions in the near future as they are all lingering questions. We won't be able to utilise capsules to their full potential until this is completed. EEG emotion detection algorithm based on the attention mechanism as well as a pre-trained convolution capsule structure to more accurately identify different emotions are presented in [36]. The Transformer Capsule System proposed in [37] primarily consists of two modules: the EEG Transformer module, which extracts EEG features, and the Emotion Capsule module, which improves the features and categorizes the emotion states.

VI. Conclusion

One of the most critical jobs in AI still involves processing unstructured input for machine vision applications. The difficult feature engineering effort that eventually resulted in high dimensionality has been removed with the advent of Deep Learning. Although Deep Learning models like CNNs have excelled in this area, their implementation calls for a significant amount of data and processing power. In order to overcome the difficulties faced by CNNs, capsules were developed, and they have thus far performed admirably. However, given the field's relative youth, further research is necessary before its maximum potential can be realised. Therefore, the models that have had a big impact on the area of 1D biomedical signal analysis in literature were studied in this research. The research conducted a futuristic analysis of capsule networks. The computer vision community hopes to construct strong machine vision algorithms by building on the triumphs and mistakes of CapsNets via extensive field research. Future research may apply one dimensional CapsNet to other 1D biomedical signals, which can be used to assess the health condition.

References

[1] A. C. Guyton and J. E. Hall, "Guyton and hall textbook of medical physiology," 2011.

[2] R. M. Rangayyan, *Biomedical signal analysis*. John Wiley & Sons, 2015.

[3] M. Akay, *Biomedical signal processing*. Academic press, 2012.

[4] A. Luthra, *ECG made easy*. Jaypee Brothers Medical Publishers, 2019.

[5] M. K. Chaitanya and L. D. Sharma, "Electrocardiogram signal filtering using circulant singular spectrum analysis and cascaded savitzky-golay filter," *Biomedical Signal Processing and Control*, vol. 75, p. 103583, 2022.

[6] L. D. Sharma, R. K. Sunkaria, and A. Kumar, "Bundle branch block detection using statistical features of qrs-complex and k-nearest neighbors," in *2017 conference on information and communication technology (CICT)*. IEEE, 2017, pp. 1–4.

[7] L. D. Sharma and R. K. Sunkaria, "Novel t-wave detection technique with minimal processing and rr-interval based enhanced efficiency," *Cardiovascular Engineering and Technology*, vol. 10, no. 2, pp. 367–379, 2019.

TABLE I: Comparison of 1D-CapsNet and other cutting-edge (deep learning) studies

Reference	Year	Network	Input	Objective	Performance Metrics
[33]	2018	CAPSNET	Protein structures	Protein family structure prediction and classification.	ACC
[34]	2018	CAPSNET	Images of hand written digits, images ofhouse number, photos of animals and automobiles	Test CapsNets using datasets similar to MNIST and investigate the underlying embedding space and the sources of inaccuracy.	ACC
[38]	2019	CAPSNET	Fingerspelling alphabets	CapsNet's difficulty to perform well on other datasets aside MNIST	ACC, Speed of classification
[39]	2019	CAPSNET	Complex ultrasonic data	Non-image classification for Selfdriving cars.	ACC
[40]	2019	CAPSNET	Fluorescence microscopic images	Protein categorization	ACC
[41]	2019	CAPSNET	EEG signals	Emotion Recognition	ACC
[42]	2019	CAPSNET	Motor imagery EEG signals	Classify two class motor imagery	ACC
[31]	2020	1-D CAPSNET	ECG Signals	Identification of coronary artery disease	ACC, SEN,SPY
[32]	2020	CAPSNET	ECG Signals	ECG arrhythmia classification	ACC
[43]	2020	CAPSNET	Combination of MDL Molecular Access as well as molecular descriptors	Classification of hERG Blockers /Nonblockers	ACC, SEN,SPY

[8] M. K. Chaitanya, L. D. Sharma, A. Roy, and J. Rahul, "A review on artificial intelligence for electrocardiogram signal analysis," *Big Data Analytics and Artificial Intelligence in the Healthcare Industry*, pp. 38–72, 2022.

[9] J. S. Kumar and P. Bhuvaneswari, "Analysis of electroencephalography (eeg) signals and its categorization–a study," *Procedia engineering*, vol. 38, pp. 2525–2536, 2012.

[10] L. D. Sharma, R. K. Saraswat, and R. K. Sunkaria, "Cognitive performance detection using entropy-based features and lead-specific approach," *Signal, Image and Video Processing*, vol. 15, no. 8, pp. 1821–1828, 2021.

[11] L. D. Sharma, H. Chhabra, U. Chauhan, R. K. Saraswat, and R. K. Sunkaria, "Mental arithmetic task load recognition using eeg signal and bayesian optimized k-nearest neighbor," *International Journal of Information Technology*, vol. 13, no. 6, pp. 2363–2369, 2021.

[12] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of emg signal analysis: detection, processing, classification and applications," *Biological procedures online*, vol. 8, no. 1, pp. 11–35, 2006.

[13] N. Agrawal, A. Kumar, V. Bajaj, and G. K. Singh, "Design of bandpass and bandstop infinite impulse response filters using fractional derivative," *IEEE transactions on industrial electronics*, vol. 66, no. 2, pp. 1285–1295, 2018.

[14] N. Agrawal, A. Kumar, and V. Bajaj, "A new design method for stable iir filters with nearly linear-phase response based on fractional derivative and swarm intelligence," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 1, no. 6, pp. 464–477, 2017.

[15] A. Kumar, N. Agrawal, I. Sharma, S. Lee, and H.-N. Lee, "Hilbert transform design based on fractional derivatives and swarm optimization," *IEEE Transactions on Cybernetics*, vol. 50, no. 5, pp. 2311–2320, 2018.

[16] N. Agrawal, A. Kumar, and V. Bajaj, "Design of infinite impulse response filter using fractional derivative constraints and hybrid particle swarm optimization," *Circuits, Systems, and Signal Processing*, vol. 39, pp. 6162–6190, 2020.

[17] N. Agrawal, A. Kumar, B. Kuldeep, S. Lee, and H.-N. Lee, "Weighted least square design technique for hilbert transformer using fractional derivative," *Signal, Image and Video Processing*, vol. 15, pp. 1461–1468, 2021.

[18] N. Agrawal, A. Kumar, and V. Bajaj, "A new method for designing of stable digital iir filter using hybrid method," *Circuits, Systems, and Signal Processing*, vol. 38, pp. 2187–2226, 2019.

[19] M. Nielson, "Neural networks and deep learning [www document]," *URL http://neuralnetworksanddeeplearning. com/chap2. html.(Accessed 6.17. 19)*, 2019.

[20] S. Pattanayak, J. S. Pattanayak, and S. John, *Pro deep learning with tensorflow*. Springer, 2017.

[21] H. Mhaskar and C. Micchelli, "How to choose an activation function [www document]," *URL https://papers. nips. cc/paper/874-how-to-choose-anactivation-function. pdf.(Accessed 7.6. 19)*, 1994.

[22] J. Wu, "Introduction to convolutional neural networks," *National Key Lab for Novel Software Technology. Nanjing University. China*, vol. 5, no. 23, p. 495, 2017.

[23] C.-C. J. Kuo, "Understanding convolutional neural networks with a mathematical model," *Journal of Visual Communication and Image Representation*, vol. 41, pp. 406–413, 2016.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.

[25] D. Scherer, A. Müller, and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in *International conference on artificial neural networks*. Springer, 2010, pp. 92–101.

[26] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *Advances in neural information processing systems*, vol. 30, 2017.

[27] J. Su, D. V. Vargas, and K. Sakurai, "One pixel attack for fooling deep neural networks," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 828–841, 2019.

[28] G. E. Hinton, A. Krizhevsky, and S. D. Wang, "Transforming auto-encoders," in *International conference on artificial neural networks*. Springer, 2011, pp. 44–51.

[29] G. E. Hinton, S. Sabour, and N. Frosst, "Matrix capsules with em routing," in *International conference on learning representations*, 2018.

[30] M. Engelin, "Capsnet comprehension of objects in different rotational views: A comparative study of capsule and convolutional networks," 2018.

[31] E. Butun, O. Yildirim, M. Talo, R.-S. Tan, and U. R. Acharya, "1d-cadcapsnet: One dimensional deep capsule networks for coronary artery disease detection using ecg signals," *Physica Medica*, vol. 70, pp. 39–48, 2020.

[32] T. Neela and S. Namburu, "Ecg signal classification using capsule neural networks," *IET Networks*, vol. 10, no. 3, pp. 103–109, 2021.

[33] D. R. de Jesus, J. Cuevas, W. Rivera, and S. Crivelli, "Capsule networks for protein structure classification and prediction," *arXiv preprint arXiv:1808.07475*, 2018.

[34] P. Nair, R. Doshi, and S. Keselj, "Pushing the limits of capsule networks," *arXiv preprint arXiv:2103.08074*, 2021.

[35] A. Marchisio, B. Bussolino, A. Colucci, M. A. Hanif, M. Martina, G. Masera, and M. Shafique, "X-traincaps: accelerated training of capsule nets through lightweight software optimizations," *arXiv preprint arXiv:1905.10142*, 2019.

[36] S. Liu, Z. Wang, Y. An, J. Zhao, Y. Zhao, and Y.-d. Zhang, "Eeg emotion recognition based on the attention mechanism and pre-trained convolution capsule network," *Knowledge-Based Systems*, p. 110372, 2023.

[37] Y. Wei, Y. Liu, C. Li, J. Cheng, R. Song, and X. Chen, "Tc-net: A transformer capsule network for eeg-based emotion recognition," *Computers in Biology and Medicine*, vol. 152, p. 106463, 2023.

[38] H. P. Nguyen and B. Ribeiro, "Advanced capsule networks via context awareness," in *International Conference on Artificial Neural Networks*. Springer, 2019, pp. 166–177.

[39] M. Pöpperli, R. Gulagundi, S. Yogamani, and S. Milz, "Capsule neural network based height classification using low-cost automotive ultrasonic sensors," in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 661–666.

[40] X. Zhang and S.-G. Zhao, "Fluorescence microscopy image classification of 2d hela cells based on the capsnet neural network," *Medical & biological engineering & computing*, vol. 57, no. 6, pp. 1187–1198, 2019.

[41] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion recognition from multiband eeg signals using capsnet," *Sensors*, vol. 19, no. 9, p. 2212, 2019.

[42] K.-W. Ha and J.-W. Jeong, "Motor imagery eeg classification using capsule networks," *Sensors*, vol. 19, no. 13, p. 2854, 2019.

[43] Y. Wang, L. Huang, S. Jiang, Y. Wang, J. Zou, H. Fu, and S. Yang, "Capsule networks showed excellent performance in the classification of herg blockers/nonblockers," *Frontiers in pharmacology*, vol. 10, p. 1631, 2020.