Transformer Models for Processing Biological Signal

Ivan Kuzmanov, Nevena Ackovska, Ana Madevska Bogadnova
Ss. Cyril and Methodius University,
Faculty of Computer Science and Engineering
Skopje, North Macedonia
ivan.kuzmanov@students.finki.ukim.mk,
nevena.ackovska@finki.ukim.mk,
ana.madevska.bogdanova@finki.ukim.mk

***Abstract*—The transformer neural network architecture is a deep learning model, that has been developed recently and as such it's potential is still being investigated. It is a powerful model due to the their self-attention mechanism that finds use in several domains, but our focus is on transformers used for biological signals processing. Various hybrid model architectures suitable for this type of task are considered in this study: the basic transformer, temporal fusion transformer, time series transformer, convolutional vision transformer and informer. A brief description of the architecture is given. The reasons why they are appropriate for processing biological signals, what makes them unique, along with their strengths and weaknesses, are discussed. Finally, a literature review is made involving actual studies that use these model types for biosignal processing.**

Keywords*—*transformer, ECG, EEG, EMG, PPG, biological signals, processing

I. Introduction

A transformer is a deep learning model architecture developed as an computationally efficient alternative to recurrent and convolutional neural networks. Their primary use is in natural language processing (NLP). They are introduced in a paper published in 2017, titled "Attention Is All You Need" [1]. The innovation of transformers lies in the self-attention mechanism. This mechanism allows the model to selectively focus on certain parts of the input sequence. It actualized this by assigning weight to emphasize the significance of the different input parts.

There are two parts of a transformer model: the encoder and the decoder. The encoder encodes the input to a fixed length representation, which the decoder decodes in the output. In the literature this type of neural network is referred to as a sequence to sequence (seq2seq) model [2]. This is an important capability of transformers. However, this ability doesn't limit them. The architecture is robust and versatile enough to be used on tasks that aren't strictly seq2seq.

Recurrent neural network (RNN) is a type of network that contains recurrent connections between the neurons. The most popular RNN is called Long Short Term Memory (LSTM) and was developed to solve the vanishing gradient problem, that plagues this type of networks. The neurons in an LSTM neural network are cells capable of memorizing states. The flow of information in each cell is controlled via three gates: input gate, forget gate, and output gate. This architecture allows the models to detect both short term and long term dependencies in a sequential data [3]. Before the advent of transformers these type of models were widely used for NLP problems.

Convolution neural network (CNN) are based on the operation convolution, a operation that outputs a function by combines two other functions. This type of networks are the most suited for image processing and vision tasks. A typical CNN has three type of layers: convolutional layer that perform the convolutions over the input and produce an output in the form of features maps, pooling layers that downgrade the spatial resolution and a fully connected layer that flattens the input from the previous layer and performs the prediction [4]. This is a broad overview of the architecture, however, there are many modified versions some of which include the transformer self-attention mechanism. CNNs are mostly used in vision tasks, such as object detection, facial recognition, image and video processing. It's not uncommon to see this architecture used for signal processing as well.

The transformer architecture holds many advantages over the recurrent and convolutional neural networks, such as greater parallelization that allows for shorter training times, far fewer parameters, the potential for transfer learning allows repurposing existing models for a similar task, the ability to process sequences of variable length and greater interpretability. However, they have their fair share of disadvantages as well, such as struggling to learn complex patterns and long term dependencies.

There are several variants developed based on the transformer model, each attempting to improve or specialize the existing architecture on a specific type of problem. This type of models have achieved state-of-the-art results on many benchmark NLP datasets. One such model is the BERT, which is an acronym Bidirectional Encoder Representations from Transformers [5]. The innovation of the BERT architecture lies in its ability to pre-train models that can be easily fine tuned for various language modeling tasks. The popular chat application ChatGPT is another example of a neural network that utilizes a modified transformer architecture, the Generative Pre-trained Transformer (GPT) [6]. The capabilities of this application include language modeling, question answering, translation, sentiment analysis, etc. The architecture is versatile enough to be adapted to variety of problems.

This paper is structured as follows: the second section gives a short literature review based on the type of transformers used for processing biological signals, the third section focuses on an in-depth look of the actual architecture of the transformers used when working with biosignals and finally the conclusion is presented in the fourth section.

II. Transformer Architecture Models

The transformer architecture is powerful when used by itself, however, improvements have been made by combining it with other neural networks. The idea behind this type of model is to use the strengths and cover the weaknesses of the different architecture types. All of the previously mentioned architectures except the basic transformer are a part of the family hybrid transformers. Often times transformers that belong to this family attempt to use the attention mechanism in tandem with recurrence and/or convolution, since those are powerful and well research architecture.

In this section we will compare some variants of transformers adapted to processing biological signals. The first architecture described is the initial as introduced in the 2017 paper (Vaswani et al). While this architecture might not be the most suitable for biosignal processing it provides a baseline for comparison and provides context for the modifications made in the other architectures. The other explained architectures are: temporal fusion transformer, time series transformer, convolutional vision transformer, and informer. The methodology used to describe the considered models in this research paper is as follows: the basic transformer serves a baseline from which the other models are derived and as such it's architecture is explored in depth. The unique architecture of the derived models is noted and illustrated, and their strengths and weakness are discussed, as well as directly compared in a table format.

A. Basic Transformer

The basic transformer consists of two parts: an encoder and a decoder. The encoder is a stack of identical layers that "encodes" the input into a sequence of hidden representations. While the decoder is stack that "decodes" the sequence of hidden representations and produces an output in the form of a sequence of predictions. Figure 1 presents the transformer model architecture as described in "Attention is all you need" [1]. The picture shows the encoder and decoder side by side and the steps the input passes to calculate the output.

The self-attention mechanism is a component in every encoder and decoder layer that allows the model to attend to different parts of the input sequence without relying on recurrent connections. At each time step the weighted sum is calculated by the input sequence. This means that the focus may be different at different time steps. The outputs of every self-attention layer are fed to a feed-forward network. This allows the models to track both short term and long term dependencies.

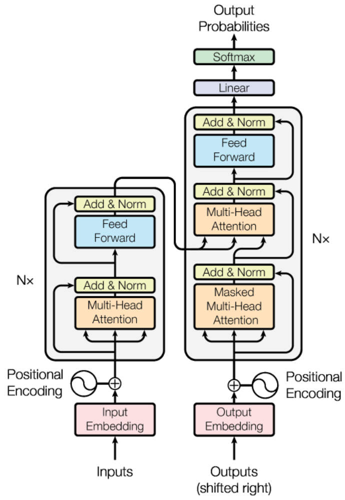

Fig. 1. Transformer architecture [1]

Residual connections and normalization is used to enhance the learning process and make the model more robust. For this type of model it is necessary to include positional encoding in the model architecture for the model. Otherwise the model won't be able to understand the order of the sequence, since it's lacking both recurrent and convolutional layers.

B. Temporal Fusion Transformer (TFT)

Temporal Fusion Transformer (TFT) is a modified version of the transformer architecture. It's specifically designed to handle time series data. It's a recent model proposed in a paper funded by Google, titled "Temporal Fusion Transformers for interpretable multi-horizon time series forecasting" [7]. Figure 2 illustrates the architecture of this type of transformers, taken from the same study [7]. This model architecture is designed to factor in both short term dependencies via recurrent layers and long term dependencies via the self-attention layer. The temporal fusion transformer is a flexible model that finds use in real world application.

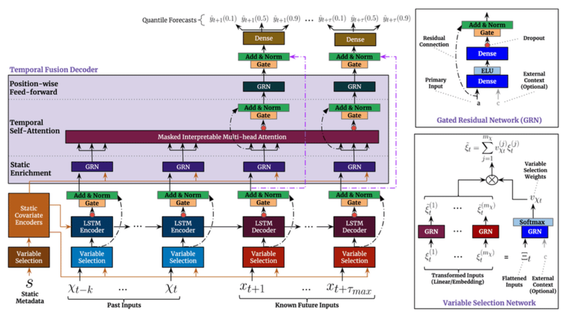

Fig. 2. Temporal fusion transformer architecture [7]

This model is made of 5 components: "gating mechanisms, variable selection, static covariate encoders, temporal processing, and prediction intervals" [7]. The gating mechanisms gives the model the ability to skip unused components, which makes the adaptable and flexible. The variable selection chooses relevant features as an input, which helps reduce noise and improve accuracy. The static covariate encoders integrate ”static features in the network”. The prediction intervals give a likely range of target values.

The TFT is meticulously designed for handling time series and predicting multiple time steps in the future, which means TFT is a multi-horizon model. It is also capable of taking as input multiple series with several features. This capability could be used, for example to process ECG from several different leads at the same time.

C. Time Series Transformer (TST)

The time series transformer (TST) is another variant of a transformer developed to process time series data. It was introduced for the first time in a 2021 [8]. This type of transformers are similar to the temporal fusion transformers. However, the architectures differ in the attention mechanism and the use of convolutional layers. The model is efficient in finding local patterns in the signal and using them to make predictions. A typical input in a TST is a one dimensional time series, for example an EEG signal. However, for training this type of model a vast dataset of clean signals is usually needed.

D. Convolutional Vision Transformer (CvT)

Convolutional neural networks (CNN) are types of networks that specialize in vision tasks. The main strength of this model type is their pattern recognition ability. The principle on which they function is finding simple patterns in the lower layer and more complex patterns in the higher layers. A picture is a 2D array of pixels, while a signal is 1D array of values. By following the same principles this model can also be used on 1D arrays. Convolutional vision transformers combine the convolutional properties of CNNs and attention of the transformer networks. They are introduced in 2021 in a study title ”CvT: Introducing Convolutions to Vision Transformers” [9]. The primary purpose of this type of model are vision tasks, however, they can be adapted to biosignal processing. Figure 3 displays the architecture of this type of model as given in the study it was introduced in.

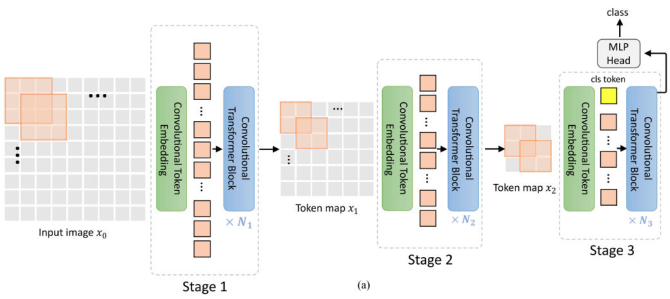

Fig. 3. Convolutional vision transformer architecture [9]

Unlike vision transformers (ViT), transformer-based model for image tasks without the use convolutions, the convolutional vision transformer uses convolutions. This type of model has a flexible input size and is adequate at generalisation, however, it’s generally used in vision tasks and might not be the best choice for biosignal processing.

E. Informer Transformer

The informer is a transformer model described in the paper ”Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting” [10] and a diagram of it can be seen at Figure 4. It’s a convolution based model with a recurrence mechanism called ProbSparse Self-Attention that allows it to handle irregularly spaced data. This model is specifically designed for time series forecasting.

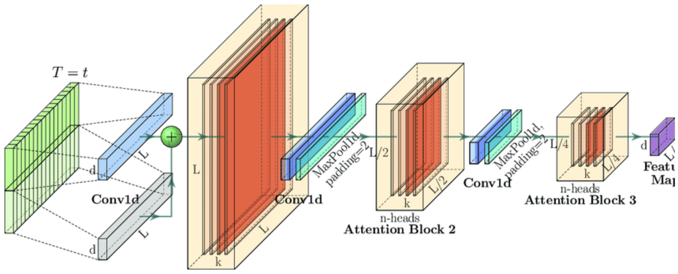

Fig. 4. Informer architecture [10]

The informer is capable of working with long time series and capturing long term dependencies with an irregular interval. However, they require a lot of computational resources for training and optimization.

Figure 5 shows a table in which the characteristics of the models are compared. The time series transformer is a simple model mostly used with a time series data with a fixed length input that mostly finds use in regular univariate time series forecasting. The temporal fusion architecture is a more complex model that can handle irregular and missing data. It’s a model that can process irregular multivariate time series. Informer is an even more complex and powerful model than TFT that can be adapted to a variety of problems. However needs a lot of data and computational resources to train. The convolutional vision transformer was originally designed for other tasks so it may not preform as well as model developed specifically for time series processing.

	Temporal Fusion	Time Series	Convolutional Vision	Informer
Data type	irregular multivariate time series	regular univariate or multivarate time series	images	irregular multivariate time series
Tasks	forecasting	forecasting	image classification, object detection	forecasting, imputation, anomaly detection
Pros	1. robust to missing data and noise 2. interpretability and transparency	1. simple model 2. tracking temporal dependecies	1. scalability 2. adaptibility to different domains	1. variable length input 2. interpretability
Cons	1. fixed length input 2. sensitivity to loss function and hyperparameters	1. restricted usability 2. fixed length input	1. may require large amount of data 2. not specifically made for time series data	1. computationally expensive 2. large amount of data is required

Fig. 5. Architecture Types Comparison

III. Transformer models in biosignal processing

Transformers are an effective model when dealing with NLP problems, however their capabilities can be used to solve a variety of deep-learning tasks. There have been several attempts made to use this architecture in conjuncture with biomedical data and as such there have been models developed for medical images analysis and segmentation [11], structuring health records [12], augmentation of medical datasets [13], etc. However in this study, our focus is only on biological signals (biosignals) which is a small subset of the term biomedical data that refers to ”space-time records of a biological events”. It should be noted that the transformer architecture is relatively new and its potential is still being investigated. In this section we describe works that utilize transformers for biological signals processing. Each of the signals is accompanied by an illustration courtesy of the UCI machine learning repository.

1) **Electrocardiogram (ECG)** An electrocardiogram denoted as ECG is a graph of the electrical activity of the heart over time. An ECG is commonly used to detect arrhythmia, which is a complication of the heart’s rhythm. There exist multiple types of arrhythmia that can cause the heart to beat too fast, too slow or with an irregular pattern. The detection of such conditions are a subject of interest for many studies working with this type of signals. An ECG is generated by analysing the electrical current detected by electrodes placed on the skin. Different views of the heart’s activity can by generated depending on the placement of the electrodes, called leads. A standard ECG is recorded by 12 leads, typically in a hospital setting. However, in the type of studies mentioned often times a single lead ECG is used to provide valuable information about the cardiovascular system. Figure 6 shows 15 seconds window of a typical ECG signal recorded from lead II, which is the most common choice for a single lead ECG.

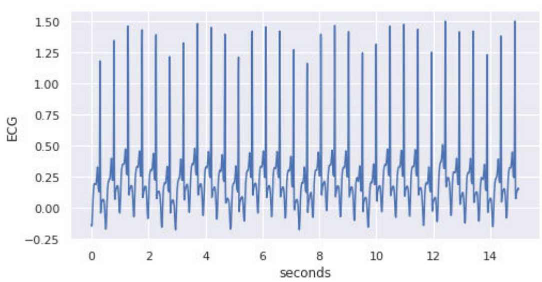

Fig. 6. Visualization of an ECG waveform (15 seconds)

In [14] transformers are used as part of a proposal for cardiac abnormalities detection and classification. The proposed network combines handcrafted ECG features and features extracted directly from the ECG. A hybrid architecture is used combining two separate parts with the qualities wide and deep, whose results are later concatenated as part of the model. The tasks of arrhythmia and heartbeat, detection and classification via transformer has been the theme of a few studies.

Another study [15] uses a modified transformer network for heartbeat classification. The architecture used in this paper is a fusing transformer, a model made by combining multiple smaller models. There are also several changes made to the design, among them the most important is the exclusion of the decoder part, since there is no need for translation when working with this type of signal.

A convolutional vision transformer has been used for congestive heart failure diagnosis in [16]. There are subtle differences in the ECG from patient to patient. These differences are the reason why detection of congestive heart failure is problematic. This neural network is capable of extracting high dimensional abstract features from the ECG and robust enough to function with a certain degree of noise.

Another research [17] uses the ECG for stress detection by creating a hybrid transformer by combining convolution layers with the attention mechanism. These architectures are considered viable because the convolution layers have successfully been used for automatic feature extraction for emotion recognition and the transformers are suitable to learning spatio-temporal representations.

2) **Electroencephalogram (EEG)** An electroencephalogram (EEG) is a graph of the electrical activity of the brain over time. Electrodes are placed on the scalp and is usually a non-invasive procedure. It’s used to diagnose conditions affecting the brain by detecting abnormal patterns. The human brain is a complex structure that has many parts with different responsibilities. The neurons communicate with each other via electrical impulses even during sleep, as such EEG is tool that can help diagnose sleep disorders. An EEG can be recorder from one or multiple channels. The term channel refers to simultaneous recordings from the same patient. The number of channels used varies depending on psychological activity the patient is performing and the disorder the test is trying to diagnose. An EEG from a single channel is shown on Figure 7.

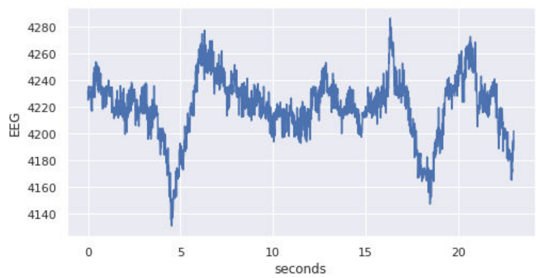

Fig. 7. Visualization of an EEG waveform

Transfer learning is a machine learning method that refers to the use of pre-trained models and adapting

them on the problem at hand. It's an useful tool that can drastically improve training times and it's doesn't require as much data, however it's not all powerful and there are several drawbacks when used incorrectly, such as a domain mismatch and negative transfer. There already exist NLP transform models that are pre-trained on language datasets and can easily be adapted to specific problems. A study [18] attempts to create a model on a vast amount of data of EEG signals, that can be later used for transfer learning in related fields.
In [19] the transformer architecture is used for emotion recognition from EEG. In this study an experiment was conducted and several variants of frameworks considered, which can be categorized as "spatial attention, temporal attention, sequential spatial-temporal attention and simultaneous spatial-temporal attention" [19]. The best results were achieved on the last proposed framework, simultaneous spatial-temporal attention.
Different visual stimuli elicit the activation of different pattern in the brain. Classification of such patterns is the goal of a study [20]. The study works with single-trial data and the proposed solution is a convolutional transformer. The input in this network is naturally EEG since it's a measurement of the activations that the model is trying to classify. The architecture is based on multi-headed self attention and it represents an overall improvement of the classification accuracy over the previous state-of-the-art models across five different visual stimuli tasks.

3) **Electromyography (EMG)** Electromyography is a technique for measuring the electrical activity of the skeletal muscles. The resulting graph of the activity over time is called electromyograph (EMG). The electrodes used for measuring the EMG differ from the one used for ECG or EEG analysis. Measuring a EMG involves inserting a needle electrode inside the muscle. This procedure can reveal dysfunctions of the muscles or the neurons controlling them, including motor problem, degenerative disorders and nerve injuries. Figure 8 illustrates the muscle activity of a hand at rest.

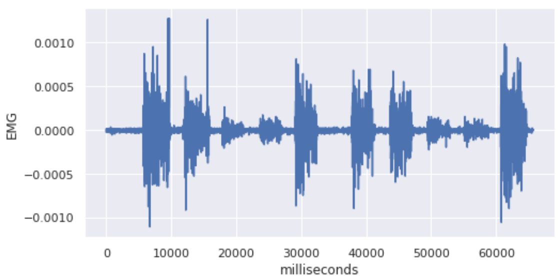

Fig. 8. Visualization of an EMG waveform of a hand at rest

Several studies have attempted to process the EMG signal, raw or filtered, as input to a transformer neural network in order to accomplish their goals. Robotics is field of study and development of autonomous machines, that draws inspiration from the human musculoskeletal system. EMG illustrates this system activity and as such can give valuable information that can later be utilized for some tasks, such as control of bionic hands. In [21] a model for hand motion classification is proposed. This model uses temporal multi-channel vision transformers. The goal in this paper is to reveal the potential of the model for dexterous robot control.
The prediction of human intention allows for better collaboration between a human and an AI assistant. One study [22] working in this domain particularly predicting the steering of a human driver. Several different driving postures and steering torques are considered. The resulting model is a multi-task time series transformer that utilizes upper limb EMG.

4) **Photoplethysmogram (PPG)** Photoplethysmogram (PPG) is used to detect "volumetric changes in blood in peripheral circulation". It's both low cost and easy to measure, since some approaches only require a contact with the surface skin. PPG finds broad use in medicine and has application in calculating or estimating various biological features, blood oxygen saturation, respiration, heart rate, vascular assessment, cardiac output, blood pressure, etc. Figure 9 displays a PPG signal from a patient sampled at 125 Hz.

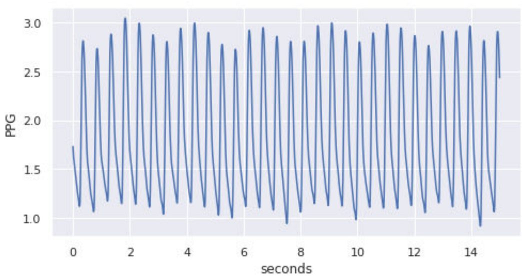

Fig. 9. Visualization of a PPG waveform (15 seconds)

The problem of cardiovascular diseases detection is investigated in [23]. The goal of this study is to utilize a transformer for PPG to ECG reconstruction. Continuously monitoring of ECG for cardiac irregularities isn't always feasible. To overcome this problem a novel architecture called "Performer", was proposed to reconstruct the ECG signal. Generally PPG is easier to measure than ECG, since it only involves a single point contact with the skin.
In [24] a KD-Informer architecture is proposed for cuff-less blood pressure (BP) estimation. Monitoring BP is vital for early detection and prevention of many cardiovascular diseases. Cuff-less BP estimation is a well researched problem and the most successful models use a combinations of ECG and PPG, since the pertinent information is contained in the difference between these two signals. However, PPG only model have been devel-

oped, such as the one in this study. A transfer learning approach is implemented, by first training the model on a small high quality dataset before transferring it to the target dataset.

IV. Conclusion

Transformers are neural network models that show promising results in several fields and their potential is still being investigated. They are usually faster to train and address several deficiencies of existing models. These models have been used for many real world applications, language, music, speech, vision, biomedical data, and others. Biological signals are a sequence of temporal readings for biological events. There have been several attempts to use this model type to solve tasks in this domain. Using the initial architecture as a baseline several specialized models have been developed to process temporal data such as a signal. A quick overview is made on the basic transformer model and it's hybrid variants: the temporal fusion, the time series, the convolutional vision transformer, and informer, as suitable models for processing of biological signals.

This review was done as a part of our research in method for generalised blood pressure estimation from ECG and PPG signals. This is a well researched problem that has been studied utilizing different machine learning and deep learning methods. We believe that the relatively new transformer architecture may be utilized to offer new insight to this task. The review focuses specifically on transformers that work with biological signals, monitored recordings taken from the human body sampled at a regular interval. It's important to note that the human body is a very complex system as such there may be large variations between the signals recorded from different patients. Our opinion is that the models considered hold the potential to improve on the results achieved by more traditional approaches.

Acknowledgment

This paper has been written thanks to the support of the "Smart Patch for Life Support Systems" - NATO project G5825 SP4LIFE and by the National project IBS4LIFE of Faculty of Computer Science and Engineering, at Ss. Cyril and Methodius University in Skopje.

References

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," Advances in neural information processing systems, vol. 30, 2017.

[2] Mani Wadhwa, "Seq2seq model in machine learning," [retrieved: May, 2023]. [Online]. Available: https://www.geeksforgeeks.org/seq2seq-model-in-machine-learning/

[3] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," Neural computation, vol. 31, no. 7, pp. 1235–1270, 2019.

[4] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghayvat, "Cnn variants for computer vision: history, architecture, application, challenges and future scope," Electronics, vol. 10, no. 20, p. 2470, 2021.

[5] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.

[6] OpenAi, "Chatgpt," [retrieved: Mar, 2023]. [Online]. Available: https://openai.com/blog/chatgpt

[7] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," International Journal of Forecasting, vol. 37, no. 4, pp. 1748–1764, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207021000637

[8] G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, and C. Eickhoff, "A transformer-based framework for multivariate time series representation learning," in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 2114–2124.

[9] H. Wu, B. Xiao, N. Codella, M. Liu, X. Dai, L. Yuan, and L. Zhang, "Cvt: Introducing convolutions to vision transformers," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 22–31.

[10] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in Proceedings of the AAAI conference on artificial intelligence, vol. 35, no. 12, 2021, pp. 11 106–11 115.

[11] Y. Gao, M. Zhou, and D. N. Metaxas, "Utnet: a hybrid transformer architecture for medical image segmentation," in Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24. Springer, 2021, pp. 61–71.

[12] E. Choi, Z. Xu, Y. Li, M. Dusenberry, G. Flores, E. Xue, and A. Dai, "Learning the graphical structure of electronic health records with graph convolutional transformer," in Proceedings of the AAAI conference on artificial intelligence, vol. 34, no. 01, 2020, pp. 606–613.

[13] A. Amin-Nejad, J. Ive, and S. Velupillai, "Exploring transformer text generation for medical dataset augmentation," in Proceedings of the Twelfth Language Resources and Evaluation Conference, 2020, pp. 4699–4708.

[14] A. Natarajan, Y. Chang, S. Mariani, A. Rahman, G. Boverman, S. Vij, and J. Rubin, "A wide and deep transformer neural network for 12-lead ecg classification," in 2020 Computing in Cardiology. IEEE, 2020, pp. 1–4.

[15] G. Yan, S. Liang, Y. Zhang, and F. Liu, "Fusing transformer model with temporal features for ecg heartbeat classification," in 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2019, pp. 898–905.

[16] T. Liu, Y. Si, W. Yang, J. Huang, Y. Yu, G. Zhang, and R. Zhou, "Inter-patient congestive heart failure detection using ecg-convolution-vision transformer network," Sensors, vol. 22, no. 9, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/9/3283

[17] B. Behinaein, A. Bhatti, D. Rodenburg, P. Hungler, and A. Etemad, "A transformer architecture for stress detection from ecg," in 2021 International Symposium on Wearable Computers, 2021, pp. 132–134.

[18] D. Kostas, S. Aroca-Ouellette, and F. Rudzicz, "Bendr: using transformers and a contrastive self-supervised learning task to learn from massive amounts of eeg data," Frontiers in Human Neuroscience, vol. 15, p. 653659, 2021.

[19] J. Liu, L. Zhang, H. Wu, and H. Zhao, "Transformers for eeg emotion recognition," arXiv preprint arXiv:2110.06553, 2021.

[20] S. Bagchi and D. R. Bathula, "Eeg-convtransformer for single-trial eeg-based visual stimulus classification," Pattern Recognition, vol. 129, p. 108757, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320322002382

[21] R. V. Godoy, G. J. G. Lahr, A. Dwivedi, T. J. S. Reis, P. H. Polegato, M. Becker, G. A. P. Caurin, and M. Liarokapis, "Electromyography-based, robust hand motion classification employing temporal multi-channel vision transformers," IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 10 200–10 207, 2022.

[22] Y. Xing, W. Li, X. Mo, and C. Lv, "Multi-task driver steering behaviour modeling using time-series transformer," arXiv preprint arXiv:2207.00484, 2022.

[23] E. Lan, "Performer: A novel ppg-to-ecg reconstruction transformer for a digital biomarker of cardiovascular disease detection," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2023, pp. 1991–1999.

[24] C. Ma, P. Zhang, F. Song, Y. Sun, G. Fan, T. Zhang, Y. Feng, and G. Zhang, "Kd-informer: Cuff-less continuous blood pressure waveform estimation approach based on single photoplethysmography," IEEE Journal of Biomedical and Health Informatics, pp. 1–14, 2022.