Article

Tetanus Severity Classification in Low-Middle Income Countries through ECG Wearable Sensors and a 1D-Vision Transformer

Ping Lu [1,*], Zihao Wang [1], Hai Duong Ha Thi [2,3], Ho Bich Hai [2], VITAL Consortium [†], Louise Thwaites [2,4] and David A. Clifton [1,5]

1   Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK;
    zihao.wang@st-annes.ox.ac.uk (Z.W.); david.clifton@eng.ox.ac.uk (D.A.C.)
2   Oxford University Clinical Research Unit, Ho Chi Minh City 700000, Vietnam;
    duonghth@oucru.org (H.D.H.T.); haihb@oucru.org (H.B.H.); lthwaites@oucru.org (L.T.)
3   Hospital for Tropical Diseases, Ho Chi Minh City 700000, Vietnam
4   Centre for Tropical Medicine and Global Health, University of Oxford, Oxford OX3 7LG, UK
5   Oxford Suzhou Centre for Advanced Research, Suzhou 215123, China
*   Correspondence: ping.lu@eng.ox.ac.uk
†   Membership of the VITAL Consortium is provided in the Acknowledgments.

Abstract: Tetanus, a life-threatening bacterial infection prevalent in low- and middle-income countries like Vietnam, impacts the nervous system, causing muscle stiffness and spasms. Severe tetanus often involves dysfunction of the autonomic nervous system (ANS). Timely detection and effective ANS dysfunction management require continuous vital sign monitoring, traditionally performed using bedside monitors. However, wearable electrocardiogram (ECG) sensors offer a more cost-effective and user-friendly alternative. While machine learning-based ECG analysis can aid in tetanus severity classification, existing methods are excessively time-consuming. Our previous studies have investigated the improvement of tetanus severity classification using ECG time series imaging. In this study, our aim is to explore an alternative method using ECG data without relying on time series imaging as an input, with the aim of achieving comparable or improved performance. To address this, we propose a novel approach using a 1D-Vision Transformer, a pioneering method for classifying tetanus severity by extracting crucial global information from 1D ECG signals. Compared to 1D-CNN, 2D-CNN, and 2D-CNN + Dual Attention, our model achieves better results, boasting an F1 score of 0.77 ± 0.06, precision of 0.70 ± 0.09, recall of 0.89 ± 0.13, specificity of 0.78 ± 0.12, accuracy of 0.82 ± 0.06 and AUC of 0.84 ± 0.05.

Keywords: tetanus; electrocardiogram; classification; Transformer; Vision Transformer

1. Introduction

Tetanus, a life-threatening infectious disease caused by the bacterium Clostridium tetani, is most prevalent in low- and middle-income countries (LMICs). Although the disease occurs in high-income countries, it is most prevalent in low- and middle-income countries. The disease is common in settings characterised by poor hygiene, limited access to health care and inadequate immunisation programmes [1–3]. The lack of advanced medical equipment and health workers poses challenges in the management of complications of tetanus, such as autonomic nervous system dysfunction (ANSD) and laryngeal spasms, resulting in increased mortality rates, as discussed in [4].

The tetanus toxin disrupts signalling at synapses in the central nervous system, causing agonising muscle spasms and rigidity. In severe cases, its effects on the autonomic nervous system (ANS) can cause cardiovascular instability. Approximately 50% of patients progress to severe disease within 2 to 5 days, and if left untreated, these muscle spasms can affect breathing, requiring the use of powerful muscle relaxants and mechanical ventilation. In mechanically ventilated facilities, this ANS dysfunction is the leading cause of mortality

in tetanus patients. However, effective management of this condition remains a major challenge. Early detection of severe tetanus is of paramount importance, as it allows timely intervention and optimises resource allocation [5].

In clinical settings with high patient volumes or limited staff experience, achieving accurate classification can be a daunting task. Advanced continuous monitoring systems and the presence of sufficient health workers in high-income countries have been associated with improved outcomes for patients with tetanus [6,7].

In many resource-limited settings, the availability of close monitoring and timely emergency intervention is typically limited to high-acuity wards or intensive care units, as these facilities have the staff and equipment to provide such services. This increased demand for intensive care in LMICs places an additional burden on already limited resources and may ultimately lead to poorer outcomes for individuals requiring such specialised care [6,8,9]. In addition, a significant number of patients in LMICs, such as Vietnam, are burdened with out-of-pocket medical expenses. As a result, the additional costs associated with ICU care are significantly higher than those associated with standard ward care. Existing research has provided insights into the direct medical expenditure for ICU patients with tetanus, dengue and sepsis in Vietnam [6,8,9].

In resource-limited settings, the use of low-cost wearable sensors is emerging as a promising alternative for tetanus case management. These wearable sensors are wireless, compact and lightweight. Their primary function is to provide real-time, continuous monitoring of vital signs, with the overall goal of enabling early detection of patient deterioration [6,10]. Our previous research has highlighted that the use of ECG monitoring alone may be sufficient to classify the severity of tetanus [11,12]. It is worth noting, however, that the practical implementation of affordable wearable sensors still faces challenges, mainly due to inherent inaccuracies in the continuous physiological data they collect. These inaccuracies arise mainly from data gaps and the significant noise introduced by various factors, thereby undermining their reliability [6].

Our previous studies have investigated the improvement of tetanus severity classification using ECG time series imaging. Our aim in this study is to investigate an alternative method using ECG data without relying on time series imaging as an input, with the aim of achieving comparable or improved performance. This study employs ECG data obtained from wearable sensors utilised in an ICU in Vietnam and suggests a rapid triage tool, developed through deep learning techniques, to categorise tetanus severity based on the Ablett score. We choose a 1D-Vision Transformer to extract the global features of the ECG. The proposed 1D-Vision Transformer outperforms the previous 1D and 2D Convolution Neural Network (CNN) and 2D CNN with Dual Attention mechanisms.

This study provides the following contributions:

- We present a 1D-Vision Transformer model equipped with a self-attention mechanism that enables it to evaluate and assign importance to elements within the input ECG time series data while processing each specific element.
- This is the first time that a 1D Transformer-based method has been investigated to classify the severity of tetanus in LMICs. The proposed 1D-Vision Transformer outperforms the performance of the state-of-the-art 1D and 2D CNN methods in tetanus classification. It promises to improve clinical decision making in resource-constrained settings.
- We illustrate the relationship between the ECG signal and the proposed AI model's decision using attention scores, showing how the signal exerts varying degrees of influence through different weights.

2. Related Work

The healthcare landscape is being transformed by artificial intelligence [13–19]. In traditional machine learning (ML) methodologies, manual feature extraction is often required. For instance, datasets may necessitate the manual extraction of RR intervals, as exemplified in [20]. Support Vector Machines (SVMs) have been employed to automatically

gauge the degree of autonomic nervous system (ANS) dysfunction in tetanus patients, as detailed in [21]. However, it is worth noting that deep learning (DL) methods have exhibited superior performance compared to conventional ML techniques like SVMs, as highlighted in [22].

Transformers represent a remarkable advance in the field of computer vision and image analysis [12,23–27]. The field of 2D deep learning with time series imaging has been actively explored for tetanus severity classification, as evidenced by our previous works [11,12,28]. Our previous studies have investigated the improvement of tetanus severity classification using ECG time series imaging. In our previous study [11], we introduced a two-dimensional (2D) convolutional neural network (CNN) augmented with a channel-wise attention mechanism for binary ECG signal classification. Lu et al. [12] proposed a groundbreaking hybrid CNN-Transformer model for tetanus severity level classification using a wearable ECG. This innovative model combines the ability to capture local features via CNN and global features via the Transformer architecture. The input to this model consists of time series images specifically derived from one-dimensional ECG signals using a spectrogram representation. They designed a square time series image format that serves as a bridge between biomedical signals and advanced computer vision algorithms. In addition, Lu et al. [28] introduced a 2D-WinSpatt-Net, a novel Vision Transformer that incorporates both local spatial window self-attention and global spatial self-attention mechanisms. This is the first time that continuous wavelet transform (CWT) has been used for the representation of tetanus ECG information in the form of time series images. This innovative approach resulted in improved tetanus severity classification accuracy, even with shorter tetanus ECG signals of only 20 s. This is a remarkable achievement, especially when compared to the 60 s and 5 min ECG recordings commonly used for heart rate variability.

3. Method

The proposed framework includes both data pre-processing and feature extraction using a 1D-Vision Transformer. Figure 1 provides an overview of the framework and illustrates its role in tetanus severity classification using the 1D-Vision Transformer method.
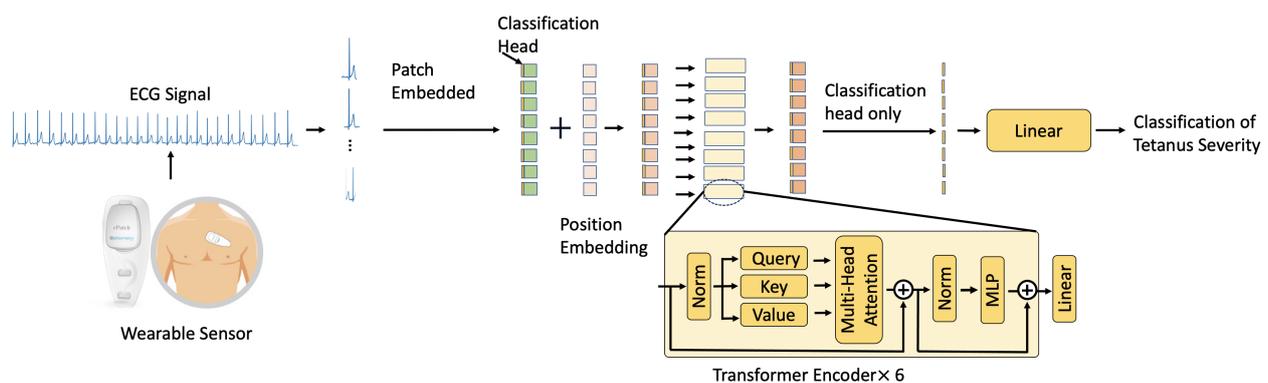

Figure 1. Overview of the tetanus severity classification framework using the proposed method.

3.1. Data Pre-Processing

During the initial data processing phase, our primary goal was to remove noise from the ECG signal. There are two main types of noise that can interfere with ECG signal analysis, as described in [29], low-frequency and high-frequency noise. We acquired single-lead ECG signals using an low-cost, portable monitoring device. To improve data quality, we used a Butterworth filter to remove unwanted noise. The high-pass filter was set to a cut-off frequency of 0.05 Hz, and the low-pass filter was set to a cut-off frequency of 100 Hz. We implemented this data pre-processing step using the SciPy package, as described in [30].

3.2. 1D-Vision Transformer

We segmented the ECG signal data, denoted as **a**, into flattened non-overlapping patches, represented as $\tilde{\mathbf{a}}_p \in \mathrm{R}^{N \times (P^2 \times C)}$. Here, C represents the number of channels, N corresponds to the total number of patches ($N = \frac{H}{P} \times \frac{W}{P}$), and P indicates the patch size. Subsequently, we transformed these patches into a D-dimensional embedding space through a trainable linear projection. To preserve the spatial information of these extracted patches, we combined the position embeddings with patched embeddings, as outlined below.

$$\mathbf{c}_0 = [\tilde{\mathbf{a}}_p^1 \mathbf{E}; \tilde{\mathbf{a}}_p^2 \mathbf{E}; \ldots; \tilde{\mathbf{a}}_p^N \mathbf{E}] + \mathbf{E}_{pos}, \tag{1}$$

where $\mathbf{E} \in \mathrm{R}^{(P^2 \times C) \times D}$ represents the projected patch embedding, and $\mathbf{E}_{pos} \in \mathrm{R}^{N \times D}$ stands for the learnable position embedding.

After creating the embeddings, we proceeded to apply L Transformer layers. Within each Transformer layer, as described in [31,32], three principal components were noted: Multi-Head Self-Attention (MSA), Multi-Layer Perceptron (MLP), and Layer Normalisation (LNorm). The resulting output for the l-th layer can be expressed as follows:

$$\mathbf{c}_l' = MSA(LNorm(\mathbf{c}_{l-1})) + \mathbf{c}_{l-1}, \qquad l = 1, \ldots, L, \tag{2}$$

$$\mathbf{c}_l = MLP(LNorm(\mathbf{c}_l')) + \mathbf{c}_l', \qquad l = 1, \ldots, L. \tag{3}$$

Multi-Head Self-Attention

The input matrix $\mathbf{m} \in \mathrm{R}^{n \times d}$ was subjected to a transformation, resulting in the generation of three distinct vectors: queries $\mathbf{Que} \in \mathrm{R}^{n \times d_k}$, keys $\mathbf{Key} \in \mathrm{R}^{n \times d_k}$, and values $\mathbf{Val} \in \mathrm{R}^{n \times d_v}$. Here, d_k denotes the dimensions of the queries and keys, while d_v represents the dimensions of the values. The mechanism of the scaled dot-product attention, as elucidated in [31], can be expressed through the following equation:

$$Att(\mathbf{Que}, \mathbf{Key}, \mathbf{Val}) = softmax\left(\frac{\mathbf{Que}\mathbf{Key}^T}{\sqrt{d_k}}\right)\mathbf{Val}, \tag{4}$$

Here, the term $\frac{1}{\sqrt{d_k}}$ acts as a scale factor, serving to maintain stable gradients by preventing the softmax function from venturing into regions where gradients become excessively small.

The Multi-Head Self-Attention (MSA) constitutes a fundamental component within the Transformer architecture. It comprises n parallel self-attention (SA) heads, each of which dissects the **Que**, **Key**, and **Val** matrices into distinct subspaces, concurrently executing the scaled dot-product attention operation. Subsequently, the outputs from each head are concatenated and transformed into the final MSA output through a linear projection. The corresponding formula is presented as follows:

$$MSA(\mathbf{Que}, \mathbf{Key}, \mathbf{Val}) = Concatenate(Head_1, \ldots, Head_h)W^o, \tag{5}$$

$$Head_i = Att(\mathbf{Que}W_i^Q, \mathbf{Key}W_i^K, \mathbf{Val}W_i^V), \tag{6}$$

where W^o denotes the multi-headed trainable parameter weights.

4. Experiments

4.1. Recording ECG Data in Tetanus Patients

The dataset was collected from patients with tetanus admitted to the Hospital for Tropical Diseases, situated in Ho Chi Minh City, Vietnam [6]. In our research, we used ECG data collected from people who had been diagnosed with tetanus. We used the ePatch V.1.0, a low-cost portable monitor manufactured by BioTelemetry, Malvern, PA, USA, as our monitoring device (see Figure 1). The ePatch, which weighs 7 g (ePatch information is

available at https://www.philips.co.uk/healthcare/resources/landing/epatch, accessed on 8 January 2024) was securely attached to the patient's chest skin to ensure reliable adhesion. The device records two channels of ECG data, and the sampling rate is 256 Hz.

The two channels of the ePatch device (referred to as channel 1 and 2) do not directly correlate with the ECG leads 1 and 2 of the conventional bedside monitor, as mentioned in [12]. Clinical staff used the Ablett scoring system to grade severity as follows: grades 1 and 2 (no or mild spasms) define "mild" disease, and grades 3 and 4 (spasms interfering with respiration with/without autonomic nervous system dysfunction) define "severe" disease. The details of the tetanus data information can be found in [11,12].

4.2. Implementation Details

Pre-processing. We extracted 30 ECG time series, each with a duration of 60 s, from every ECG example file. This resulted in a training dataset comprising a total of 4230 ECG time series, with 2370 samples indicative of mild tetanus and 1860 samples indicative of severe tetanus. Our validation dataset consisted of 540 ECG time series (270 mild cases and 270 severe cases), while the test dataset comprised 570 ECG time series (360 mild cases and 210 severe cases). The categorization of mild and severe tetanus cases was performed by clinicians.

Experimental Setup. Based on our experiments, the following selected hyperparameters of the proposed 1D-Vision Transformer achieved optimal results (see Table 1). Each 1-min ECG dataset comprises 15,360 data points. We applied a 1D convolution to the input signal (60-s ECG), producing 384 sets, each containing 320 data points. We then rearranged the original tensor according to the desired order, resulting in a new multidimensional rotated tensor with 320 sets, each containing 384 data points.

Table 1. Employed parameters of the proposed 1D-Vision Transformer.

Parameters		
in_channels	1	the number of channels of the image
patch size	48	the size (resolution) of each patch
num_transformer_layer	6	the number of Transformer blocks
embed_dim	384	the embedding dimension
Mlp_size	1024	the number of neurons in the hidden layer
num_heads	6	the number of heads
mlp_drouppout	0.1	the dropout for the MLP layers
embedding_dropout	0.1	the dropout for the embeddings
num_class	2	the number of classes

The model was trained with the following specifications: 100 epochs using the Adam optimiser, a learning rate of 0.001 and a batch size of 32. The torch.nn.CrossEntropyLossis was selected as the loss function. The implementation of the suggested 1D-Vision Transformer was carried out in Python 3.7 using PyTorch. The experiments were carried out on hardware equipped with the NVIDIA RTX A6000 48 GB GPU.

4.3. Baselines

In our study, we performed a comparative analysis between the 1D-Vision Transformer, our proposed method, and three different baseline approaches. These baseline methods included two 2D deep learning techniques introduced by Lu et al. [11] and a 1D-CNN method.

4.4. Evaluation Metrics

In this study, we utilised multiple performance metrics to assess the effectiveness of the binary classification task. These metrics were the F1-score, precision, recall, specificity, accuracy [22] and the area under the curve (AUC) [33]. To ensure the reliability of our

findings, each model was executed five times, with subsequent computation and reporting of the performance metric averages and standard deviations using an independent test dataset. A higher AUC value serves as an indicator of the model's superior ability to accurately distinguish between severe and mild cases of tetanus.

5. Experimental Results

5.1. Data Pre-Processing Analysis

We delved into the realm of pre-processing techniques in ECG analysis, recognising that noise removal at this stage can significantly improve classification performance. Our primary objective was to quantify the positive impact of these pre-processing steps. As shown in Table 2, we observed a significant improvement in F1-score (0.03 increase), precision (0.06 increase), specificity (0.05 increase), AUC (0.03 increase) and accuracy (0.04 increase) following the removal of data noise in the input to our proposed 1D-Vision Transformer. This demonstrates that the application of the Butterworth filter in the pre-processing step effectively removes unwanted noise, resulting in higher quality ECG data as the input to our model. This in turn leads to improved classification accuracy for tetanus severity.

Table 2. A quantitative analysis of the proposed 1D-Vision Transformer, comparing with and without data pre-processing. The outcomes are presented as the mean $\pm$ standard deviation, highlighting the top performance in bold.

1D-Vision Transformer	F1-Score	Precision	Recall	Specificity	Accuracy	AUC
without data pre-processing	0.74 ± 0.04	0.64 ± 0.07	$\mathbf{0.89 \pm 0.04}$	0.73 ± 0.08	0.78 ± 0.05	0.81 ± 0.03
with data pre-processing	$\mathbf{0.77 \pm 0.06}$	$\mathbf{0.70 \pm 0.09}$	0.89 ± 0.13	$\mathbf{0.78 \pm 0.12}$	$\mathbf{0.82 \pm 0.06}$	$\mathbf{0.84 \pm 0.05}$

5.2. Comparisons

We evaluated the proposed 1D-Vision Transformer by comparing it to three different deep learning techniques, including models using 1D (ECG signal) and 2D (time series image) data as input. In light of the experimental outcomes presented in Table 3, the 1D-Vision Transformer method using ECG (non-imaged data representation) as input achieves the best performance in diagnosing tetanus. The 1D-Vision Transformer outperforms the 1D-CNN.

Table 3. A quantitative analysis of the proposed 1D-Vision Transformer, compared to baseline methods that employ either 60-s time series image [11] or original 60-s ECG as input. The outcomes are presented as the mean $\pm$ standard deviation, highlighting the top performance in bold.

Method	**The Time Series image as the Input**					
	F1-Score	Precision	Recall	Specificity	Accuracy	AUC
2D-CNN [11]	0.61 ± 0.14	0.68 ± 0.07	0.57 ± 0.19	0.85 ± 0.02	0.75 ± 0.07	0.72 ± 0.09
2D-CNN + Dual Attention [11]	0.65 ± 0.19	$\mathbf{0.71 \pm 0.17}$	0.61 ± 0.21	$\mathbf{0.86 \pm 0.09}$	0.76 ± 0.11	0.74 ± 0.13
Method	**The ECG as the Input**					
	F1-Score	Precision	Recall	Specificity	Accuracy	AUC
1D-CNN [11]	0.65 ± 0.14	0.61 ± 0.05	0.77 ± 0.25	0.70 ± 0.13	0.73 ± 0.05	0.74 ± 0.08
Proposed 1D-Vision Transformer	$\mathbf{0.77 \pm 0.06}$	0.70 ± 0.09	$\mathbf{0.89 \pm 0.13}$	0.78 ± 0.12	$\mathbf{0.82 \pm 0.06}$	$\mathbf{0.84 \pm 0.05}$

5.3. Interpretable ECG

We used the attention scores to interpret which part of the ECG signal the model is focusing on for the classification of tetanus severity. We represent high scores with a darker shade of red, indicating that the ECG region coloured in darker red has a greater influence on the model's decision. Figure 2 displays a 60 s ECG example along with the attention scores that the proposed model relies on when categorising mild tetanus.
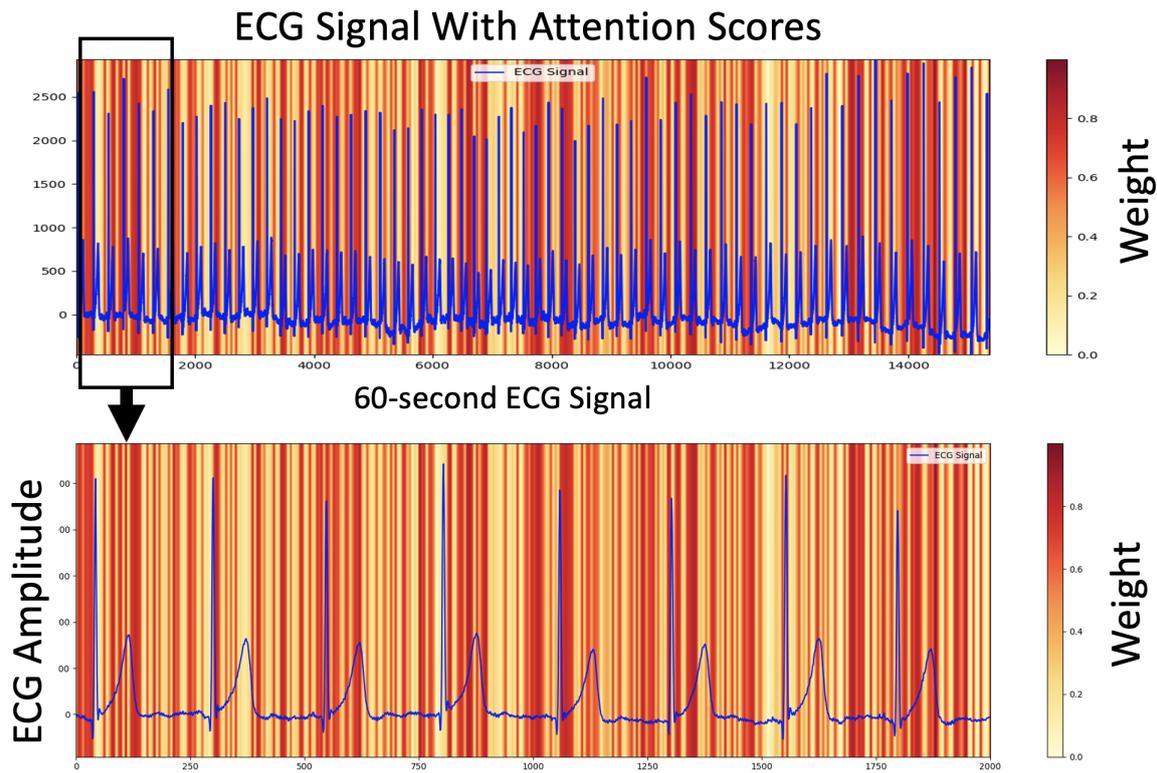
Figure 2. An example ECG waveform with corresponding attention scores: darker red signifies a greater influence on the proposed 1D-Vision Transformer model's categorisation of mild tetanus.

5.4. Misclassification

We used the same strategy to generate confusion matrices, as detailed by Lu et al. [11]. The confusion matrices depicted in Figure 3 provide a comprehensive overview of the performance of each method in our experiments as well as the types of misclassifications they exhibit between the mild and severe levels. The successful detection rate for severe tetanus diagnosis reached 165 after employing the 1D-Vision Transformer, which represents the highest accuracy achieved among these deep learning methods. The performance of the 1D-Vision Transformer surpassed that of the 1D-CNN in classifying mild and severe tetanus cases.
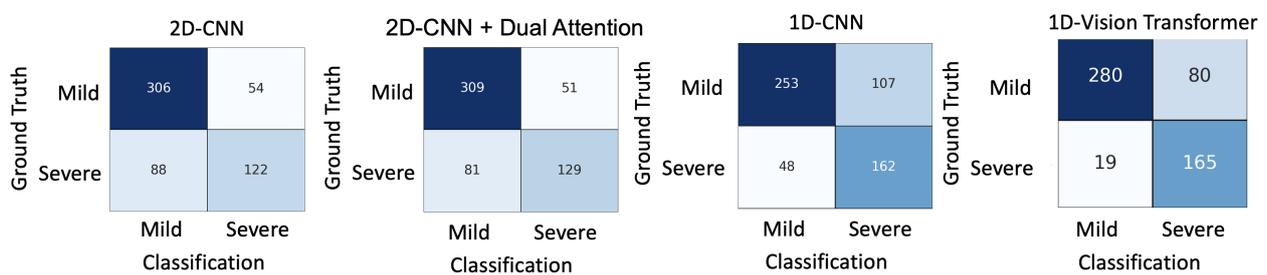
Figure 3. The confusion matrices for tetanus severity classification using different deep learning methods: 2D-CNN and 2D-CNN + Dual Attention with 60 s window log spectrograms as the inputs (without downsampling); and 1D-CNN and 1D-Vision Transformer with 60 s ECG data as the inputs, representing an image-free data representation.

6. Discussion

The proposed 1D-Vision Transformer is equipped with a self-attention mechanism that enables it to consider the importance of elements in the input ECG time series data when processing a particular element. This allows it to capture global relationships and

dependencies within the data. In other words, it can understand how different parts of the input ECG time series data relate to each other, regardless of their position. We enhanced the classification of tetanus severity levels through the utilisation of a 1D deep learning approach, surpassing the performance of 1D-CNN. In our earlier investigations, as detailed in [11,12,28], we discovered that representing 1D ECG as time series images, serving as the input for 2D deep learning methods, yielded superior performance. While the performance of the proposed 1D-Vision Transformer does not surpass that of [11,12,28], it represents a promising first step in exploring the field of 1D deep learning approaches for tetanus severity level diagnosis. Our goal is to improve the 1D-Vision Transformer for classification of mild or severe tetanus in future research efforts. In addition, the 1D-Vision Transformer can serve as a benchmark for our future 1D deep learning approaches. Furthermore, the proposed method can be applied to other biomedical signal analyses, such as sepsis or dengue.

7. Conclusions

We have proposed a 1D-Vision Transformer for tetanus severity classification. Our experimental results clearly demonstrate the superiority of our proposed method over other advanced deep learning approaches in the context of tetanus severity classification. This deep learning framework promises to significantly improve clinical decision making and streamline the allocation of limited healthcare resources, particularly in low- and middle-income countries (LMICs). In our future endeavours, we will strive to further enhance the novelty and effectiveness of the 1D-Vision Transformer-based method. Moreover, the versatility of the 1D-Vision Transformer allows its application in various classification tasks, including those involving time series data.

Thao. University of Oxford (alphabetic order by surname): Natasha Ali, David Clifton, Mike English, Ping Lu, Jacob McKnight, Chris Paton, and Tingting Zhu. Imperial College London (alphabetic order by surname): Pantelis Georgiou, Bernard Hernandez Perez, Kerri Hill-Cawthorne, Alison Holmes, Stefan Karolcik, Damien Ming, Nicolas Moser, and Jesus Rodriguez Manzano. King's College London (alphabetic order by surname): Liane Canas, Alberto Gomez, Hamideh Kerdegari, Andrew King, Marc Modat, and Reza Razavi. University of Ulm (alphabetic order by surname): Walter Karlen. The University of Melbourne (alphabetic order by surname): Linda Denehy, and Thomas Rollinson. Mahidol Oxford Tropical Medicine Research Unit (MORU) (alphabetic order by surname): Luigi Pisani, and Marcus Schultz.

References

1. Thwaites, C.L.; Yen, L.M.; Glover, C.; Tuan, P.Q.; Nga, N.T.N.; Parry, J.; Loan, H.T.; Bethell, D.; Day, N.P.J.; White, N.J.; et al. Predicting the clinical outcome of tetanus: The tetanus severity score. *Trop. Med. Int. Health* **2006**, *11*, 279–287. [PubMed]
2. Yen, L.M.; Thwaites, C.L. Tetanus. *Lancet* **2019**, *393*, 1657–1668. [CrossRef] [PubMed]
3. Thuy, D.B.; Campbell, J.I.; Thanh, T.T.; Thuy, C.T.; Loan, H.T.; Van Hao, N.; Minh, Y.L.; Boni, M.F.; Thwaites, C.L. Tetanus in southern Vietnam: Current situation. *Am. J. Trop. Med. Hyg.* **2017**, *96*, 93. [CrossRef] [PubMed]
4. Li, J.; Liu, Z.; Yu, C.; Tan, K.; Gui, S.; Zhang, S.; Shen, Y. Global epidemiology and burden of tetanus from 1990 to 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Int. J. Infect. Dis.* **2023**, *132*, 118–126. [PubMed]
5. The Importance of Diagnostic Tests in Fighting Infectious Diseases. Available online: https://www.lifechanginginnovation.org/medtech-facts/importance-diagnostic-tests-fighting-infectious-diseases.html (accessed on 6 October 2021).
6. Van H.M.T.; Hao, N.V.; Phan Nguyen Quoc, K.; Hai, H.B.; Khoa, L.D.V.; Yen, L.M.; Nhat, P.T.H.; Hai Duong, H.T.; Thuy, D.B.; Zhu, T.; et al. Vital sign monitoring using wearable devices in a Vietnamese intensive care unit. *BMJ Innov.* **2021**, *7*, S7–S11 .
7. Mahieu, R.; Reydel, T.; Maamar, A.; Tadié, J.M.; Jamet, A.; Thille, A.W.; Chudeau, N.; Huntzinger, J.; Grangé, S.; Beduneau, G.; et al. Admission of tetanus patients to the ICU: A retrospective multicentre study. *Ann. Intensive Care* **2017**, *7*, 1–7. [CrossRef]
8. Hung, T.M.; Van Hao, N.; Yen, L.M.; McBride, A.; Dat, V.Q.; van Doorn, H.R.; Loan, H.T.; Phong, N.T.; Llewelyn, M.J.; Nadjm, B.; et al. Direct Medical Costs of Tetanus, Dengue, and Sepsis Patients in an Intensive Care Unit in Vietnam. *Front. Public Health* **2022**, *10*, 1665. [CrossRef]
9. Hung, T.M.; Clapham, H.E.; Bettis, A.A.; Cuong, H.Q.; Thwaites, G.E.; Wills, B.A.; Boni, M.F.; Turner, H.C. The estimates of the health and economic burden of dengue in Vietnam. *Trends Parasitol.* **2018**, *34*, 904–918. [CrossRef]
10. Joshi, M.; Ashrafian, H.; Aufegger, L.; Khan, S.; Arora, S.; Cooke, G.; Darzi, A. Wearable sensors to improve detection of patient deterioration. *Expert Rev. Med Devices* **2019**, *16*, 145–154. [CrossRef]
11. Lu, P.; Ghiasi, S.; Hagenah, J.; Hai, H.B.; Hao, N.V.; Khanh, P.N.Q.; Khoa, L.D.V.; VITAL Consortium; Thwaites, L.; Clifton, D.A.; et al. Classification of Tetanus Severity in Intensive-Care Settings for Low-Income Countries Using Wearable Sensing. *Sensors* **2022**, *22*, 6554. [CrossRef]
12. Lu, P.; Wang, C.; Hagenah, J.; Ghiasi, S.; Zhu, T.; Thwaites, L.; Clifton, D.A. Improving Classification of Tetanus Severity for Patients in Low-Middle Income Countries Wearing ECG Sensors by Using a CNN-Transformer Network. *IEEE Trans. Biomed. Eng.* **2022**, *70*, 1340–1350. [CrossRef] [PubMed]
13. Lu, H.; Clifton, D.; Lu, P.; Hirst, J.; MacKillop, L. A Deep Learning Approach of Blood Glucose Predictive Monitoring for Women with Gestational Diabetes. *Res. Sq.* **2023**. . [CrossRef]
14. Chauhan, V.K.; Molaei, S.; Tania, M.H.; Thakur, A.; Zhu, T.; Clifton, D.A. Adversarial De-confounding in Individualised Treatment Effects Estimation. In Proceedings of the 26th International Conference on Artificial Intelligence and Statistics, PMLR, Valencia, Spain , 25–27 April 2023; Volume 206, pp. 837–849.
15. Chauhan, V.K.; Thakur, A.; O'Donoghue, O.; Rohanian, O.; Clifton, D.A. Continuous Patient State Attention Models. *medRxiv* **2022**. [CrossRef]
16. Salaun, A.; Knight, S.; Wingfield, L.R.; Zhu, T. Interpretable Machine Learning in Kidney Offering: Multiple Outcome Prediction for Accepted Offers. *medRxiv* **2023**. [CrossRef]
17. Lu, P.; Barazzetti, L.; Chandran, V.; Gavaghan, K.A.; Weber, S.; Gerber, N.; Reyes, M. Super-Resolution Classification Improves Facial Nerve Segmentation from CBCT Imaging. In Proceedings of the CURAC, Saskatoon, SK, Canada, 27 May 2016; pp. 143–144.
18. Lu, P.; Barazzetti, L.; Chandran, V.; Gavaghan, K.; Weber, S.; Gerber, N.; Reyes, M. Facial nerve image enhancement from CBCT using supervised learning technique. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 2964–2967.
19. Lu, P.; Bai, W.; Rueckert, D.; Noble, J.A. Multiscale graph convolutional networks for cardiac motion analysis. In Proceedings of the International Conference on Functional Imaging and Modeling of the Heart, Stanford, CA, USA, 21–25 June 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 264–272.

20. Ghiasi, S.; Zhu, T.; Lu, P.; Hagenah, J.; Khanh, P.N.Q.; Hao, N.V.; Vital Consortium, Thwaites, L.; Clifton, D.A. Sepsis Mortality Prediction Using Wearable Monitoring in Low–Middle Income Countries. *Sensors* **2022**, *22*, 3866. [CrossRef] [PubMed]

21. Tadesse, G.A.; Zhu, T.; Le Nguyen Thanh, N.; Hung, N.T.; Duong, H.T.H.; Khanh, T.H.; Van Quang, P.; Tran, D.D.; Yen, L.M.; Van Doorn, R.; et al. Severity detection tool for patients with infectious disease. *Healthc. Technol. Lett.* **2020**, *7*, 45–50. [PubMed]

22. Tadesse, G.A.; Javed, H.; Thanh, N.L.N.; Thi, H.D.H.; Thwaites, L.; Clifton, D.A.; Zhu, T. Multi-modal diagnosis of infectious diseases in the developing world. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2131–2141. [CrossRef]

23. Hatamizadeh, A.; Tang, Y.; Nath, V.; Yang, D.; Myronenko, A.; Landman, B.; Roth, H.R.; Xu, D. Unetr: Transformers for 3d medical image segmentation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2022; pp. 574–584.

24. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv* **2021**, arXiv:2102.04306.

25. Zhao, C.; Droste, R.; Drukker, L.; Papageorghiou, A.T.; Noble, J.A. Visual-Assisted Probe Movement Guidance for Obstetric Ultrasound Scanning Using Landmark Retrieval. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France, 27 September–1 October 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 670–679.

26. Zhang, J.; Li, C.; Liu, G.; Min, M.; Wang, C.; Li, J.; Wang, Y.; Yan, H.; Zuo, Z.; Huang, W.; et al. A CNN-transformer hybrid approach for decoding visual neural activity into text. *Comput. Methods Programs Biomed.* **2022**, *214*, 106586. [CrossRef]

27. Wu, H.; Chen, S.; Chen, G.; Wang, W.; Lei, B.; Wen, Z. FAT-Net: Feature adaptive transformers for automated skin lesion segmentation. *Med. Image Anal.* **2022**, *76*, 102327.

28. Lu, P.; Creagh, A.P.; Lu, H.Y.; Hai, H.B.; Consortium, V.; Thwaites, L.; Clifton, D.A. 2D-WinSpatt-Net: A Dual Spatial Self-Attention Vision Transformer Boosts Classification of Tetanus Severity for Patients Wearing ECG Sensors in Low-and Middle-Income Countries. *Sensors* **2023**, *23*, 7705. [CrossRef] [PubMed]

29. Byeon, Y.H.; Kwak, K.C. Pre-configured deep convolutional neural networks with various time-frequency representations for biometrics from ECG signals. *Appl. Sci.* **2019**, *9*, 4810.

30. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]

31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [CrossRef]

32. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.

33. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [CrossRef]