Title: A Convolved Self-Attention Model for IMU-based Gait Detection and Human Activity Recognition

Authors: Shuailin Tao, Wang Ling Goh, Yuan Gao

Summary:

This paper introduces a convolved self-attention neural network model for gait detection and human activity recognition (HAR) tasks using wearable inertial measurement unit (IMU) sensors. The proposed model aims to improve accuracy by incorporating prior time step knowledge through a convolved window inside the self-attention module. It also introduces a streamlined fully connected (FC) layer without hidden layers for the feature mixer, resulting in a significant reduction of network parameters. The method achieves better accuracy compared to other state-of-the-art neural networks on HAR datasets UCI-HAR and MHEALTH.

Methods:

1. Convolved Self-Attention: The paper proposes a convolved self-attention model, which combines the benefits of convolutional neural networks (CNN) and self-attention mechanisms. By using a convolved window, the model captures prior time step information and utilizes it for improved accuracy in gait detection and HAR tasks.

1. Streamlined Fully Connected (FC) Layer: The paper introduces a streamlined FC layer for the feature mixer, which does not include hidden layers. This design choice reduces the overall network parameters, as hidden layers typically occupy the majority of parameters in a transformer encoder.

1. IMU Data Segmentation and Convolution: The IMU data is processed by segmenting it into consecutive time steps and applying convolutional operations. This allows the model to capture local position information and relevant features for gait detection and HAR.

1. Self-Attention Model: The self-attention layer is employed in the proposed model to capture long-term dependencies in the input sequence. Unlike recurrent neural networks (RNN), self-attention can be applied to specific parts of the input sequence without the constraint of sequential processing.

1. Transformer Encoder: The paper mentions the limitations of transformer encoders, such as a large number of parameters and inferior performance compared to CNN when training datasets are small. The proposed convolved self-attention model addresses these limitations by incorporating prior knowledge and achieving better results.

Overall, the paper presents a novel approach that combines convolved self-attention with IMU-based gait detection and HAR. The proposed model demonstrates improved accuracy and reduced network parameters, making it suitable for implementation on edge devices.