Activity Recognition Using Imaging the Time Series

Mohammad Sakka

November 23, 2023

The proposed methodology for activity recognition leverages Gramian Angular Fields (GAF) and a Convolutional Neural Network (CNN) with a classification layer. The key steps are as follows:

1. **Transforming Time Series Data into Images (GAF):**
   - Time series data is transformed into images using the Gramian Angular Fields (GAF) method.
   - GAF encoding is a pivotal step, capturing temporal information through angular representations.
2. **Convolutional Neural Network (CNN):**
   - A CNN is employed to extract salient features from the GAF images.
   - The CNN leverages the rich information captured by different sensors in the time series data.
3. **Classification Layer:**
   - A classification layer refines the extracted features and fused sensor data.
   - The final layer enables accurate prediction and classification of the type of action being performed.

1 Introduction

Activity recognition is a fundamental aspect in the realm of human behavior analysis, finding applications in various domains such as healthcare, sports, and security. Traditional approaches often rely on signal processing techniques to extract features from raw sensor data. However, recent advancements in computer vision have paved the way for more sophisticated methodologies that can capture intricate patterns and dynamics inherent in time series data.

Our proposed methodology draws inspiration from a benchmark approach [1] that successfully utilizes contemporary computer vision techniques to address the challenges of activity recognition. This inspiration stems from the recognition of the inherent complexities in human activities and the need for robust methods to decipher these intricacies.

The importance of harnessing computer vision techniques in activity recognition cannot be overstated. Unlike traditional signal processing methods, computer vision allows for the extraction of high-level features from raw data, enabling the model to discern complex patterns and relationships. By transforming time series data into visual representations, such as Gramian Angular Fields (GAF), we leverage the power of image-based analysis, capturing temporal dependencies and nuances that might be challenging to extract through conventional means.

This shift towards computer vision-based methodologies not only enhances the accuracy of activity recognition but also broadens the applicability of such systems across diverse scenarios. The ability to comprehend and interpret human activities in a visual context facilitates more nuanced and context-aware predictions, contributing to the overall efficacy of the recognition system.

In the subsequent sections, we will delve into the details of our methodology, highlighting the key steps involved in transforming time series data into images, employing Convolutional Neural Networks (CNNs) for feature extraction, and introducing a fusion layer for the seamless integration of data from multiple sensors. This holistic approach aims to enhance the accuracy of activity recognition, marking a significant departure from conventional methods.

2 Methodology

Our approach draws inspiration from a benchmark methodology [1] that harnesses contemporary computer vision techniques to address the research objective. The methodology begins by transforming time series data into images using the Gramian Angular Fields (GAF) method, a pivotal step in encoding temporal information. Following this, a Convolutional Neural Network (CNN) is employed to extract salient features from these images, leveraging the rich information captured by different sensors.

To enhance the integration of data from multiple sensors, a fusion layer is introduced, facilitating the aggregation of sensor-derived information. This fusion layer plays a crucial role in consolidating diverse data sources, enabling the model to comprehend the holistic context of the activities.

The final step of the approach involves training a classification layer, which is pivotal for predicting the type of action being performed. This layer refines the extracted features and the fused sensor data, ultimately enabling accurate classification of activities.

3 Transforming Time Series Data into Images (GAF)

The transformation of time series data into images is a pivotal step in our methodology, and we employ the Gramian Angular Fields (GAF) method to achieve this. This process involves encoding temporal information into visual representations, providing a richer understanding of the underlying patterns within the time series data.

3.1 Gramian Angular Fields (GAF)

The GAF method is a powerful technique that converts time series data into images by capturing the angular relationships between data points. The process can be succinctly described as follows:

1. **Normalization:** Prior to GAF transformation, the time series data is normalized to ensure that the values fall within a standardized range. This step is essential for

mitigating the impact of varying scales and ensuring that the GAF representation is robust across different datasets.

2. **Gramian Matrix Calculation:** The normalized time series data is then used to construct a Gramian matrix. This matrix captures the pairwise relationships between the data points, emphasizing the angular patterns formed by their interactions. The Gramian matrix is computed using operations such as dot products or cosine similarities.

3. **Image Generation:** The Gramian matrix is transformed into an image, typically a heatmap, through mapping its values onto a color scale. This visual representation retains the temporal structure of the original time series, with darker and lighter regions indicating different angular relationships between data points.

3.2 Significance of GAF Transformation

The GAF transformation is pivotal for encoding temporal information in several ways:

- **Pattern Recognition:** GAF images highlight temporal patterns and dependencies that might be challenging to discern in raw time series data. The visual nature of the representation allows for the identification of recurring motifs and irregularities.

- **Feature Extraction:** GAF images serve as effective feature vectors, capturing complex temporal features in a condensed and interpretable form. These features are then utilized by subsequent stages of the methodology, such as the Convolutional Neural Network (CNN), for robust activity recognition.

- **Interpretability:** The visual nature of GAF images enhances the interpretability of the model's decision-making process. Researchers and practitioners can gain insights into the temporal dynamics of different activities by analyzing the generated images.

4 Transforming Time Series Data into Images (GAF)

the GAF transformation plays a crucial role in our methodology by converting raw time series data into visual representations that encode temporal information. This process enables effective pattern recognition, feature extraction, and interpretability, ultimately contributing to the accuracy and robustness of our activity recognition system.

5 Convolutional Neural Network (CNN)

In summary, a CNN processes GAF images derived from time series data, capturing temporal dependencies and patterns. It leverages convolutional and pooling layers to extract hierarchical features and can be adapted for multisensory integration when dealing with data from different sensors. The training process enables the model to learn discriminative features for subsequent classification or regression tasks.

6 Classification Layer

the classification layer refines the high-level features extracted by the preceding layers, incorporates multisensory information if applicable, and produces class probabilities for accurate activity classification. The layer's design, activation function, and loss function are pivotal in ensuring that the neural network effectively learns and generalizes from the training data to make accurate predictions on unseen data.

7 Github Link

https://github.com/MohammadSakka/CV-project

8 Experiments and Evaluation

We tried to apply the methodology with some simplification due to the lack of time, we converted the timeseries to images then we applied cnn to classify the activities. and instead of fusing many sensors data, we have done training with only one sensor which is the Ankle accelaration measurement sensor, we used the GOTOV Dataset [2]

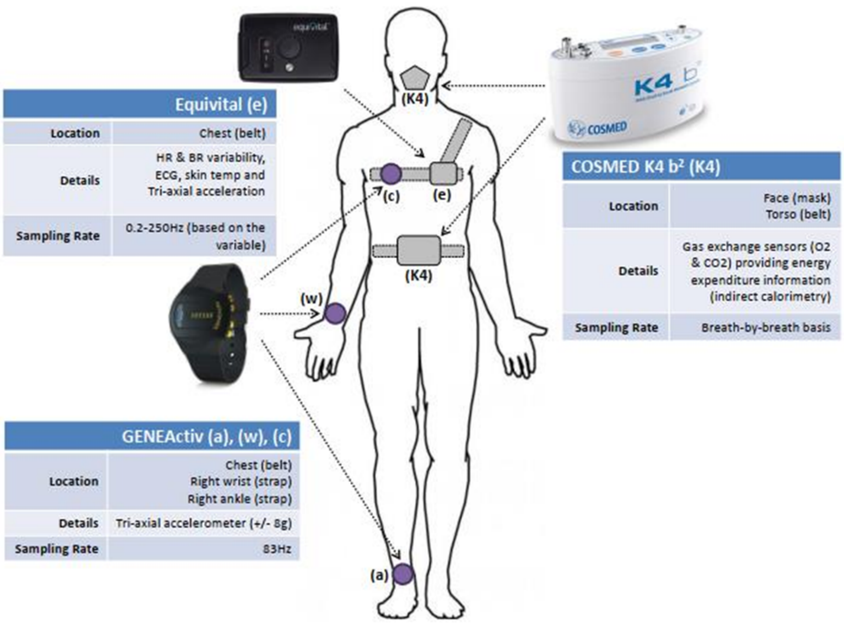

Figure 1: GOTOV Data

hence finally, we had a model suffers from low accuracy, we will try to fix and enhance in the demo and the presentation

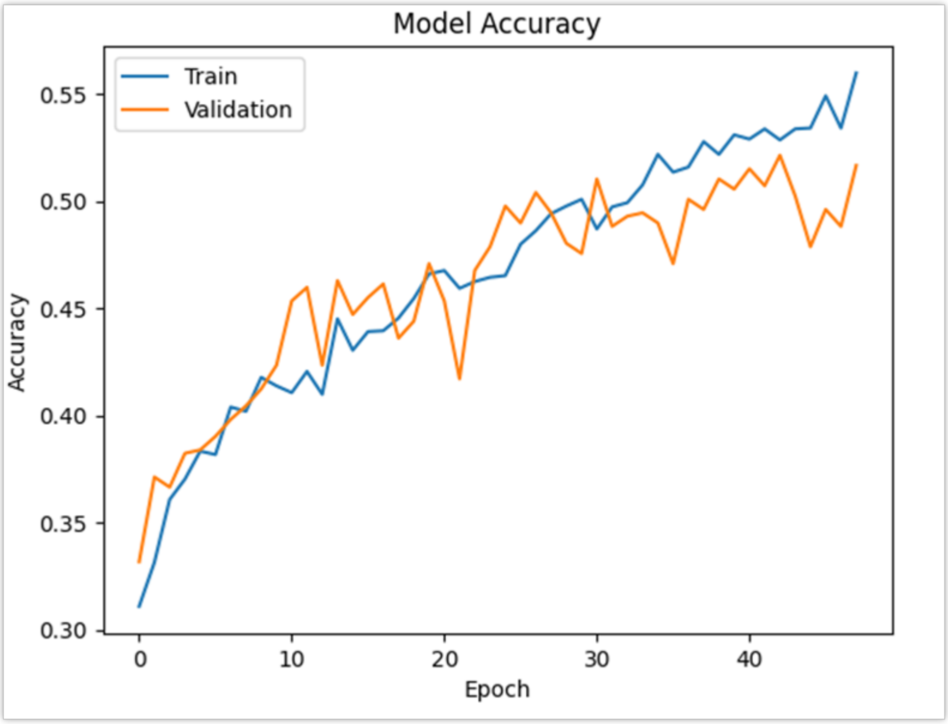

Figure 2: implementation

9 Conclusion

We tried to apply the methodology with some simplification due to the lack of time, we converted the timeseries to images then we applied cnn to classify the activities

10 References

- [1] Qin, Z., Zhang, Y., Meng, S., Qin, Z. and Choo, K.K.R., 2020. Imaging and fusing time series for wearable sensor-based human activity recognition. Information Fusion, 53, pp.80-87

- [2] Paraschiakos, S. et al. (2021) "GOTOV Human Physical Activity and Energy Expenditure Dataset on Older Individuals." 4TU.ResearchData. doi: 10.4121/12716081.V2.