Introduction

The research of human activity recognition (HAR) aim to identify human activity from the data acquired by sensors on human. In earlier studies, researchers used traditional machine learning methods to deal with this problem. Later, RNN became very popular as it is suitable for processing time-series data. But there are also problems with these approaches:

• Traditional machine learning methods’ ability to handle raw data and the performance to recognize activity was limited.
• In HAR, only a few parts of time-series data are relevant to human activities. This makes it difficult to capture important parts of human activity.

To solve these problems, we combined attention mechanism with the LSTM, a type of RNNs, so that it can focus on the more important parts. Every human activity consists of many atom motions, such as running and walking. Some of these atom motions are common to both running and walking, and thus not appropriate for distinguishing human activity. There is also motion that only happens when you are running, which is more suitable for classification. The degree to which an atom motion uniquely belong to one single activity is called uniqueness of this motion in this paper. Notice that our definition of atom motion is very broad. Every movement for a short period can be considered as atom motion.

Objectives: To design a model that can calculate the uniqueness attention of atom motion so that we can use this attention to focus on the parts of human activities that are more important for classification

Method

The model we design that can calculate attention is shown in Fig. 1. x_t denotes a sensor reading at time t. We use LSTM to encode these x_t, and for each x_t output h_t. The genatt module will calculate attention weight att_t for each h_t and combine these into H = \sum_t att_t \cdot h_t. The right elliptic region in Fig. 1 represents the internal structure diagram of its left genatt module.

This generic model don’t specify how to calculate attention weights. In order to compare with the attention model proposed below, we specify two way to calculate the attention weight, and thus form two benchmark models: LSTM with Last Attention, and LSTM with Mean Attention. LSTM with Last Attention only sets the attention weight, and thus form two benchmark models: LSTM with Last Attention, and LSTM with Mean Attention. LSTM with Last Attention only sets the attention weight to all of these vectors.

\[
\text{prob}_t = \text{softmax}(h_t)
\]
\[
\text{score}_t = \max_x (\text{prob}_t) - \max_x (\text{prob}_t)
\]
\[
\text{att}_t = \text{uniq} (\text{prob}_t) = \frac{\text{score}_t}{\text{prob}_t}
\]
\[
H = \sum_{t=1}^{T} \text{att}_t \cdot h_t
\]

In Fig. 2, each h_t’s attention weight is only determined by itself and softmax module. h_t is first processed by softmax module, and then turned into prob_t. If the current atom motion is unique to an activity, the difference between the largest and second largest value in vector of prob_t should be quite large. So uniq module use this difference to calculate the uniqueness attention of h_t. Then, we use the calculation defined in the last formula to get H, which will finally result in the activity recognition result through the softmax module.

\[
\text{prob}_t = \text{softmax}(h_t)
\]
\[
\text{score}_t = \max_x (\text{prob}_t) - \max_x (\text{prob}_t)
\]
\[
\text{att}_t = \text{uniq} (\text{prob}_t) = \frac{\text{score}_t}{\text{prob}_t}
\]
\[
H = \sum_{t=1}^{T} \text{att}_t \cdot h_t
\]

In Fig. 2, each h_t’s attention weight is only determined by itself and softmax module. h_t is first processed by softmax module, and then turned into prob_t. If the current atom motion is unique to an activity, the difference between the largest and second largest value in vector of prob_t should be quite large. So uniq module use this difference to calculate the uniqueness attention of h_t. Then, we use the calculation defined in the last formula to get H, which will finally result in the activity recognition result through the softmax module.

Contact Author: (shili@zju.edu.cn)

Other contact Information: WeChat: isteincn

\[
\text{prob}_t = \text{softmax}(h_t)
\]
\[
\text{score}_t = \max_x (\text{prob}_t) - \max_x (\text{prob}_t)
\]
\[
\text{att}_t = \text{uniq} (\text{prob}_t) = \frac{\text{score}_t}{\text{prob}_t}
\]
\[
H = \sum_{t=1}^{T} \text{att}_t \cdot h_t
\]

Result

Table 1. F_1 results of the PAMAP2 dataset

<table>
<thead>
<tr>
<th>Models</th>
<th>Locomotion</th>
<th>Gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM with Uniqueness Attention</td>
<td>0.8796</td>
<td>0.892</td>
</tr>
<tr>
<td>LSTM with Last Attention</td>
<td>0.7568</td>
<td>0.873</td>
</tr>
<tr>
<td>LSTM with Mean Attention</td>
<td>0.7922</td>
<td>0.883</td>
</tr>
<tr>
<td>Continuous + Temporal Attention</td>
<td>0.8629</td>
<td>0.892</td>
</tr>
</tbody>
</table>

Table 2. F_1 results of the Opportunity dataset

<table>
<thead>
<tr>
<th>Models</th>
<th>Locomotion</th>
<th>Gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM with Uniqueness Attention</td>
<td>0.892</td>
<td>0.904</td>
</tr>
<tr>
<td>LSTM with Mean Attention</td>
<td>0.875</td>
<td>0.899</td>
</tr>
<tr>
<td>LSTM with Last Attention</td>
<td>0.875</td>
<td>0.899</td>
</tr>
</tbody>
</table>

Conclusion

Based on our understanding on the relationship between atom motion and human activity, we propose a model: LSTM with Uniqueness Attention.

• This model outperforms other model on two dataset, which indicates its effectiveness.
• This model is better than these two LSTM benchmarks suggests that the uniqueness attention mechanism does enable our model to focus on more important parts of the input.
• By visualizing our uniqueness attention, we show the characteristics of this attention, and explain how similar are these characteristics to human intuition.

Other contact Information: WeChat: isteincn

ManuscriptID 342