GR-67

Representation Learning for Motion Sequence

Sponsors – Dr. Ying Xie & Dr. Linh Le
Advisor – Dr. Meng Han
Team Members – Saisangararamaleengam Alagapan, Alexandru Malos, Roshni Kishor, Venkateswara Reddy Mosali

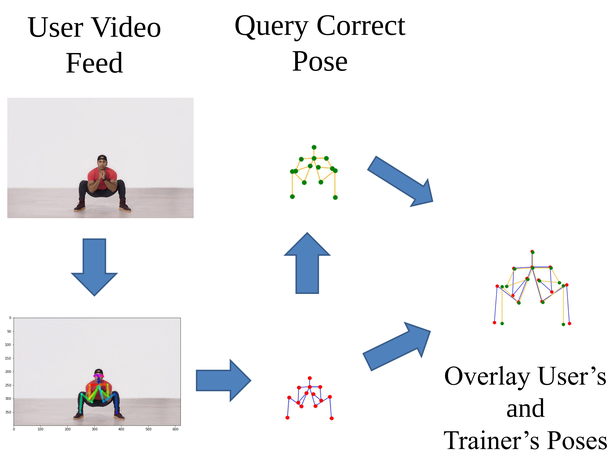
Abstract

The project proposes a new deep learning architecture that is used to align human poses to be used in exercise/rehabilitation assistant system. In short, the assistant system aims to provide users with visual feedback for their physical exercises. The feedback is generated by first extracting user's poses, overlaid with correct trainer's poses and then display for the user to observe and fix their errors.

Introduction

The project is regarding the first step in a smart workout assistant system. The COVID-19 pandemic has brought to the forefront the importance of promoting a healthy lifestyle. We wanted to look at preventative measures such as identifying and promoting ways for individuals to improve their health. This system addresses the health and wellness issues in the United States.

The Overall System

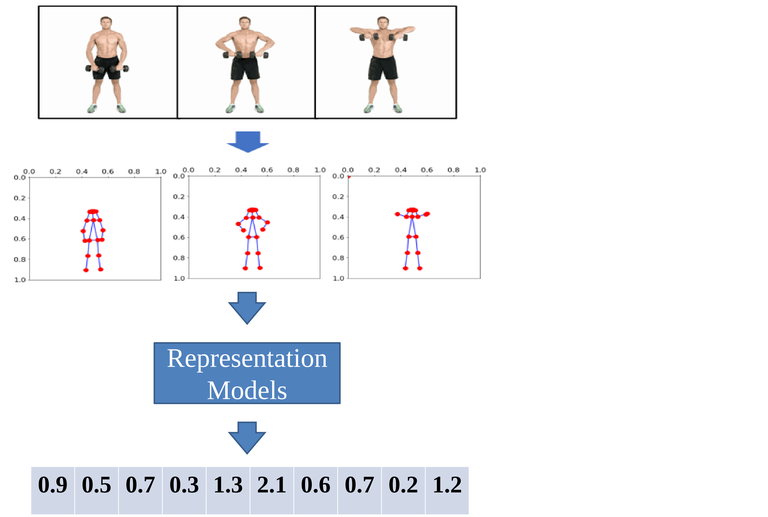


Overview

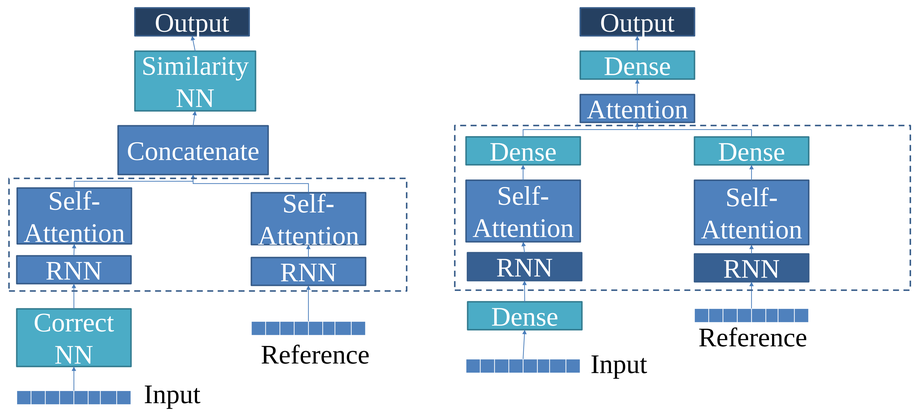
The system aims to assist people in physical and/or rehabilitation exercises by learning and providing visual feedback to the users, either during or after their training sessions. For a given video feed from user, our system works in three main steps such as:

- Capture/extract the user's pose
- Querying the correct pose
- Aligning the user's pose and queried pose

We transformed the pose sequences to a better representation for indexing and querying by running nine different representation models. Each pose is represented by an array of coordination of 14 main joints in the human body. A pose sequence is a sequence of arrays, and these models transforms each sequence of array into a single vector.



Model Flowcharts (Sqgru and Sqgru-att)



Results Comparison

	Loss Rate	N_epochs	Shuffle	Test
sqgru	0.0011	100	15	0.0148
sqgru-att	0.1494	100	15	0.1606

Conclusion

In this research project, we ran different representation models to achieve least loss rate, better accuracy and efficiency of the pose sequences. We believe that Sqgru model stands out from the rest of the other models by providing better accuracy and minimum loss rate.

Materials and Methods

The main objective was to choose the best representation model among the nine models based on the least loss rate. We collected a total count of 373 workout/exercise videos and converted them to pose sequences. We also utilized several deep learning models to learn the sequence embedding such as:

- Deep Auto-Encoders
- Transformers
- Siamese 3D Convolutional Neural Networks

References

Brownlee, J. (2020, August 27). A Gentle Introduction to LSTM Autoencoders. Machine Learning Mastery. https://machinelearningmastery.com/lstm-autoencoders/.

Salam, I. us. (2019, August 20). Autoencoders-Guide and Code in TensorFlow 2.0. Medium. https://medium.com/red-buffer/autoencoders-guide-and-code-in-tensorflow-2-0-a4101571ce56

KENNESAW STATE UNIVERSITY
COLLEGE OF COMPUTING AND SOFTWARE ENGINEERING