**Comment** : Specifically, in this work, the authors create a neural network using LSTM, which uses the input with only one time-step (with 12 features). Using one time-step as the input is not very meaningful for LSTM. Since the advantages of LSTM is for extracting information from sequential data (multiple features over several time steps).

***Response*:** The authors are thankful to the reviewer for this important suggestion and comment. Now, the experimentation is carried out by considering the multiple timesteps, and results are presented below for reference. In this experimentation, two hidden layers of LSTM are considered, with a dropout layer present between them with a total of 18343 as the trainable parameters. The model weights are updated for every 3500 samples (35-sec interval) and obtained RMSE for Roll, Pitch, Yaw, and their average, for few indicative values, are shown in the below table.

Table 1: RMSE Values (in radians) for different number of timesteps

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No of Timesteps | Roll | Pitch | Yaw | Average |
| 1 | 0.33 | 0.38 | 0.46 | 0.39 |
| 2 | 0.28 | 0.21 | 0.26 | 0.25 |
| 5 | 0.52 | 0.38 | 0.52 | 0.48 |
| 10 | 0.59 | 0.38 | 0.51 | 0.49 |

It can be observed that when 2-time steps are considered, the results are comparatively better. And hence now the manuscript is updated with the values obtained in new experimentation.

**Comment** : Specifically, using a neural network with a single LSTM layer and 6 hidden units seems not appropriate. If it is appropriate, then the problem by itself is not complicated enough to be useful. Furthermore, using LSTM to utilize sequential inputs but specifying only one time-step in the input, then it is not very meaningful. There are lots more to explore:

- Normalizing the input values

- Different LSTM architecture

- Different number of time steps to test

***Response*:** Authors are thankful to the reviewer for this important suggestion. The experimentation is carried out by trying different architectures are different combinations of number of hidden layers and hidden layers. Few indicative results are shown below for reference.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No of LSTM hidden units | Trainable  parameters | No of samples used in  incremental weight update | RMSE (in radian) | | | |
| Roll | Pitch | Yaw | Average |
| 20 | 2703 | 2000 | 0.38 | 0.12 | 0.29 | 0.26 |
| 20 | 2703 | 5000 | 0.60 | 0.25 | 0.56 | 0.47 |
| 50 | 12753 | 2000 | 0.44 | 0.20 | 0.36 | 0.34 |
| 50 | 12753 | 5000 | 0.37 | 0.25 | 0.28 | 0.30 |
| 200 | 1,71,003 | 2000 | 0.28 | 0.12 | 0.23 | 0.21 |
| 200 | 1,71,003 | 5000 | 0.35 | 0.23 | 0.26 | 0.28 |
| 300 | 3,76,503 | 2000 | 0.30 | 0.15 | 0.24 | 0.23 |
| 300 | 3,76,503 | 5000 | 0.39 | 0.22 | 0.23 | 0.28 |
| 1000 | 40,55,003 | 2000 | 0.29 | 0.13 | 0.24 | 0.22 |
| 1000 | 40,55,003 | 5000 | 0.35 | 0.35 | 0.25 | 0.32 |

It can be observed from the table that as the number of LSTM units is increasing, the improvement in the results is observable. However, the member of trainable parameters is drastically increasing, making the model complex, and more time is required for training and update. The table also indicates the results obtained when different intervals (2000 and 5000 samples) of weight update are considered. Even though the results with an interval of 2000 samples are better, the model weights are updated very frequently, making the process complex and time consuming compared to that when the interval of 5000 samples are used for incremental learning.

Similar experimentation is carried out with two hidden layers of LSTM cells.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No of LSTM hidden units | | Trainable parameters | No of samples using in  incremental weight update | RMSE (in radian) | | | |
| Layer1 | Layer2 | Roll | Pitch | Yaw | Average |
| 20 | 10 | 3,913 | 2000 | 0.43 | 0.28 | 0.49 | 0.40 |
| 20 | 10 | 3,913 | 5000 | 0.58 | 0.24 | 0.36 | 0.39 |
| 50 | 20 | 18,343 | 5000 | 0.26 | 0.22 | 0.27 | 0.25 |
| 50 | 20 | 18,343 | 2000 | 0.55 | 0.39 | 0.59 | 0.51 |
| 200 | 100 | 2,91,103 | 5000 | 0.52 | 0.33 | 0.39 | 0.41 |
| 200 | 100 | 2,91,103 | 2000 | 0.31 | 0.18 | 0.24 | 0.25 |

After all these experimentations, it was observed that the model with two LSTM hidden layers with 50 and 20 units respectively and a dropout layer between them perform relatively better. A time interval of 30 seconds is considered for model weights update. Further, this model is now considered in the manuscript and for results comparison with existing algorithms.

**Comment** : For online/incremental learning:

- For how many time steps you will continue to learn or learn and update for every single new time step?

- How would you validate that one or some new steps are not outlying and affect your performance?

***Response*:** Interval of every 30 seconds (3000 samples) is considered for model weight update in incremental learning mechanism. (line #274-275)

Validating the results with outliers and accordingly correcting the model is out of scope of the current research work. However, in future authors would like to carry out the work by using the rotary tables where hardware validation can be done and required actions can be taken.