Results

The devices and the recordings

During our recordings, We used two devices:

The first device Is a MRU which is used to get accurate positions during the recorded movement, we used a sampling rate of 40Hz and from the recordings we used only the following 3 columns*: “Height\_GNSS”*, *“Long\_GNSS”* and *“Lat\_GNSS”* from which we got the position of each step in a right handed coordination system, which we then transferred to the xyz system for our comfort.

From these recording we inferred the distance covered using the L2 norm as follows:

We use the formula on every 2 consecutive steps and we sum up the results of all the relevant steps of the movement to get the total distance covered.

The second device is a set of 4 IMU which we use to get 3 acceleration recordings from each IMU for the same movement. we used a sampling rate of 31.25Hz and from the recordings we got the acceleration in the x,y,z axis for each IMU. It is important to note that 2 of the IMU’s, numbers 2 and 4 work with a different reference system which is inverted to the “normal” one, in order to fix this we multiply the z and x axis of these IMU’s by -1. The recordings we got from the IMU’s is the input of out network.

Our recordings comprise of 10 different recordings, each one is about 30 seconds long and during it we used the IMU at the same time as the RMU. First, we activated the RMU and let it run uninterrupted for ~20 minutes for calibration purposes. After the calibration was complete, we started the recording by activating the IMU and RMU units and walking around the courtyard in random direction without stopping for the duration of the recordings. Because it is difficult to start and end both devices’ recordings at the same time we cut the first and last step of the recording.

Training

We use ResNet34 as our network for the project, the only difference we made to the network is changing the 2D conv layers to 1D conv layers because our data is time sensitive and therefore we only work on 1 dimension, forward in time.

For our training we used the RIDI dataset which is comprised of multiple recordings of movement done by multiple people in different situations (handheld, bag, leg etc). since the dataset was recorder at a sample rate of 200Hz we had to process it first so it will be useful to us. Since our sample rate was 31.25Hz and 200/31.25 = 6.4 we took every 6 measurements of the RIDI data and calculated its mean so we will get a dataset which was sampled at approximately the same rate as our recordings.

In order to evaluate our network we use 3 different stats. The first is a classic MSE loss, the second is the Accuracy, which tells us out of all the steps, how many of these we calculated the distance right (we define right as within a 10% error). The last stat is the relative error between the calculated position and the network’s output, it is calculated as follows:

We use the RIDI dataset for the training and the recordings of the 10 steps for the test. For each step we can test each individual IMU unit, the average of them and another method we devised which we refer to as “the voting method” in which for each axis in each step we choose the value which the majority of the IMU’s agrees on using a predefined method for “agreement”.

Test 1

We used a dataset called “tang\_handheld2”, after we processed it as mentioned before we got to training.

We used a window size of 50 samples while the batch size is 20 and for 50 epochs. We started with a learning rate of 7e-3 and we use a scheduler which multiplies the learning rate by a factor of 0.1 every 20 epochs.

The results for the first 3 recordings can be seen in the following table:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | IMU1 | IMU2 | IMU3 | IMU4 | Mean | Voting |
| 1 | MSE Loss | 0.8568 | 0.5720 | 0.6289 | 0.7147 | 0.6706 | 0.902 |
| Accuracy | 33.33% | 38.1% | 42.86% | 28.57% | 57.14% | 33.33% |
| Relative error | 13.0948% | 8.0493% | 9.0522% | 10.3219% | 10.7521% | 12.9237% |
| 2 | MSE Loss | 1.1431 | 0.9150 | 0.7989 | 1.1014 | 0.5767 | 0.5202 |
| Accuracy | 34.78% | 43.48% | 47.83% | 43.48% | 56.52% | 47.62% |
| Relative error | 13.6696% | 11.3789% | 10.3734% | 13.1977% | 9.1366% | 0.4042% |
| 3 | MSE Loss | 1.6638 | 1.9755 | 1.8783 | 2.1180 | 1.2873 | 5.6734 |
| Accuracy | 24% | 48% | 20% | 48% | 40% | 4.76% |
| Relative error | 10.1401% | 10.3140% | 11.9506% | 11.0702% | 1.7729% | 25.0191% |

1 – Without the fix.

2 – Without the fix.

3 - With the fix.

Test 2

We used a dataset called “hao\_handheld1”, after we processed it as mentioned before we got to training.

We used a window size of 100 samples while the batch size is 20 and for 50 epochs. We started with a learning rate of 7e-3 and we use a scheduler which multiplies the learning rate by a factor of 0.1 every 20 epochs.

The results for the first 3 recordings can be seen in the following table:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | IMU1 | IMU2 | IMU3 | IMU4 | Mean | Voting |
| 1 | MSE Loss | 0.9365 | 1.3260 | 1.0124 | 2.3100 | 1.3417 | 0.9519 |
| Accuracy | 50% | 30% | 30% | 10% | 20% | 60% |
| Relative error | 7.5422% | 12.7954% | 3.5241% | 18.3919% | 9.3721% | 5.1456% |
| 2 | MSE Loss | 3.7633 | 5.0991 | 3.8455 | 5.0378 | 5.2962 | 2.5852 |
| Accuracy | 27.27% | 18.18% | 36.36% | 18.18% | 27.27% | 30% |
| Relative error | 7.8003% | 9.1072% | 4.3779% | 11.9221% | 20.2802% | 5.7085% |
| 3 | MSE Loss | 14.2370 | 10.2985 | 12.0901 | 11.6376 | 15.1682 | 13.5381 |
| Accuracy | 16.67% | 25% | 16.67% | 16.67% | 25% | 0% |
| Relative error | 28.8402% | 22.009% | 25.2295% | 24.0436% | 28.6846% | 30.8017% |

1. With the fix.
2. With the fix.
3. With the fix.