**Question 06**

**Why you select the LSTM as the neural network? Please justify in the paper. If CNN can be used instead?**

In our research, we selected Long Short-Term Memory (LSTM) as the neural network for Wireless Capsule Endoscopy (WCE) localization because of its suitability for sequence data and its ability to capture temporal dependencies in the input data.

WCE localization involves predicting the 3D coordinates of the capsule as it moves through the gastrointestinal tract, which can be viewed as a sequential problem. LSTM, as a type of recurrent neural network (RNN), is well-suited for handling sequences and time series data. It can effectively model the dynamics and dependencies in the received path loss data as the capsule traverses the body.

While Convolutional Neural Networks (CNNs) are excellent for image-based tasks and have been used in various medical imaging applications, they are primarily designed for spatial data and may not be the optimal choice for sequential data like path loss measurements. CNNs are better suited for tasks like image classification and object detection, where spatial features play a crucial role.

Finally, we justify the choice of LSTM by highlighting its experimental results that demonstrate low root mean squared error (RMSE) achieved using LSTM in the context of WCE localization.

**References:**

1. Pham, T.D. Time–frequency time–space LSTM for robust classification of physiological signals. Sci Rep 11, 6936 (2021). https://doi.org/10.1038/s41598-021-86432-7

DOI

https://doi.org/10.1038/s41598-021-86432-7