Fig. 1. The network structure of three tested neural networks. Dash line indicates the operation of dropout.
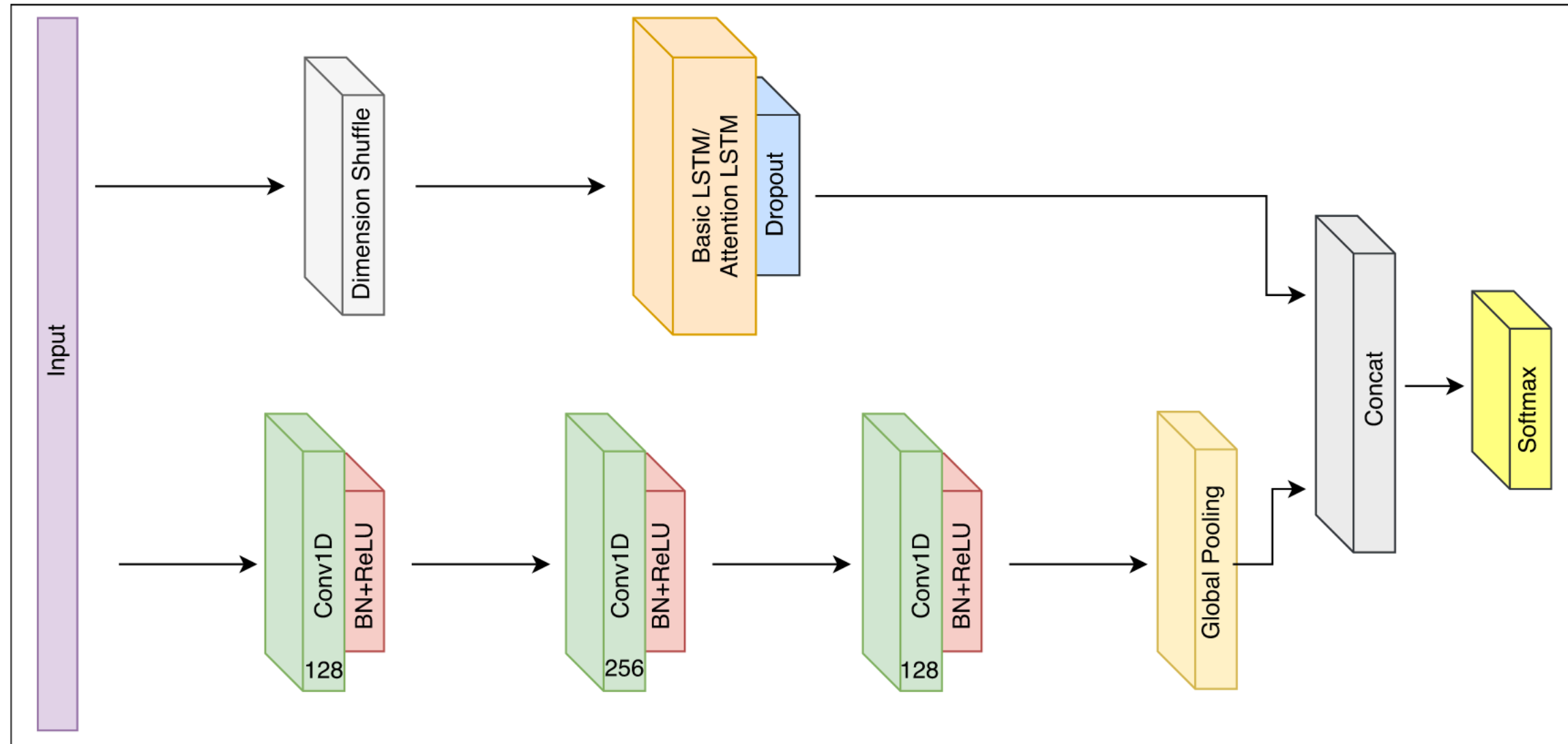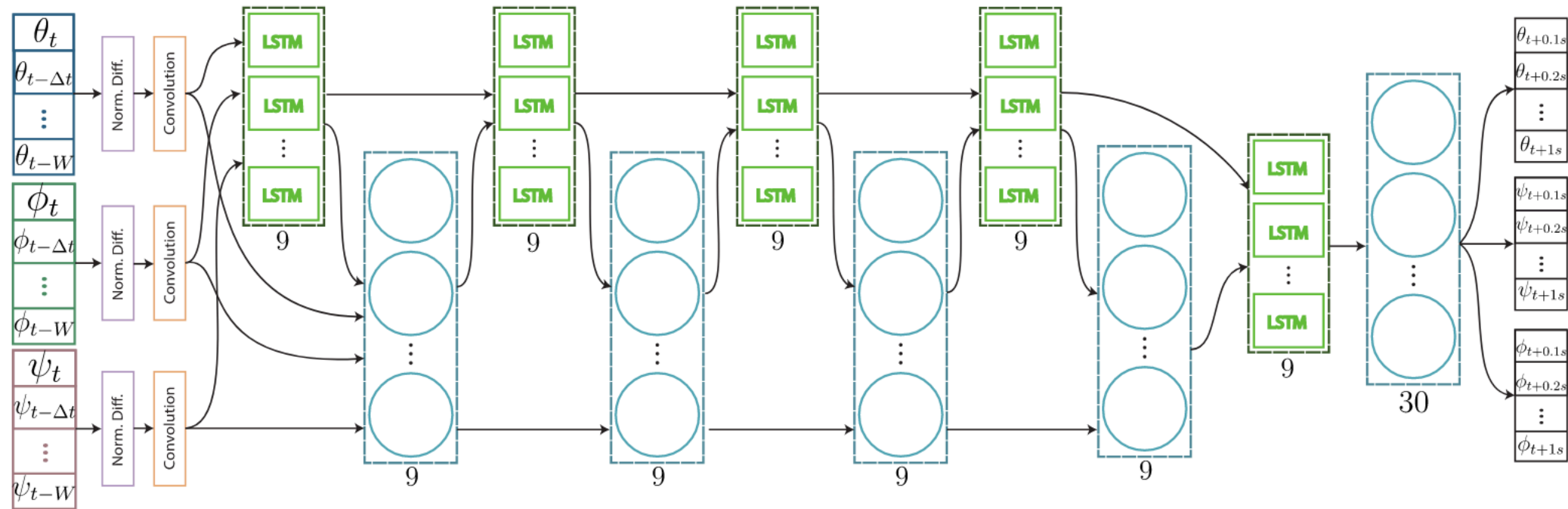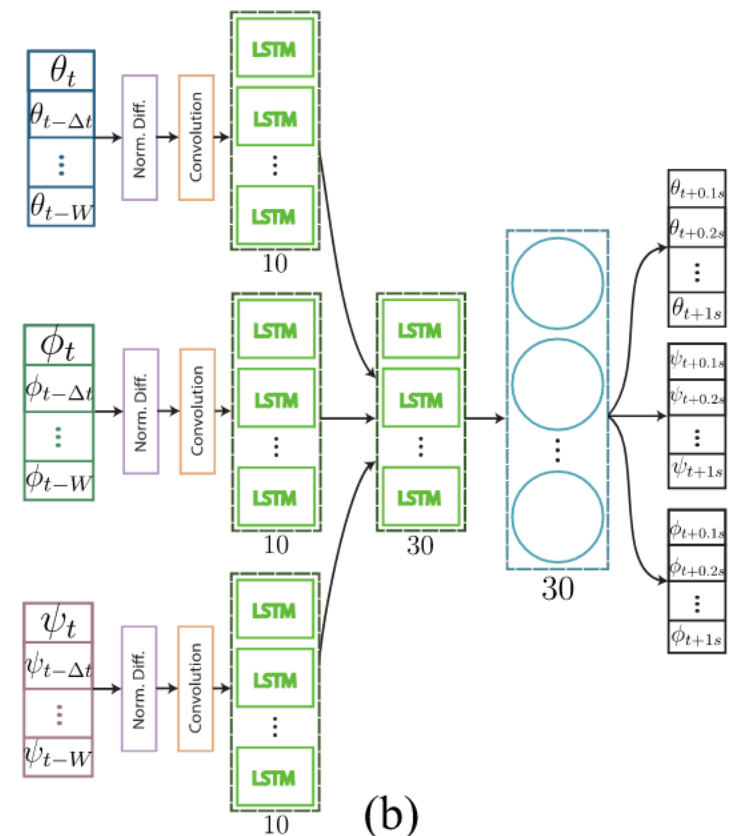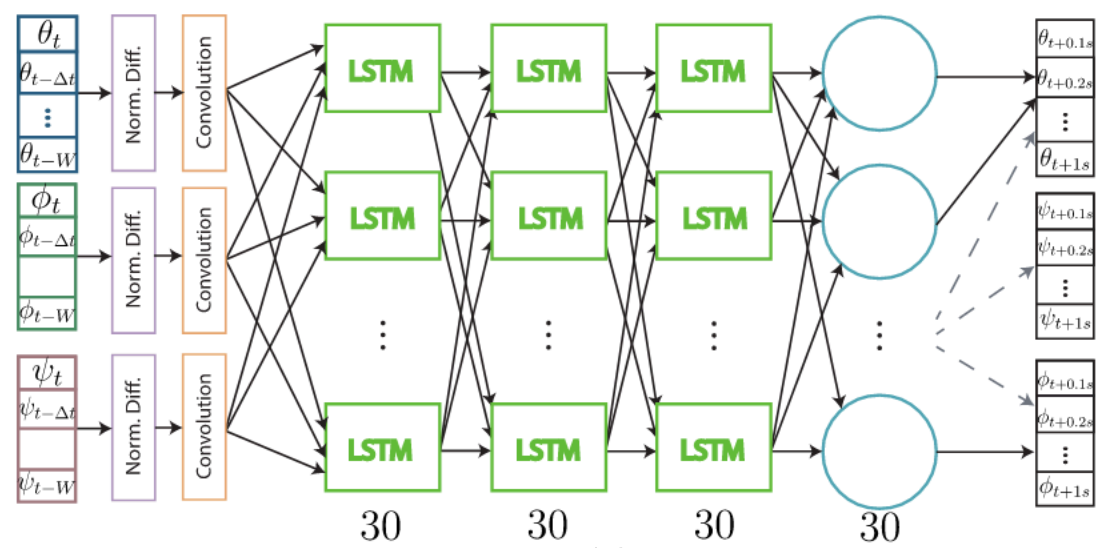
FIGURE 1. The LSTM-FCN architecture. LSTM cells can be replaced by Attention LSTM cells to construct the ALSTM-FCN architecture.
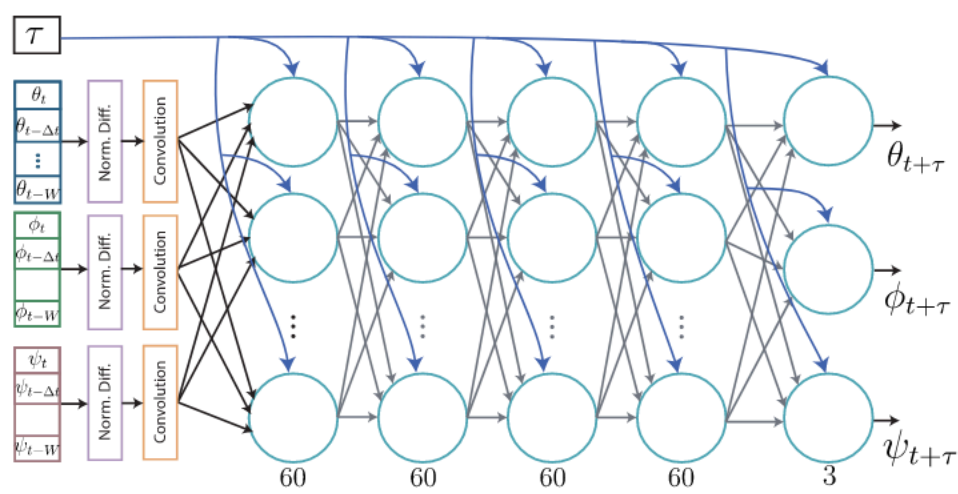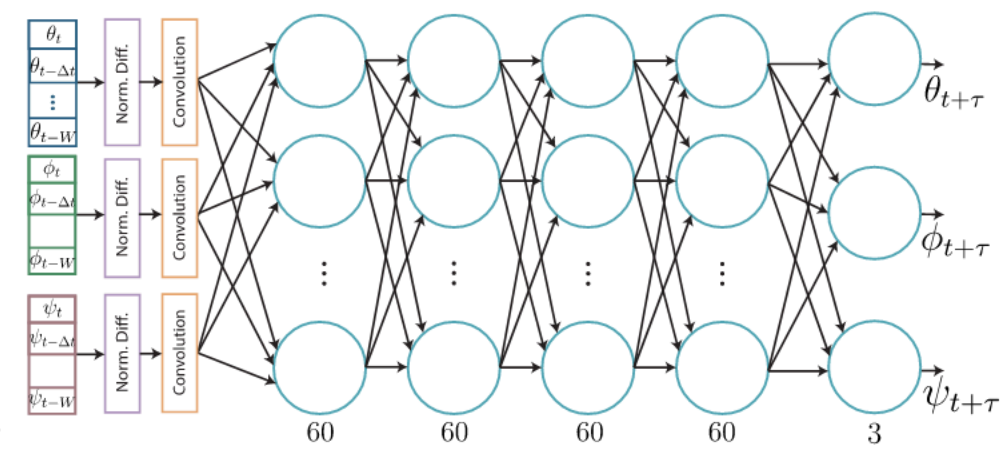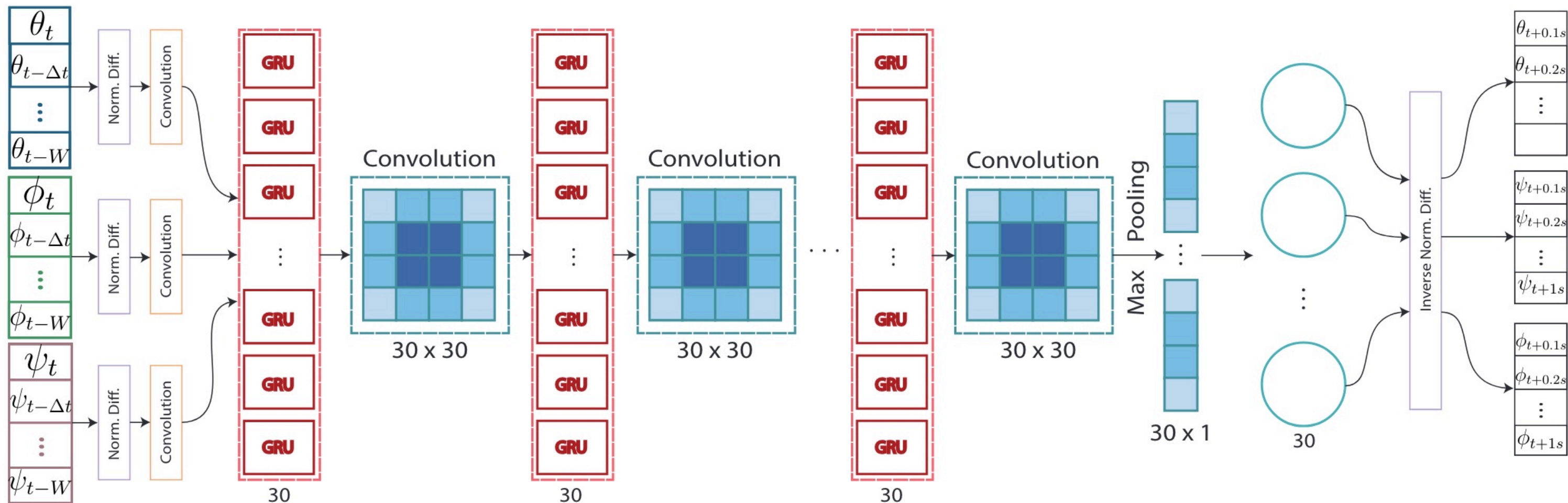
(a)

(b)

(c)

(d)

(e)

Fig. 9: Schematic overview of our proposed deep neural network for head-motion prediction. The input is first transformed to normalized difference to ensure generalization, and passed through a convolution unit that acts as a low-pass filter. Six sequential GRU and convolution units compute the most distinct features at different granularities. A subsequent max pooling layer downsamples the output of the last convolution layer und preserves the most significant features. After a dense feed-forward network, the differences are inverted to absolute values. The output predicts a whole sequence of future orientation values ($\bar{\xi}^{\text{pred}}_{t+n\cdot 100\text{ms}}$ $\forall n \in \mathbb{Z}, n = 1, ..., 10$).