



Figure 5. Neural network architecture. This is the architecture of the conditional transition density network described in section 4.2. We use Bidirectional Convolution LSTM and Convolution 3D Neural Networks as building blocks with notation '40@3x3' meaning 40 feature maps with kernel size 3 by 3. The input size of the networks is $7 \times patch_size \times patch_size$: 5 patches for $Y_{t-2:t+2}$, plus binary masks M_{t-1} and M_t encoding the positions of the particles at times $t-1$ and t respectively. The output probability map is $patch_size \times patch_size$. We used $patch_size = 28$ here. The new birth network in section 4.3 uses a similar architecture.