

# Supplementary Materials: Hydrodynamics-Informed Neural Network for Simulating Dense Crowd Motion Patterns

Anonymous Authors

## 1 OVERVIEW

This document contains supplementary material that complements our primary submission.

- We provide more experimental details related to training, testing and evaluation phases.
- We compare different parameter settings to explain why we select the specific group of parameters for our experiments.
- We exhibit some video clips from our dataset.
- We present additional experimental results from both the ablation studies and comparative experiments.

## 2 EXPERIMENTAL DETAILS

In total, 79 video clips are used for training and 47 video clips are allocated for testing. During the training phase, all the simulation models are trained for 1000 epochs. The learning rate starts from 0.001 and decays by 0.1 at epoch 500. Following training and testing, the simulations corresponding to the 47 video clips compose a prediction set, which is used to evaluate the performance of each simulation model. The prediction sets are evaluated by objective metrics including Inception Score (IS), Fréchet Inception Distance (FID), and Structural Similarity (SSIM). IS and FID are calculated based on a classifier called InceptionV3, which is trained on 345 video clips and tested on 112 video clips.

## 3 PARAMETER SETTINGS

In our primary submission, our governing equation for dense crowd simulation is given as:

$$f_{acc_x} = \lambda_1 f_{con_x} + \lambda_2 f_{vis_x} + \lambda_3 f_{ali_x} + \lambda_4 f_{nav_x} + \lambda_5 f_{coh_x} \quad (1)$$

$$f_{acc_y} = \lambda_1 f_{con_y} + \lambda_2 f_{vis_y} + \lambda_3 f_{ali_y} + \lambda_4 f_{nav_y} + \lambda_5 f_{coh_y} \quad (2)$$

where  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  and  $\lambda_5$  are parameters. For experiments, we select the parameters as 1, -0.2, 0.5, 1 and 1 respectively. To explain why we choose this group of parameters, we conduct some extra experiments using different parameter settings and evaluated them by objective metrics. Here, we mainly focus on  $\lambda_2$  and  $\lambda_3$ , the other parameters are still set to be 1.

**Table 1: Comparative performance of different parameter settings evaluated by Inception Score (IS), Fréchet Inception Distance (FID), and Structural Similarity (SSIM). Note that we conduct training for only 200 epochs here.**

Parameters	IS $\uparrow$	FID $\downarrow$	SSIM (%) $\downarrow$
$\lambda_2 = -1, \lambda_3 = 1$	<b>1.328</b>	0.148	<b>37.74</b>
$\lambda_2 = -1, \lambda_3 = 0.5$	1.042	<b>0.074</b>	38.36
$\lambda_2 = -0.2, \lambda_3 = 1$	1.274	0.108	38.35
$\lambda_2 = -0.1, \lambda_3 = 1$	1.293	0.076	38.39
$\lambda_2 = -0.2, \lambda_3 = 0.5$ (ours)	1.323	0.083	37.81

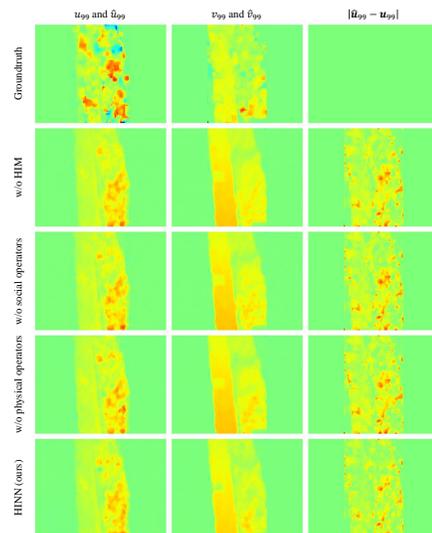
As shown in Tab.1, although parameter setting  $\lambda_2 = -0.2, \lambda_3 = 0.5$  may not yield the best performance in terms of IS, FID and SSIM individually, it demonstrates strong overall performance. Parameter settings  $\lambda_2 = -1, \lambda_3 = 1$  and  $\lambda_2 = -1, \lambda_3 = 0.5$  excel on certain metrics but perform poorly on others. Parameter settings  $\lambda_2 = -0.2, \lambda_3 = 1$  and  $\lambda_2 = -0.1, \lambda_3 = 1$  performs slightly worse than the first mentioned setting. Thus, we choose  $\lambda_1 = 1, \lambda_2 = -0.2, \lambda_3 = 0.5, \lambda_4 = 1$  and  $\lambda_5 = 1$  for experiments in primary submission.

## 4 VIDEO CLIPS

We select video clips that contain dense crowd motions under six classic motion patterns from our Dense Crowd Flow Dataset (DCFD). These video clips are named according to their motion patterns and stored in the archive.

## 5 EXPERIMENTAL RESULTS

To conduct a more compelling comparison between our HINN and ablated models or other simulation models, we present additional comparative figures that are not included in primary submission due to page limitations. Fig.1 to Fig.5 are comparisons for ablation studies and Fig.6 to Fig.11 are for comparative experiments. For typographical reasons, we opt not to display the "curve" pattern in ablation studies again, which has already been shown in primary submission.



**Figure 1: Comparison between groundtruth and simulations of crowd motion under a "line" pattern at time-step 99. The first and second column displays the horizontal and vertical components of the groundtruth and simulations respectively. The third column presents error maps of generated velocity for each model compared to the groundtruth.**

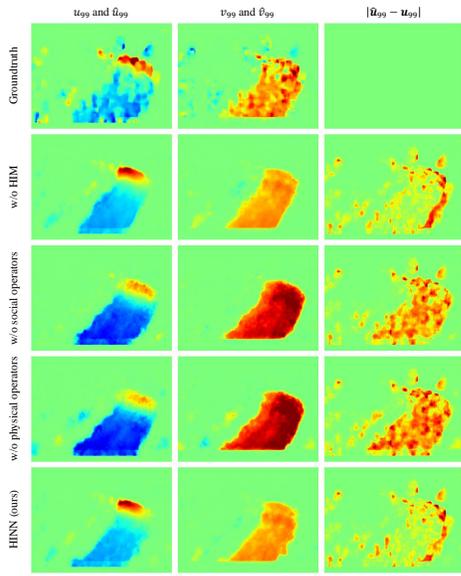


Figure 2: Comparison between groundtruth and simulations of crowd motion under a "curve" pattern at time-step 99. The first and second column displays the horizontal and vertical components of the groundtruth and simulations respectively. The third column presents error maps of generated velocity for each model compared to the groundtruth.

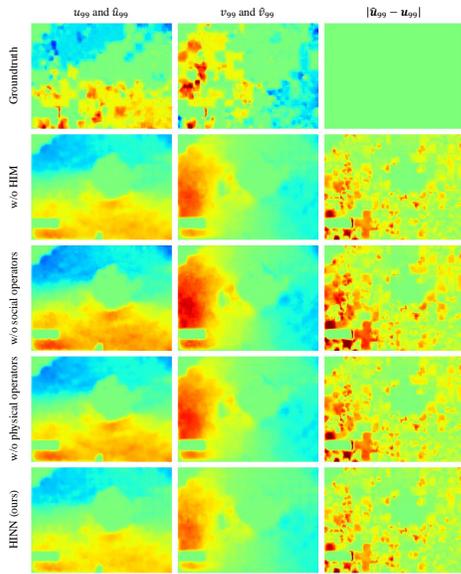


Figure 3: Comparison between groundtruth and simulations of crowd motion under a "circle" pattern at time-step 99. The first and second column displays the horizontal and vertical components of the groundtruth and simulations respectively. The third column presents error maps of generated velocity for each model compared to the groundtruth.

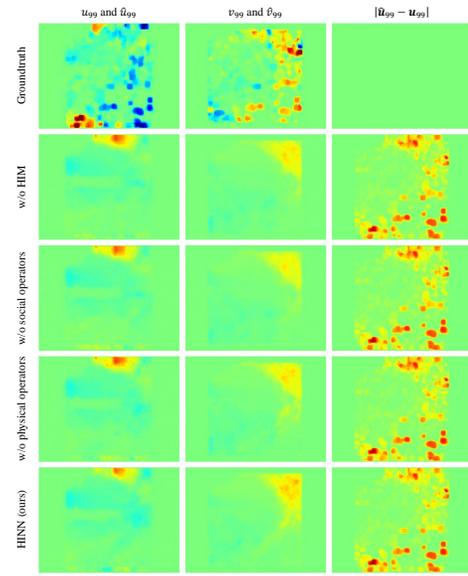


Figure 4: Comparison between groundtruth and simulations of crowd motion under a "cluster" pattern at time-step 99. The first and second column displays the horizontal and vertical components of the groundtruth and simulations respectively. The third column presents error maps of generated velocity for each model compared to the groundtruth.

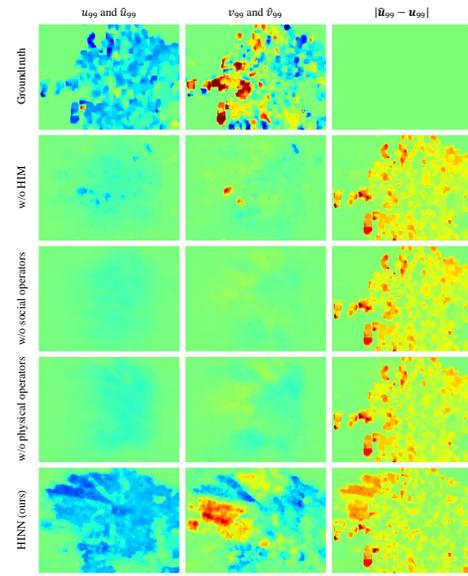


Figure 5: Comparison between groundtruth and simulations of crowd motion under a "scatter" pattern at time-step 99. The first and second column displays the horizontal and vertical components of the groundtruth and simulations respectively. The third column presents error maps of generated velocity for each model compared to the groundtruth.

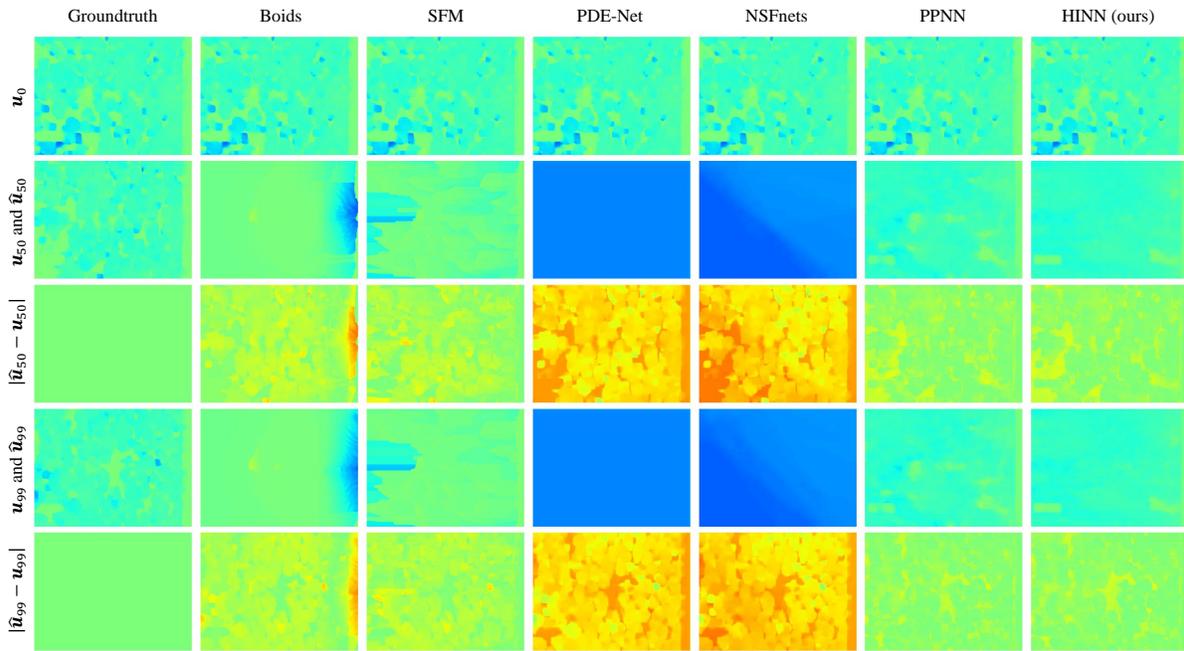


Figure 6: Comparison between groundtruth and simulations of the horizontal component of crowd motion under a "line" pattern at different time-steps. The first row depicts the given initial state. The second and fourth rows present the groundtruth and corresponding simulations at time-step 50 and time-step 99 respectively. The third and fifth rows display error maps for each model compared to the groundtruth at time-steps 50 and 99.

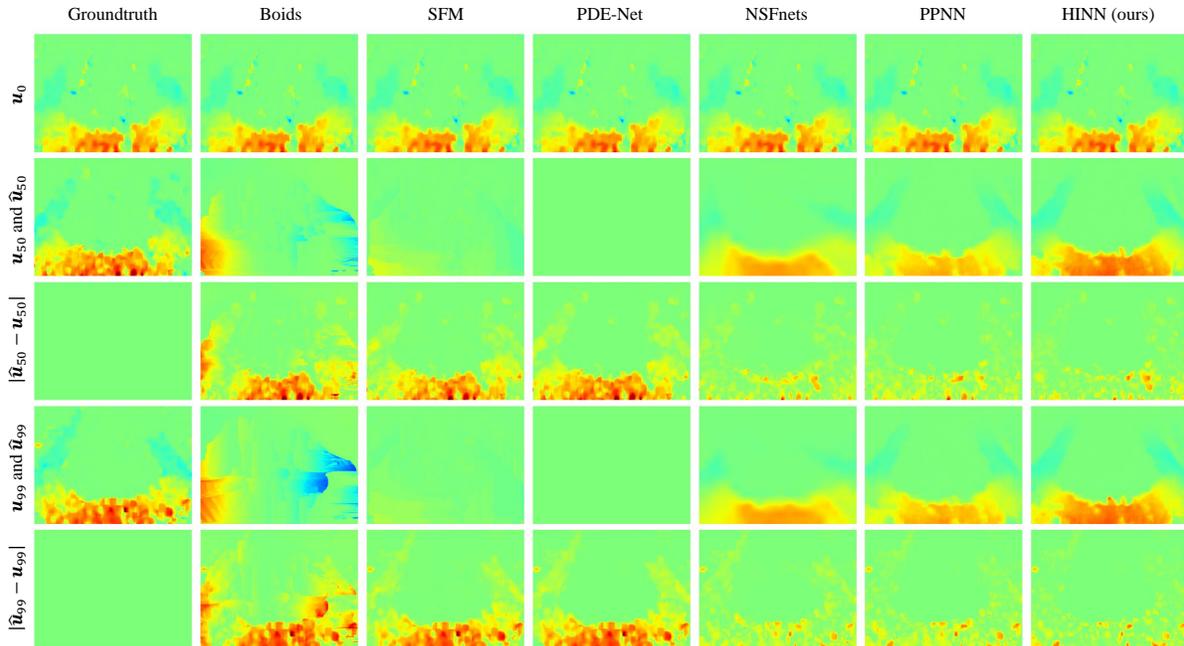


Figure 7: Comparison between groundtruth and simulations of the horizontal component of crowd motion under a "curve" pattern at different time-steps. The first row depicts the given initial state. The second and fourth rows present the groundtruth and corresponding simulations at time-step 50 and time-step 99 respectively. The third and fifth rows display error maps for each model compared to the groundtruth at time-steps 50 and 99.

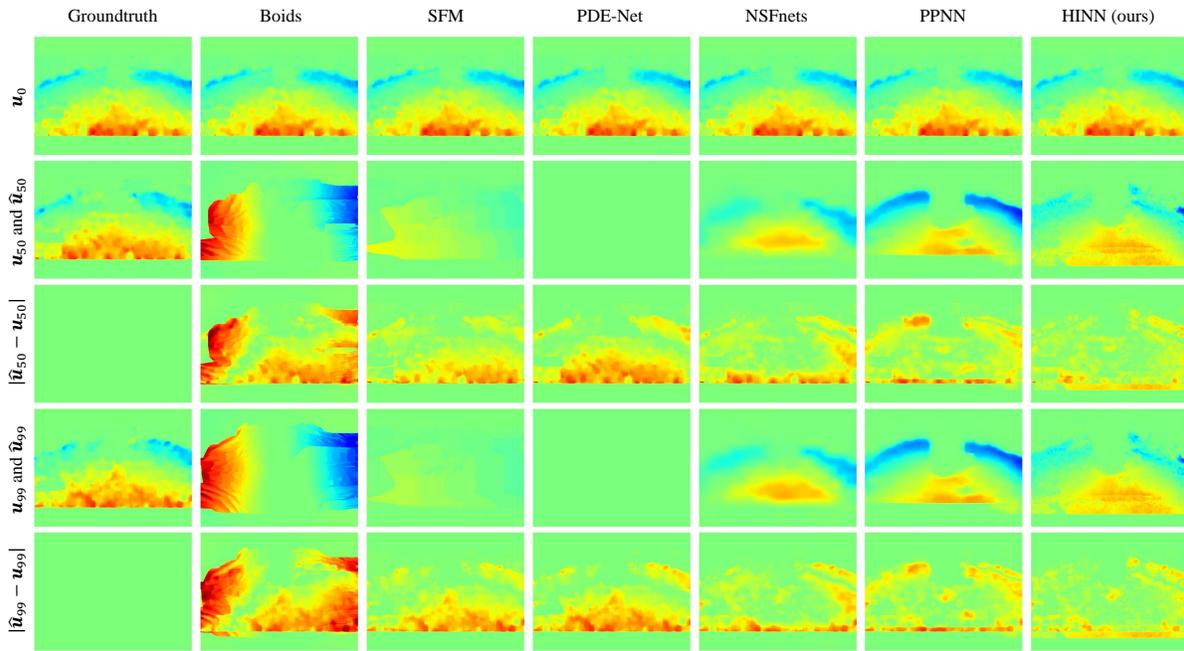


Figure 8: Comparison between groundtruth and simulations of the horizontal component of crowd motion under a "circle" pattern at different time-steps. The first row depicts the given initial state. The second and fourth rows present the groundtruth and corresponding simulations at time-step 50 and time-step 99 respectively. The third and fifth rows display error maps for each model compared to the groundtruth at time-steps 50 and 99.

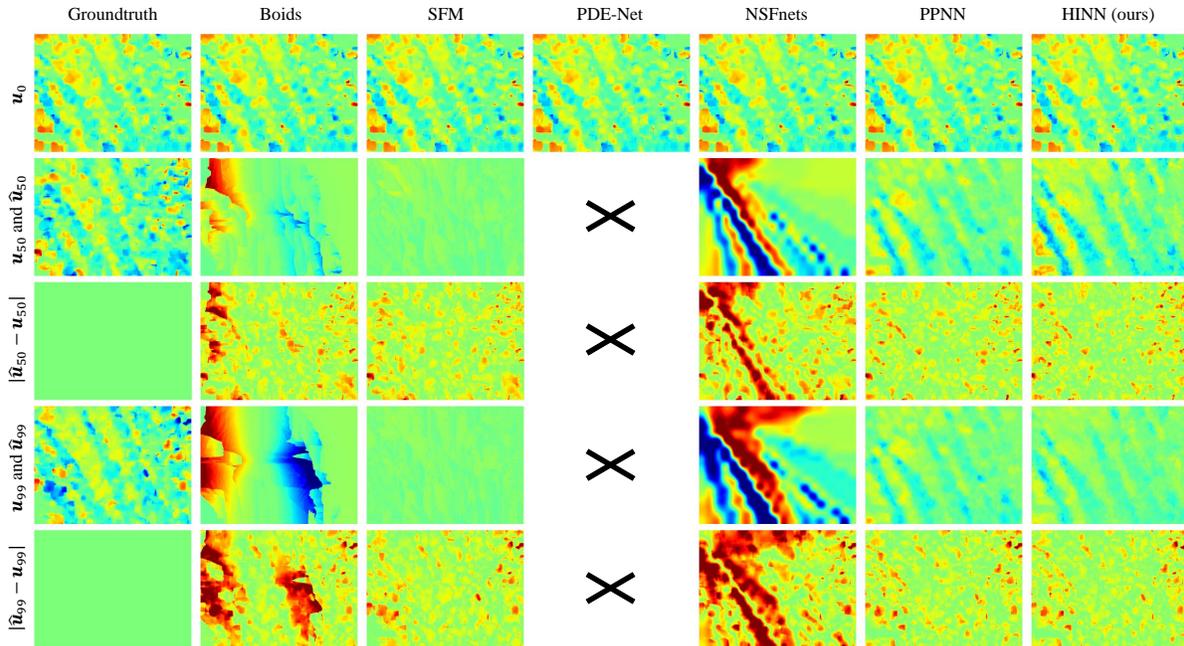


Figure 9: Comparison between groundtruth and simulations of the horizontal component of crowd motion under a "cross" pattern at different time-steps. The first row depicts the given initial state. The second and fourth rows present the groundtruth and corresponding simulations at time-step 50 and time-step 99 respectively. The third and fifth rows display error maps for each model compared to the groundtruth at time-steps 50 and 99.

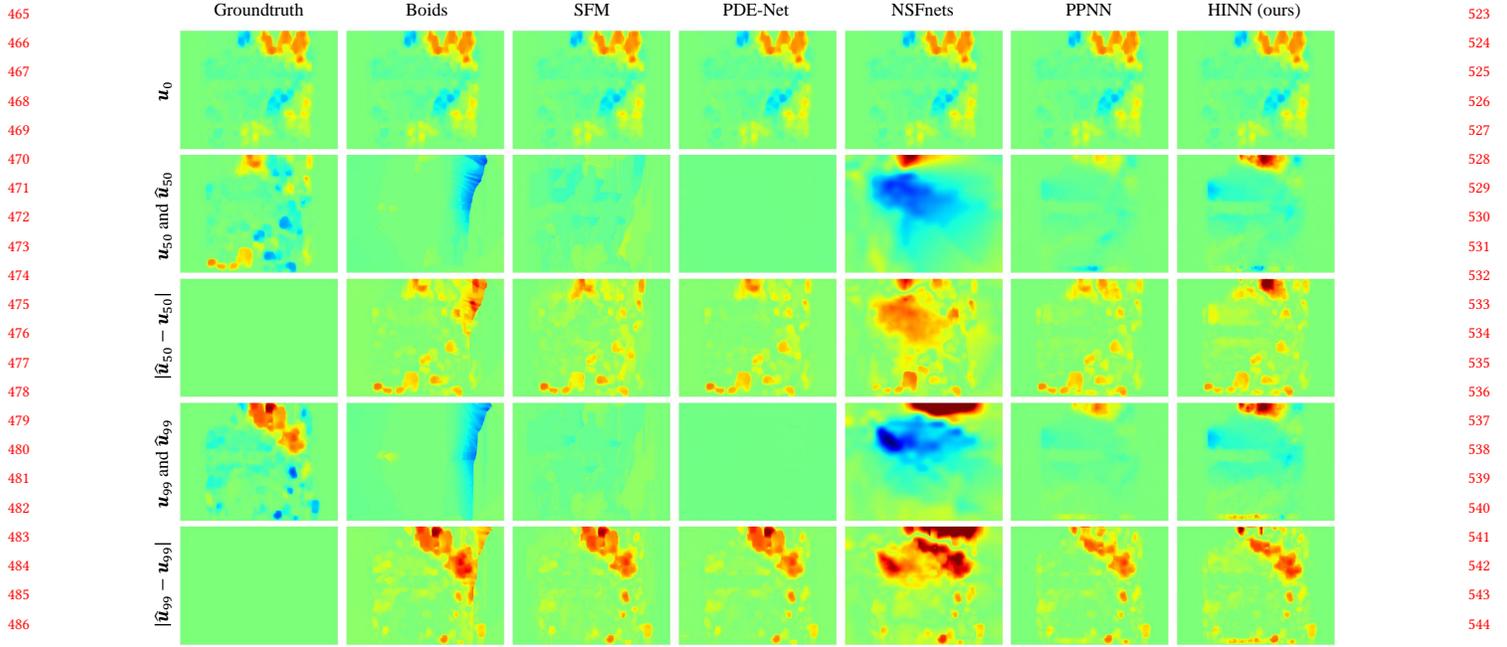


Figure 10: Comparison between groundtruth and simulations of the horizontal component of crowd motion under a "cluster" pattern at different time-steps. The first row depicts the given initial state. The second and fourth rows present the groundtruth and corresponding simulations at time-step 50 and time-step 99 respectively. The third and fifth rows display error maps for each model compared to the groundtruth at time-steps 50 and 99.

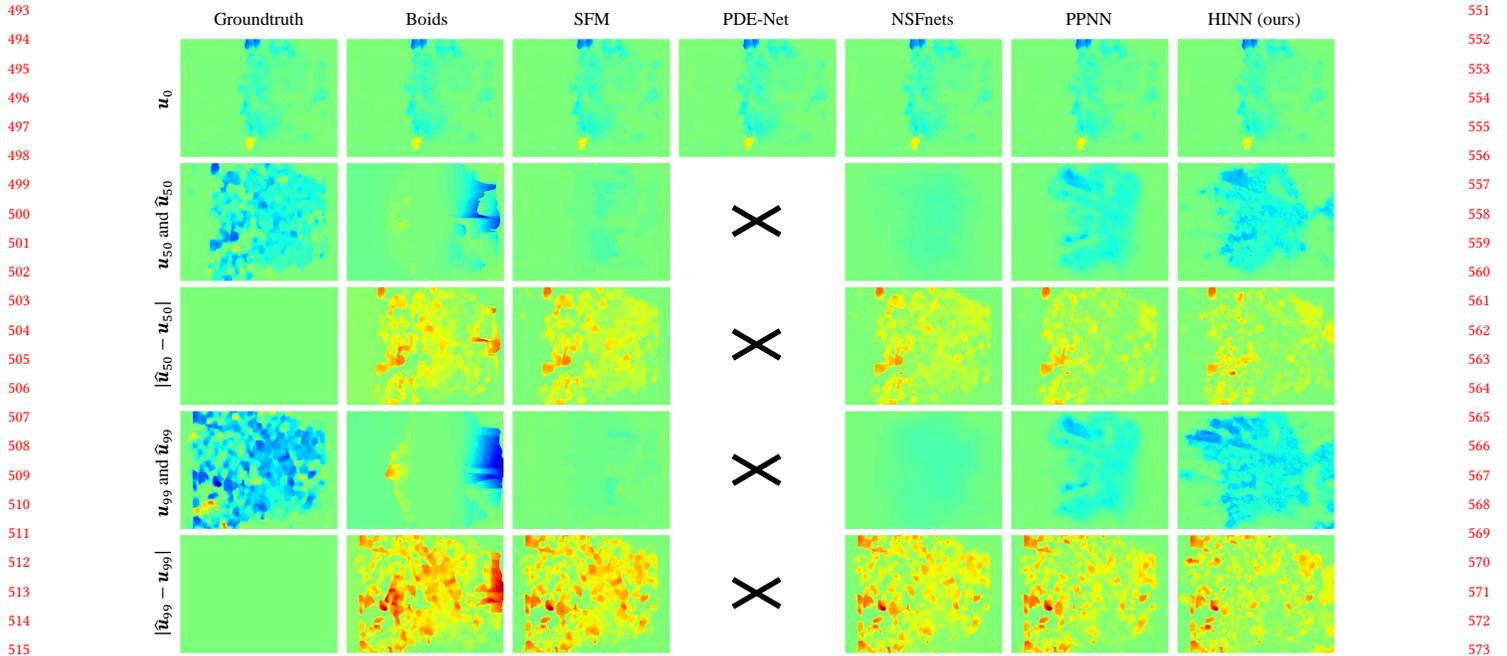


Figure 11: Comparison between groundtruth and simulations of the horizontal component of crowd motion under a "scatter" pattern at different time-steps. The first row depicts the given initial state. The second and fourth rows present the groundtruth and corresponding simulations at time-step 50 and time-step 99 respectively. The third and fifth rows display error maps for each model compared to the groundtruth at time-steps 50 and 99.