Multi Domain Learning for Motion Magnification
Supplementary Material

1 Additional Experiments

Frequency selectivity When the deep learning method is not trained directly with temporal filters, using temporal filters on intermediate features can produce incorrect results [1]. So to avoid that, video is first pre-processed with temporal filter to suppress unwanted motion. For this, [2] method's output at a small magnification factor (magnification factor=4) is given as an input to our method. Visual results are shown in the supplementary video.

Temporal Interpolation Motion magnification is the task of frame extrapolation for a magnification factor greater than 1. For this, the proposed method extrapolate phase change in the direction of change in previous phases. In the case of frame interpolation, the proposed method needs to interpolate the phase between two frames. Slow motion videos are generated with magnification factor between 0 to 1 (note:- for generation of these results, the model has not fine-tuned). By changing magnification factor (between 0 to 1) any number of frames can be interpolated, but as the interpolation increases motion smoothness reduces. Since the output is generated from the same weights trained for the motion magnification task, magnification between 0 to 1 is not properly defined. Visual results are shown in the supplementary video for slow motion by 2 and 4.

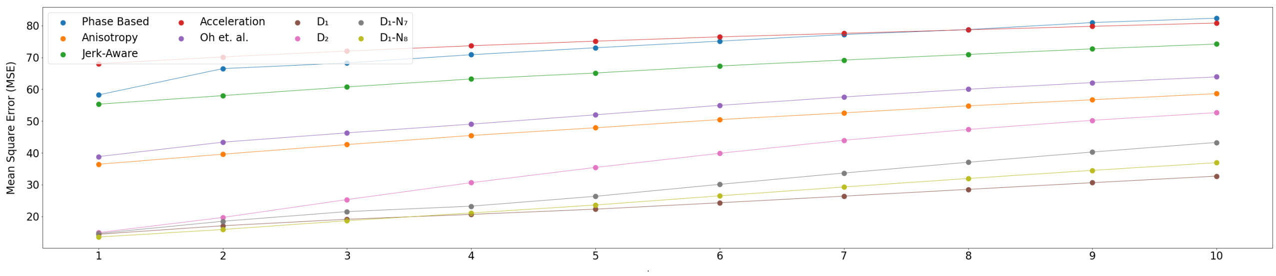

Figure S 1: Effects of increase in noise value (sigma) in input. The average mean square error (MSE) is computed across the predicted output and ground truth, over 25 different videos. Comparison is done with the Anisotropy method [4], Jerk-aware method [5], Acceleration method [3], Oh *et al.* method [1], Phase based method [2], the proposed model D_1, D_1-N_7 , D_1-N_8 and D_2. D_1-N_7 and D_1-N_8 are the D_1 models trained without amplitude and phase manipulator respectively.

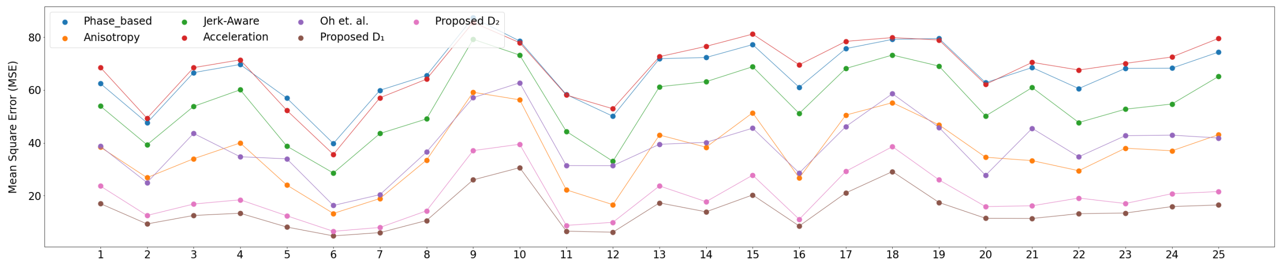

Figure S 2: Mean Square Error (MSE) of Anisotropy method [4], Jerk-aware method [5], Acceleration method [3], Oh *et al.* method [1], Phase based method [2] and the proposed methods on 25 synthetically generated videos containing different subtle motion of circles with various backgrounds.

Table 1: Parameters used for result generation. All the results are generated with variables and steps given by the respective authors. Source code and pre-trained model are downloaded from their official page, click here [25] [19] [18] [23].

Methods	**Video**	**Magnification Factor**	**Frequency**
Ours (D_1, D_2)	Gun	5	N/A
Ours (D_1, D_2)	Drill	5	N/A
Ours (D_1, D_2)	Balloon	5	N/A
Ours (D_1, D_2)	baby	20,9	N/A
Ours (D_1, D_2)	guitar	10	N/A
Ours (D_1, D_2)	thermocol beads	5	N/A
Ours (D_1, D_2)	Physical Accuracy	10	N/A
Ours(D_1, D_2)	Circle videos with different backgrounds	60	N/A
Ours (D_1, D_2)	Cat toy (for $\times 2$ slow motion)	0.5	N/A
Ours (D_1, D_2)	Cat toy (for $\times 4$ slow motion)	0.25,0.5,0.75	N/A
Oh et al	Gun	4	N/A
Oh et al	Drill	10	N/A
Oh et al	Balloon	10	N/A
Oh et al	baby	20	2.5 Hz
Oh et al	thermocol beads	5	N/A
Oh et al	Physical Accuracy	5	N/A
Oh et al	Circle videos with different backgrounds	60	N/A
Jerk-Aware	Gun	10	20
Jerk-Aware	Drill	25	3
Jerk-Aware	Balloon	25	3
Jerk-Aware	baby	50	2.5
Jerk-Aware	thermocol beads	25	3
Jerk-Aware	Physical Accuracy	20	15
Jerk-Aware	Circle videos with different backgrounds	200	15
Anisotropy	Gun	100	20
Anisotropy	Drill	100	3
Anisotropy	Balloon	100	3
Anisotropy	baby	150	2.5
Anisotropy	thermocol beads	100	3
Anisotropy	Physical Accuracy	200	3
Anisotropy	Circle videos with different backgrounds	400	15
Acceleration	Gun	10	20
Acceleration	Drill	4	3
Acceleration	Balloon	4	3
Acceleration	baby	100	2.5
Acceleration	thermocol beads	4	3
Acceleration	Physical Accuracy	20	15
Acceleration	Circle videos with different backgrounds	200	15

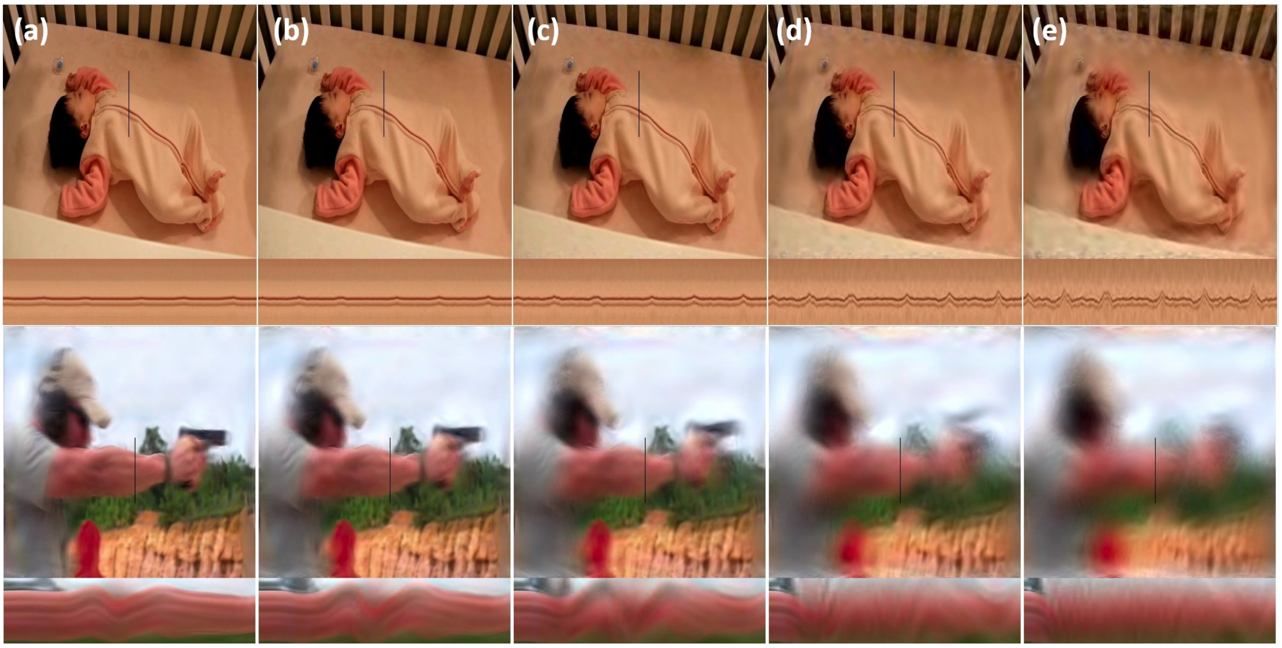

Figure S 3: **Effects of change in Magnification Factor.** Figure illustrates Acceleration method [3] output. Different values of the magnification factor in increasing order from (a) 20, (b) 40, (c) 60, (d) 100, and (e) 200 are used to generate the output shown in the respective column. As, increase in magnification factor leads to more increase in distortions than a small increment in magnification, especially in dynamic scenarios.

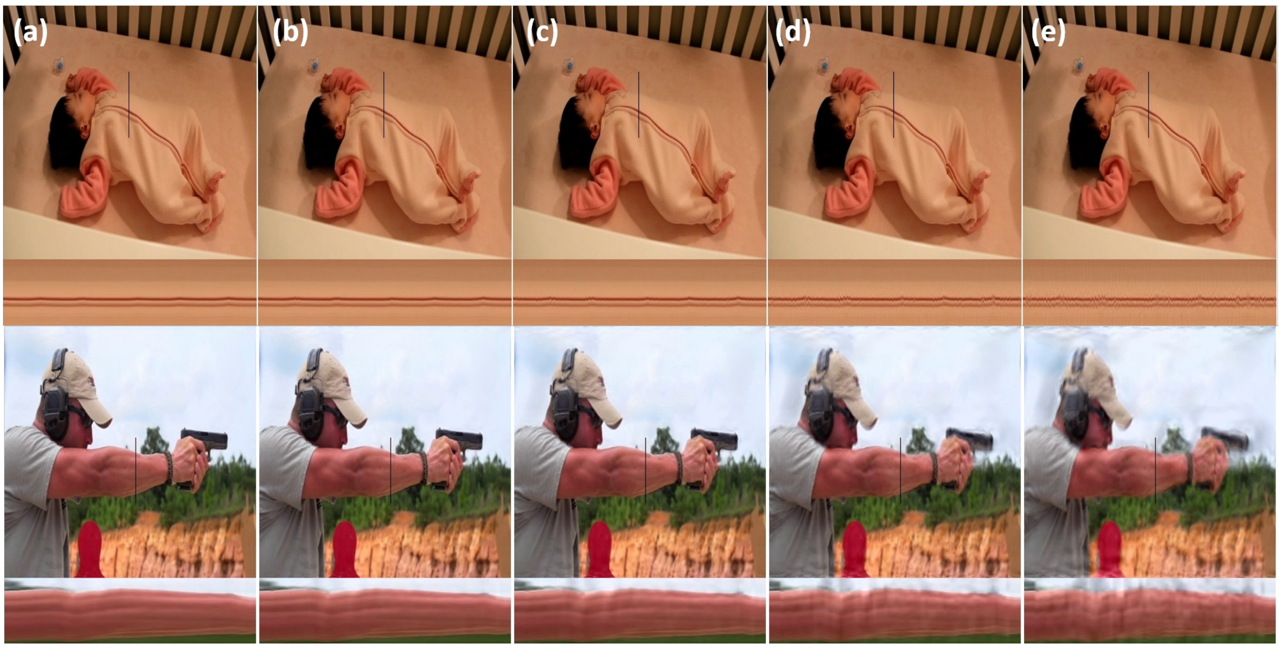

Figure S 4: **Effects of change in Magnification Factor.** Figure illustrates Anisotropy method [4] output. For columns, (a) 50, (b) 100, (c) 200, (d) 500, and (e) 1000, respective magnification factor values are used to generate magnified output. As visible from the figure, with an increase in magnification factor, there are minute changes in magnification while increments in distortions (especially in dynamic scenarios).

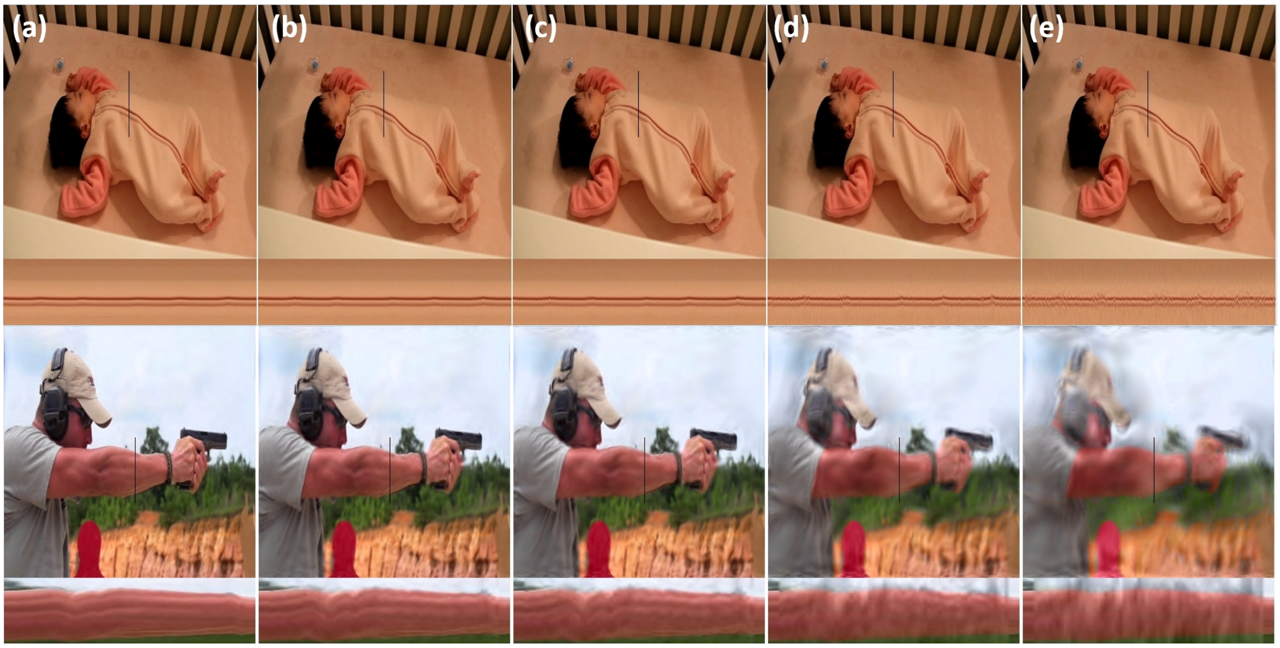

Figure S 5: **Effects of change in Magnification Factor.** Figure illustrates Jerk-Aware method [5]. Different values of the magnification factor in increasing order from (a) 10, (b) 30, (c) 50, (d) 150, and (e) 400 are used to generate the output shown in the respective column. As visible from the figure, with an increase in magnification factor, there are minute changes in magnification while increments in distortions (especially in dynamic scenarios).

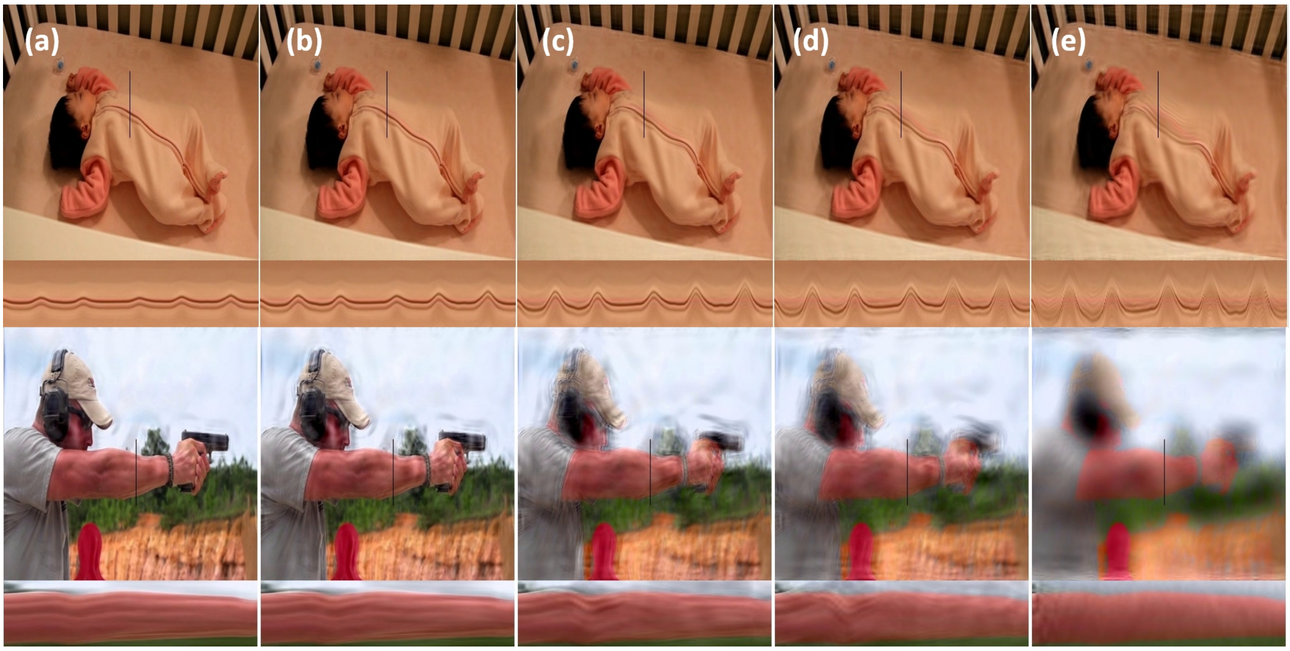

Figure S 6: **Effects of change in Magnification Factor.** Figure illustrates Phase based method [2] output. Different values of magnification factor in increasing order from (a) 10, (b) 20, (c) 40, (d) 60, and (e) 100 for baby video and (a) 1, (b) 2, (c) 5, (d) 10, and (e) 100 for gun video are used to generate the output shown in respective column (note:- for other methods same values are used for both the videos). The linear methods are not suitable for dynamic scenarios, as they are unable to ignore dynamic motion. So, they produce large distortions in the gun video (dynamic scenarios). Whereas in static scenario (baby videos), with an increase in magnification factor there is an increment in both, the amount of magnification and ringing artifacts (visible as lines overlapping the edges of moving objects) in the static scenario.

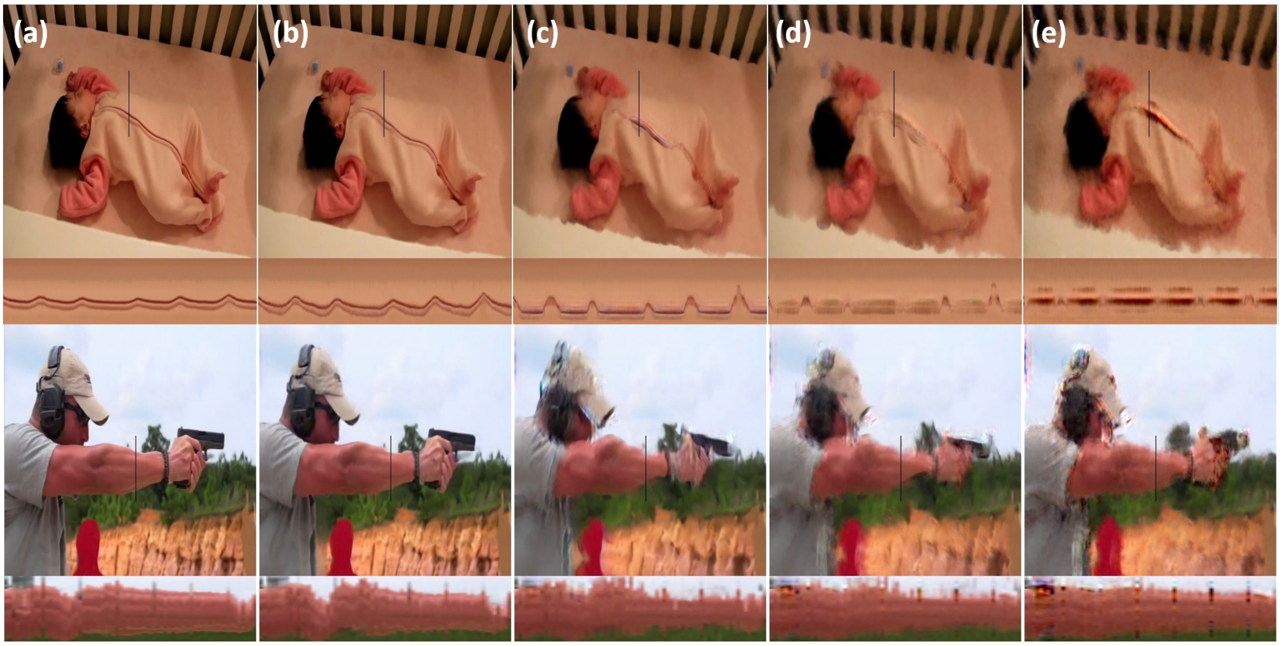

Figure S 7: **Effects of change in Magnification Factor.** Figure illustrates Oh *et al.* method [1] output. Different values of magnification factor in increasing order from (a) 10, (b) 20, (c) 50 (d) 100, and (e) 200 are used to generate the output shown in the respective column. It produces more magnification, (both in static and dynamic scenarios), but it also produces some unwanted motion (visible as large spikes in the temporal slice) and blurry distortions in the video. Distortions are increased with inclemently in the magnification factor.

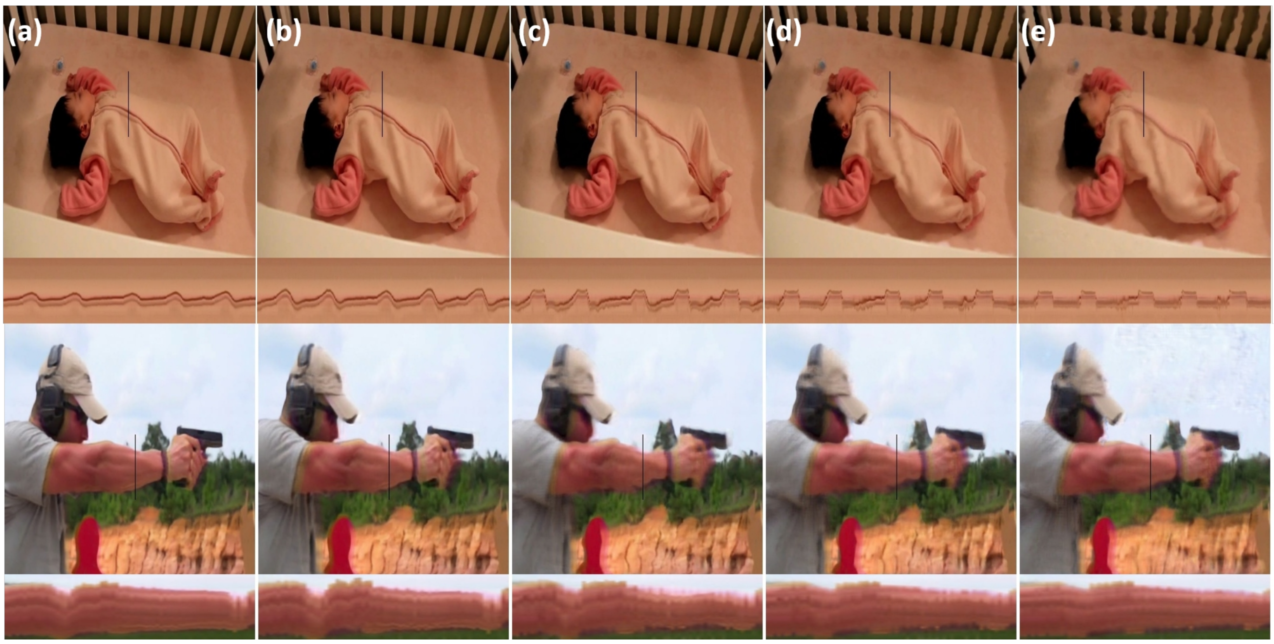

Figure S 8: **Effects of change in Magnification Factor.** Figure illustrates proposed D_1 model output. Different values of the magnification factor in increasing order from (a) 10, (b) 20, (c) 50 (d) 100, and (e) 200 are used to generate the output shown in the respective column. D_1 shows fewer distortions while increasing the amount of magnification as compared to other SOTA methods, both in static and dynamic scenarios.

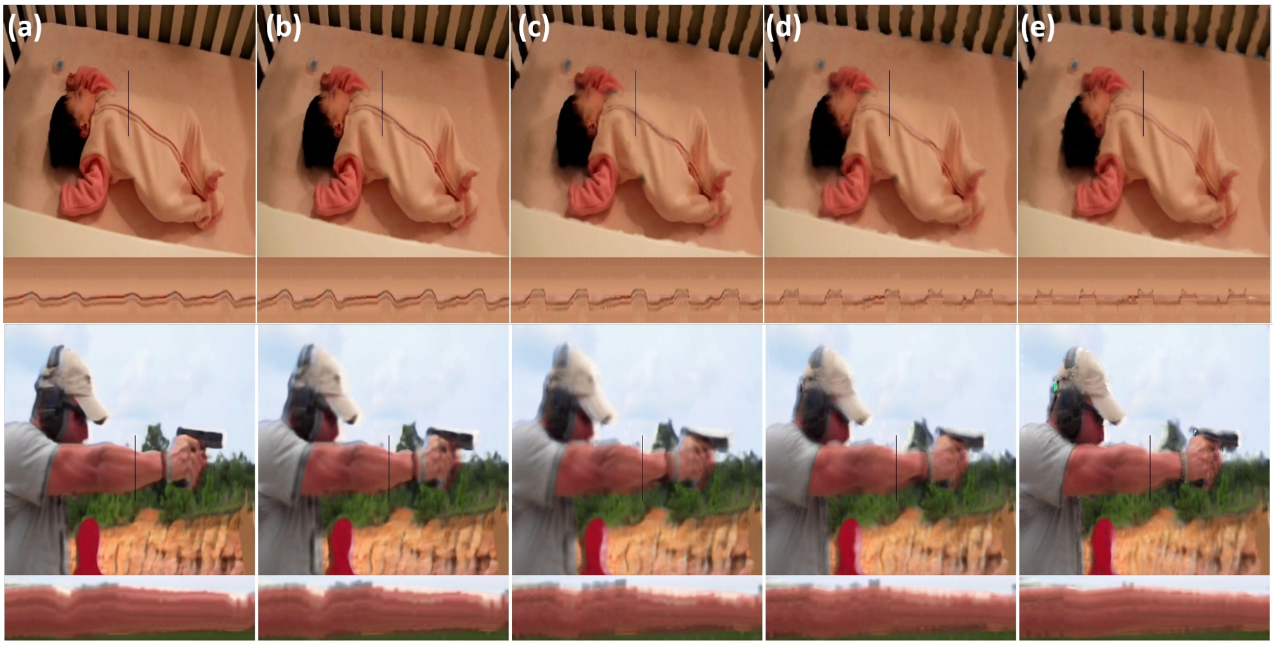

Figure S 9: **Effects of change in Magnification Factor.** Figure illustrates ours D_2 model output. Different values of magnification factor in increasing order from (a) 10, (b) 20, (c) 50 (d) 100, and (e) 200 are used to generate the output shown in the respective column. D_2 also shows a good amount of magnification, but with an increase in magnification factor, its performance degrades as compared to D_1. This is expected as D_2 has much fewer parameters than D_1, so their performance gap becomes observable in extreme scenarios.

Figure S 10: Different backgrounds used for generation of different synthetic videos for quantitative analysis. Video 1-10

Figure S 11: Different backgrounds used for generation of different synthetic videos for quantitative analysis. Video 11-20.

Figure S 12: Different backgrounds used for generation of different synthetic videos for quantitative analysis. Video 20-25.

References

[1] T.-H. Oh, R. Jaroensri, C. Kim, M. Elgharib, F. Durand, W. T. Freeman, and W. Matusik, "Learning-based video motion magnification," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 633–648.

[2] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Phase-based video motion processing," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 1–10, 2013.

[3] Y. Zhang, S. L. Pintea, and J. C. Van Gemert, "Video acceleration magnification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 529–537.

[4] S. Takeda, Y. Akagi, K. Okami, M. Isogai, and H. Kimata, "Video magnification in the wild using fractional anisotropy in temporal distribution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1614–1622.

[5] S. Takeda, K. Okami, D. Mikami, M. Isogai, and H. Kimata, "Jerk-aware video acceleration magnification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1769–1777.