Incremental Activity Modeling and Recognition in Continuous Streaming Videos

Supplementary Materials

Overall performance comparisons averaged over all activity classes of different datasets are illustrated in Figure [5] of the main paper (MP). In this supplementary material (SM), we provide per-activity basis performance comparison on multiple datasets in Figures 1 - 7. We also provide evolution of the confidence scores of some individual test actions of datasets KTH, UCF11, and VIRAT due to incremental learning in Figures 8 - 10 respectively.

In most of these experiments, overall performances of the incremental learning methods are asymptotically increasing and cross Baseline-2, in some cases touch Baseline-1. These results demonstrate the effectiveness of our proposed incremental learning framework, which requires less amount of tedious manual labeling.

Expt. No.	Dataset	Feature	Comments
1. Page 3 MP Figure [5] SM Figure 1	KTH	STIP	• For some activities like boxing, handclapping, hadwaving, and jogging, performance of IL methods are asymptotically increasing and cross Baseline-1 and Baseline-2. • For activities running and walking, performance is also asymptotically increasing but do not cross Baseline-1 and Baseline-2. • Performance curve is parabolic for handwaving and running.
2. Page 4 MP Figure [5] SM Figure 2	KTH	Gist3D	• For activities such as boxing, handclapping, and walking performance of both of the IL methods cross the performance of Baseline-1. • For activity jogging performance of IL methods cross the performance of Baseline-2. • For running, tennis, and walk_dog performance of the IL methods are below the Baseline-1 and Baseline-2.
3. Page 5 MP Figure [5] SM Figure 3	KTH	Action Bank	• For activities boxing, running, jogging, and walking performances of both of the IL methods are same as the Baseline-1. • For activity handclapping, performance of IL-unlabeled is better than all three counterparts and asymptotically increasing. • For activity handwaving, performance of the IL methods are below Baseline-2.
4. Page 6 MP Figure [5] SM Figure 4	UCF11	STIP	• For activities such as diving, golf, jumping, riding, shooting, and swing performance of both of the IL methods cross the performance of Baseline-. • For activities such as biking and juggle performance of IL methods touches the performance of Baseline-2. • For spiking, tennis, and walk_dog performance of the IL methods are below Baseline-1 and Baseline-2.
5. Page 7 MP Figure [5] SM Figure 5	UCF11	Gist3D	• For activities diving, golf, riding, spiking, swing performance of both of the IL methods are asymptotically increasing and cross the performance of Baseline-1 and Baseline-2. • For activities biking, swing, and walk_dog performance of both of the IL methods cross Baseline-2. • For activities juggle, jumping, and shooting performance of both of the IL methods are below Baseline-2.

6. Page 8 MP Figure [5] SM Figure 6	UCF11	Action Bank	• For some activities such as juggle, riding, shooting, and walk_dog performance of both of the IL methods are asymptotically increasing and cross the performance of Baseline-1 and Baseline-2. • For activities biking, diving, and swing performance of both of the IL methods cross Baseline-2. • For activities golf, jumping, spiking, and tennis, performance of both of the IL methods are below Baseline-2.
7. Page 9 MP Figure [5] SM Figure 7	KTH	STIP	• For activities loading, opening, and closing performance of both of the IL methods are close to zero because number of segmented activities available for incremental learning are very small. • For activities into the vehicle and entering facility performance of the IL methods are asymptotically increasing and cross the Baseline-1. • For activities out from vehicle and exiting facility, performance of both of the IL methods cross Baseline-1 at some point, after that it decreases. • For activity unloading performance of both of the IL methods are below Baseline-2.

Activity-wise Performance.
Dataset: KTH, Feature: STIP

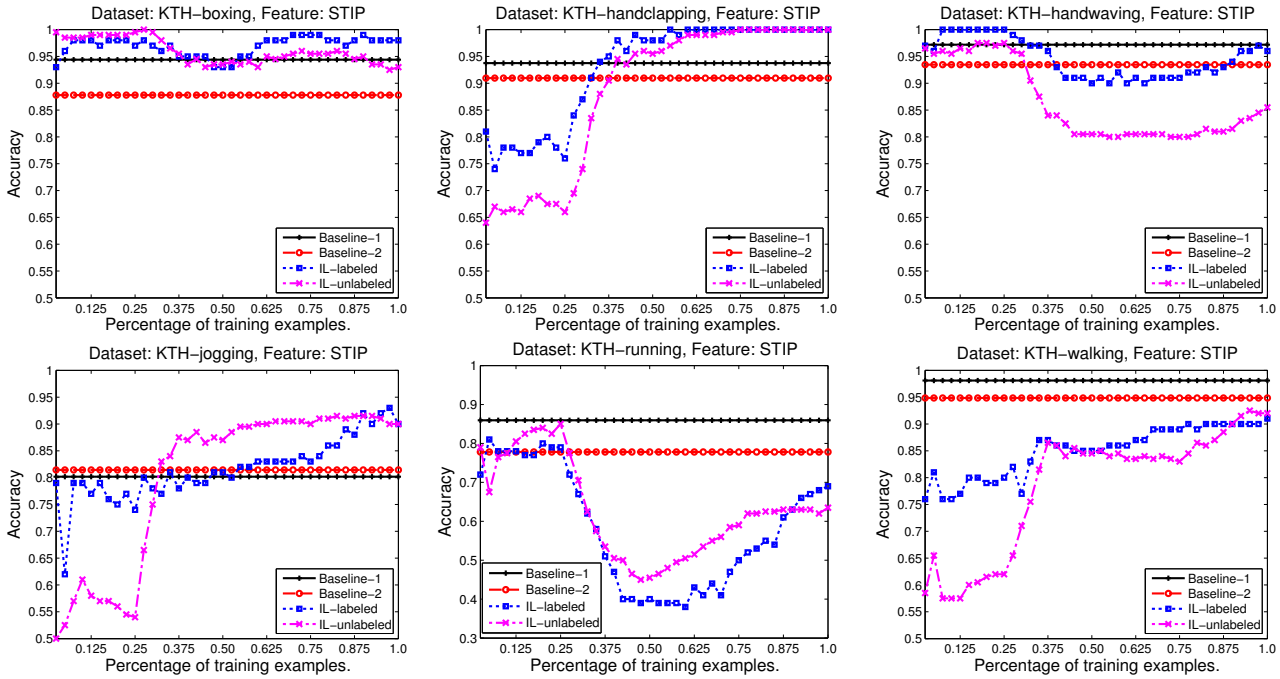

Fig. 1: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **KTH** dataset using **STIP** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

Activity-wise Performance.
Dataset: KTH, Feature: Gist3D

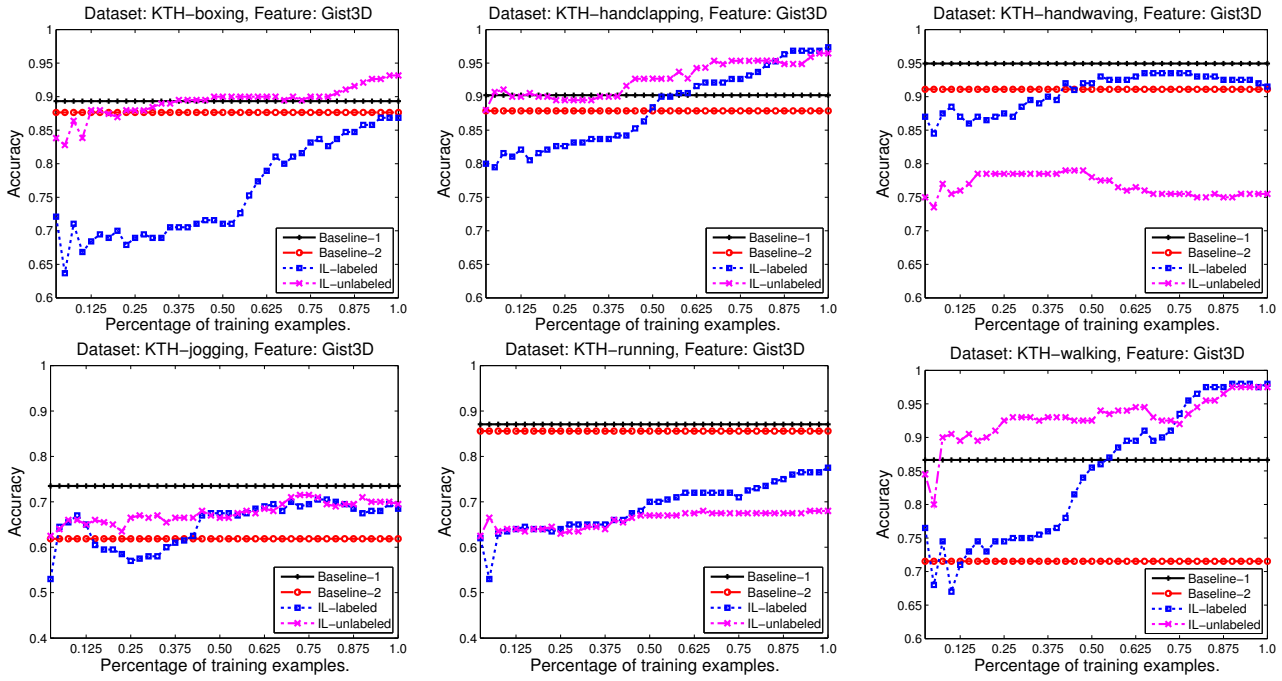

Fig. 2: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **KTH** dataset using **Gist3D** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

Activity-wise Performance.
Dataset: KTH, Feature: Action Bank

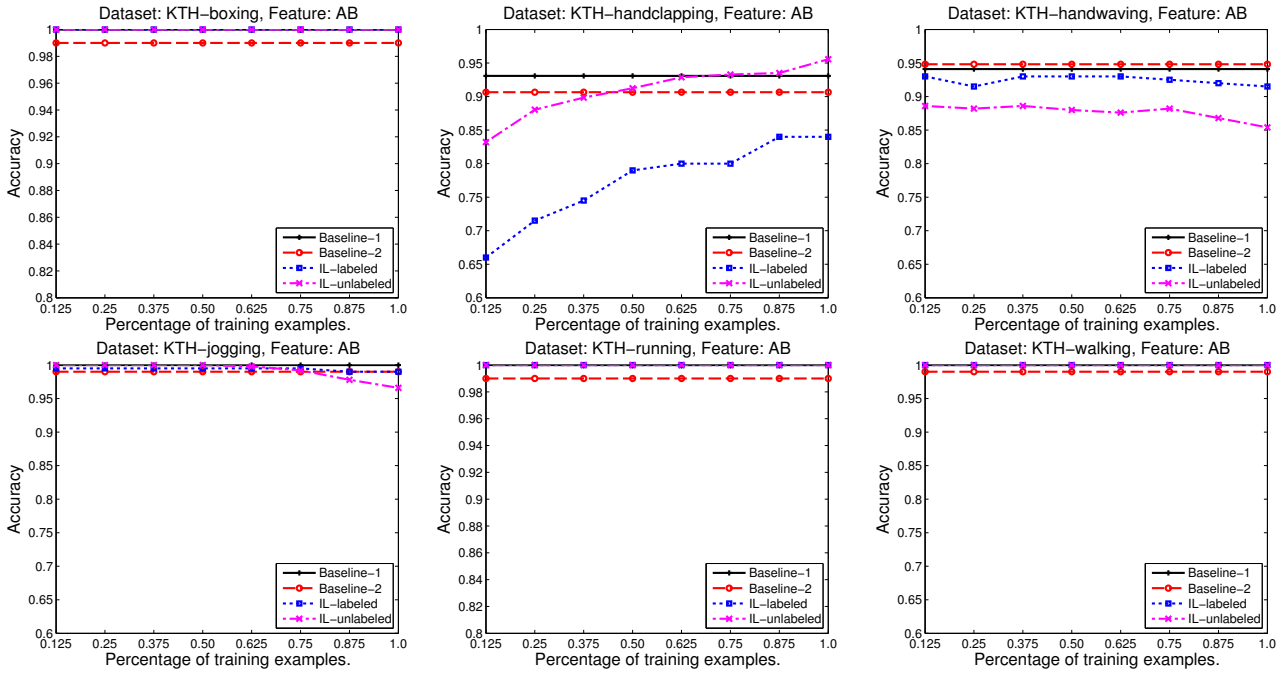

Fig. 3: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **KTH** dataset using **Action Bank** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

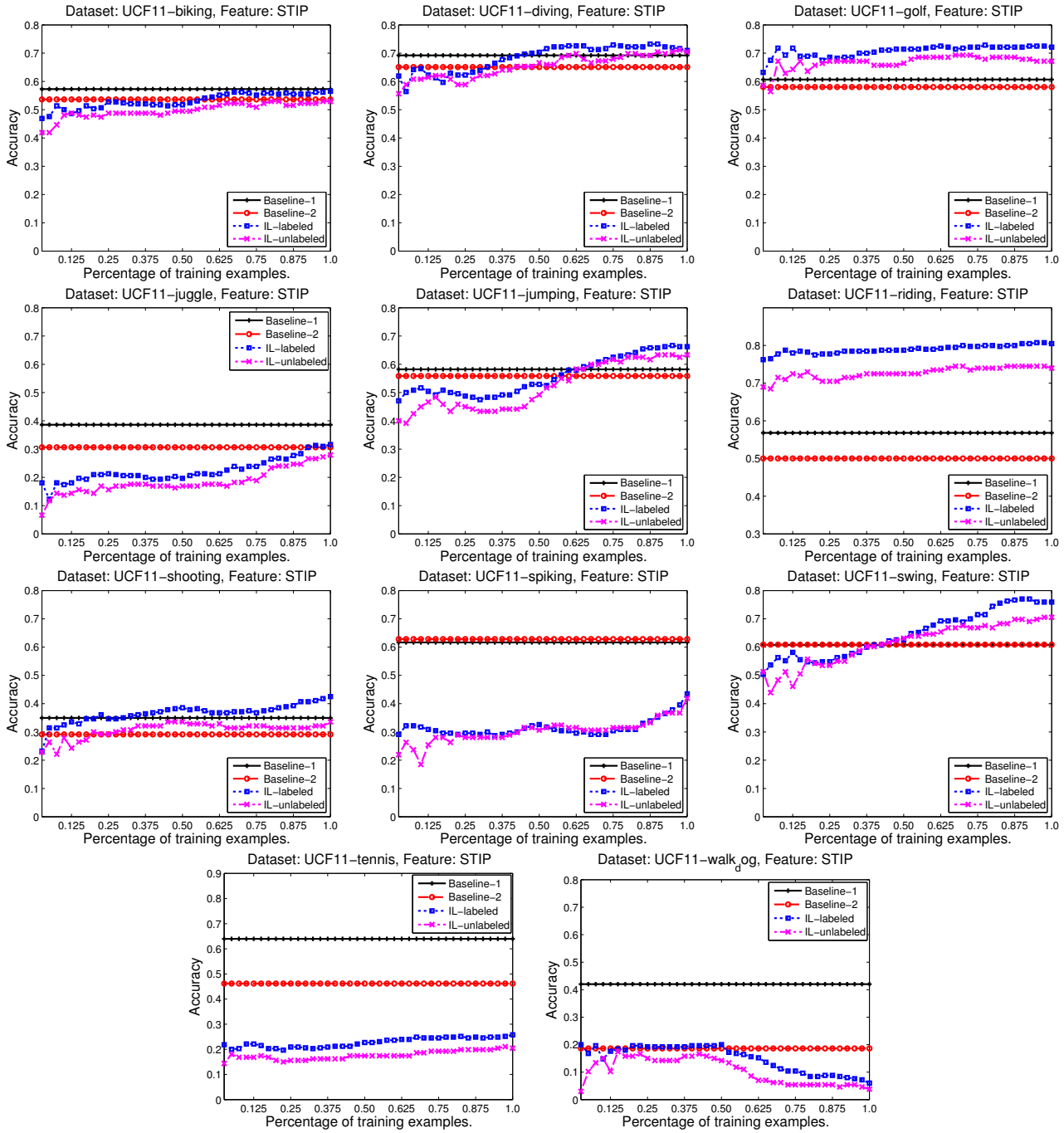

Fig. 4: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **UCF11** dataset using **STIP** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

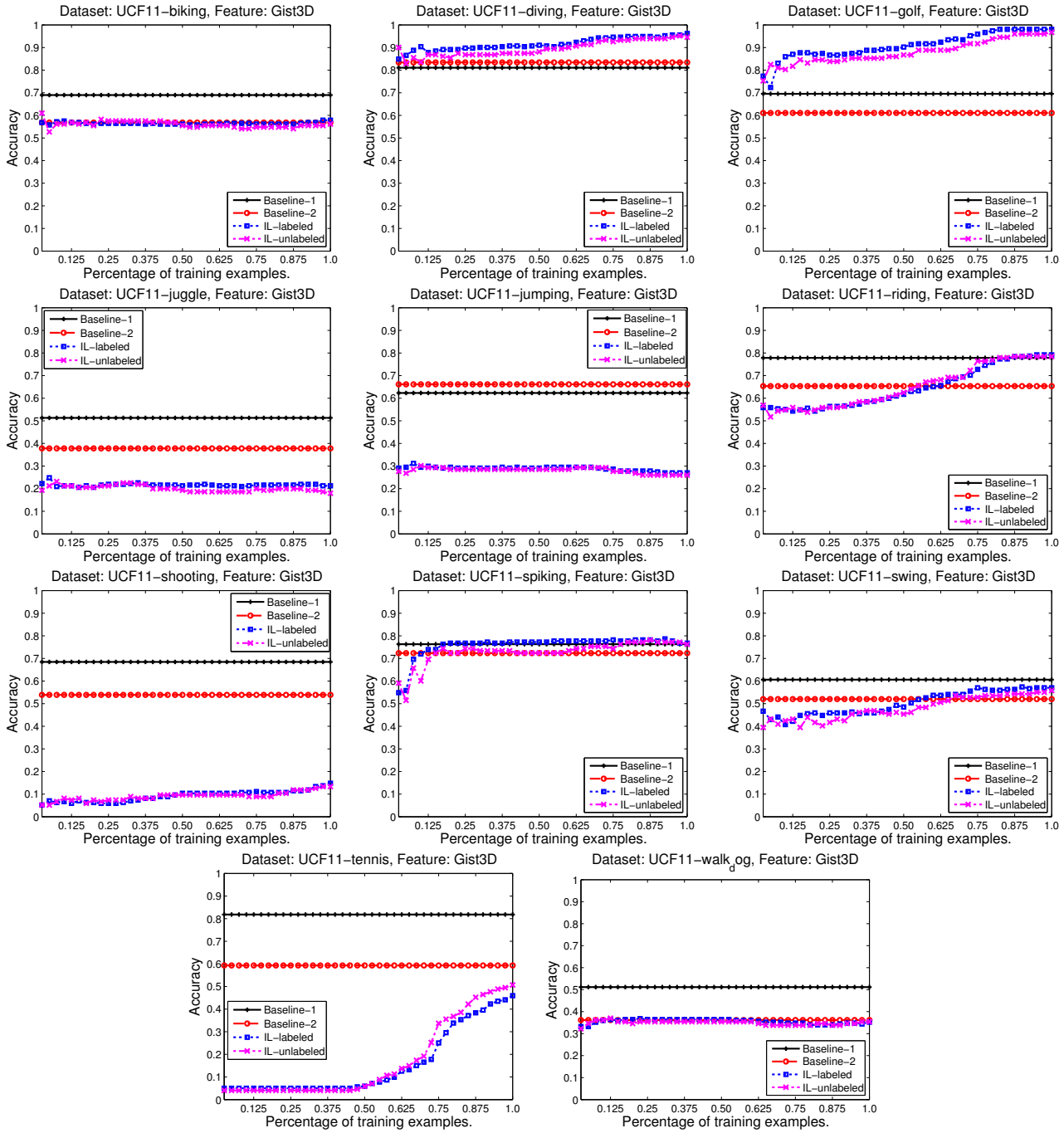

Fig. 5: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **UCF11** dataset using **Gist3D** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

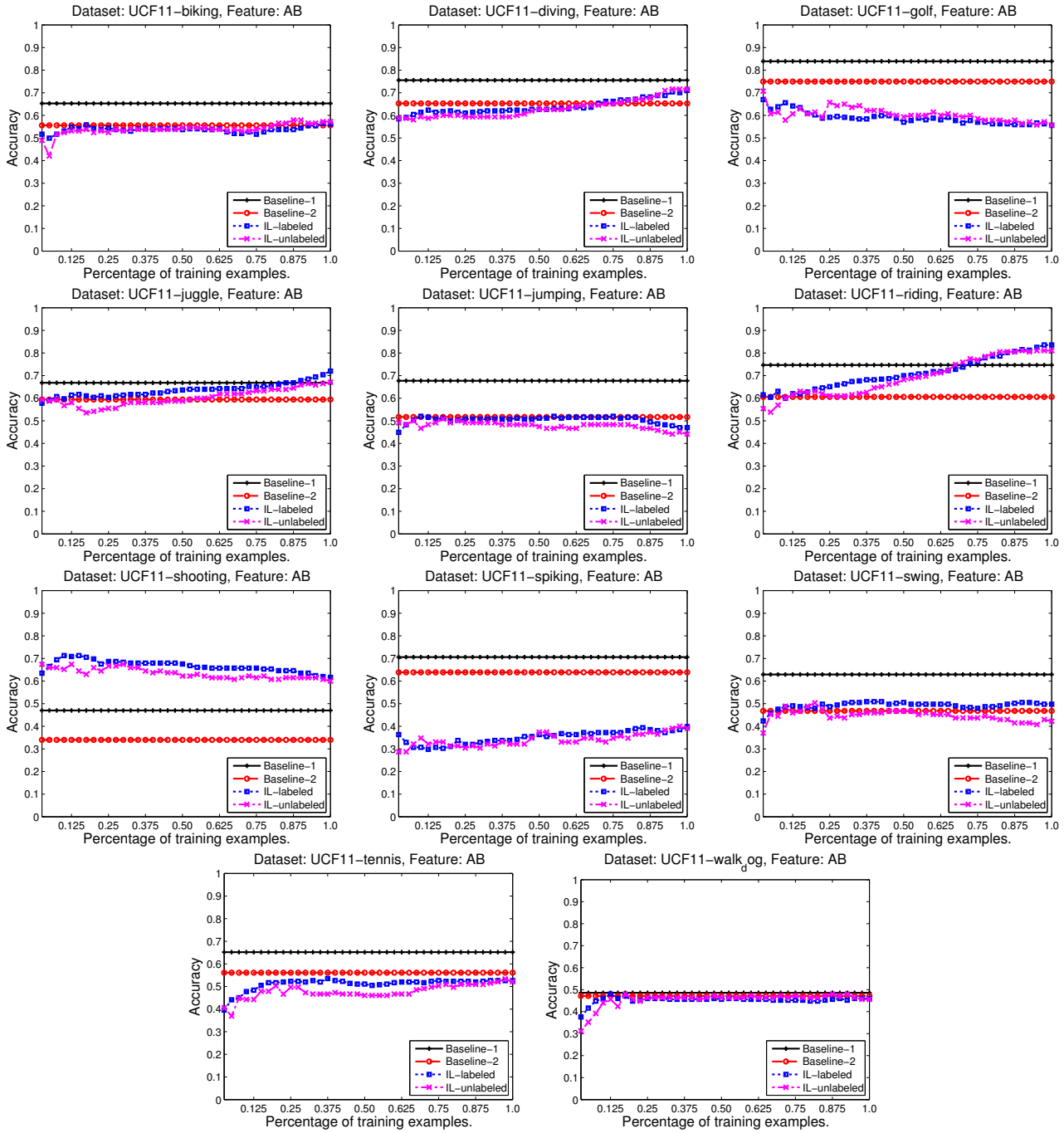

Fig. 6: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **UCF11** dataset using **Action Bank** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

Activity-wise Performance.
Dataset: ViRAT, Feature: STIP

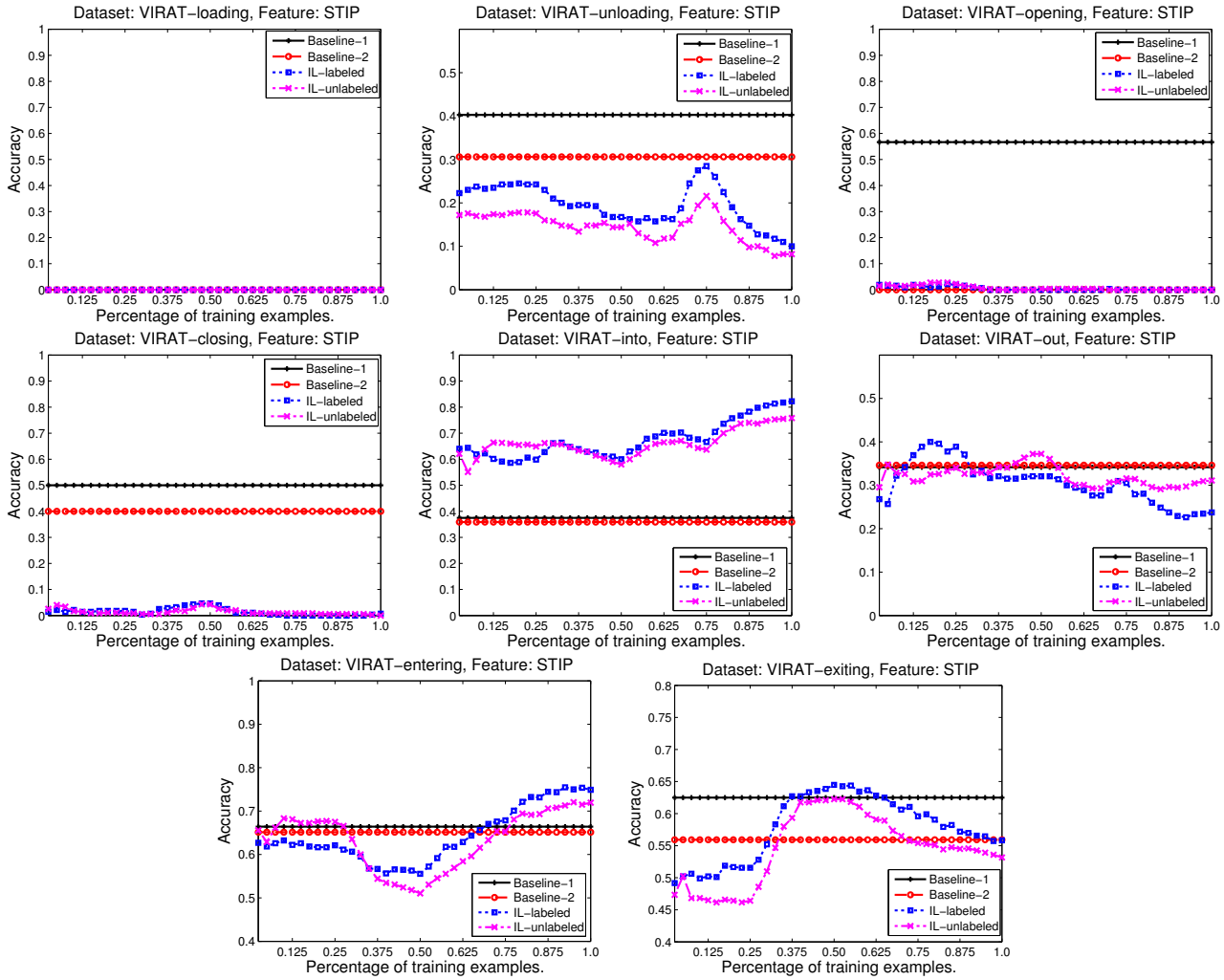

Fig. 7: Activity-wise performance comparison of our proposed method IL-unlabeled with Baseline-1, Baseline-2, and IL-labeled on **VIRAT** dataset using **STIP** feature. X-axis is the fraction of examples presented so far to the incremental learning framework and Y-axis is the accuracy of the classification. Plots are best viewable in color.

Evolution of the confidence scores of some individual test actions of datasets KTH, UCF11, and VIRAT due to incremental learning are shown in Figures 8, 9, and 10 respectively.

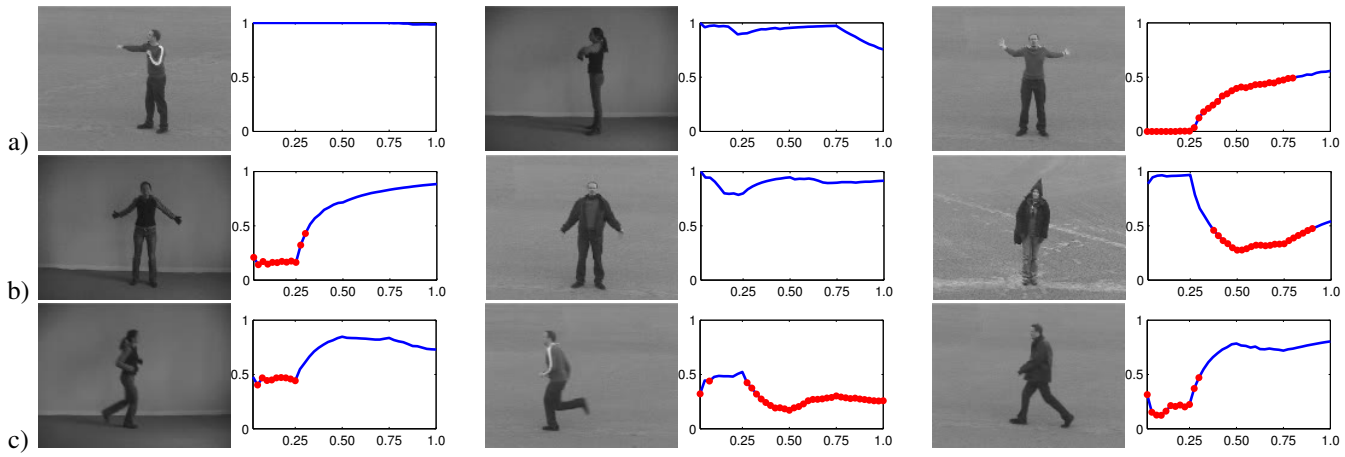

Fig. 8: This figure shows the performance of the proposed incremental activity modeling framework on individual test action clips of **KTH** dataset. Above illustrated actions are as follows (left to right, top to bottom): a) boxing-1, boxing-2, handclapping-1; b) handclapping-2, handwaving-1, handwaving-2, c) jogging, running, walking. X-axis is the fraction of the examples presented so far to the incremental learning framework and Y-axis is the normalized confidence score $\mathcal{H}(\mathbf{x})$. Blue line means correct classification of the action, while red spots means missclassificaiton of the action at that particular instant. In most of the cases, our proposed incremental learning framework increases the confidence score of an action and can retain the correct classification; in some cases, updated model rectifies the missclassifications (red to blue). In some rare cases(like running and handwaving-2), our framework failed to perform well and missclassified an action even though it was correctly classified before (blue to red). Plots are best viewable in color.

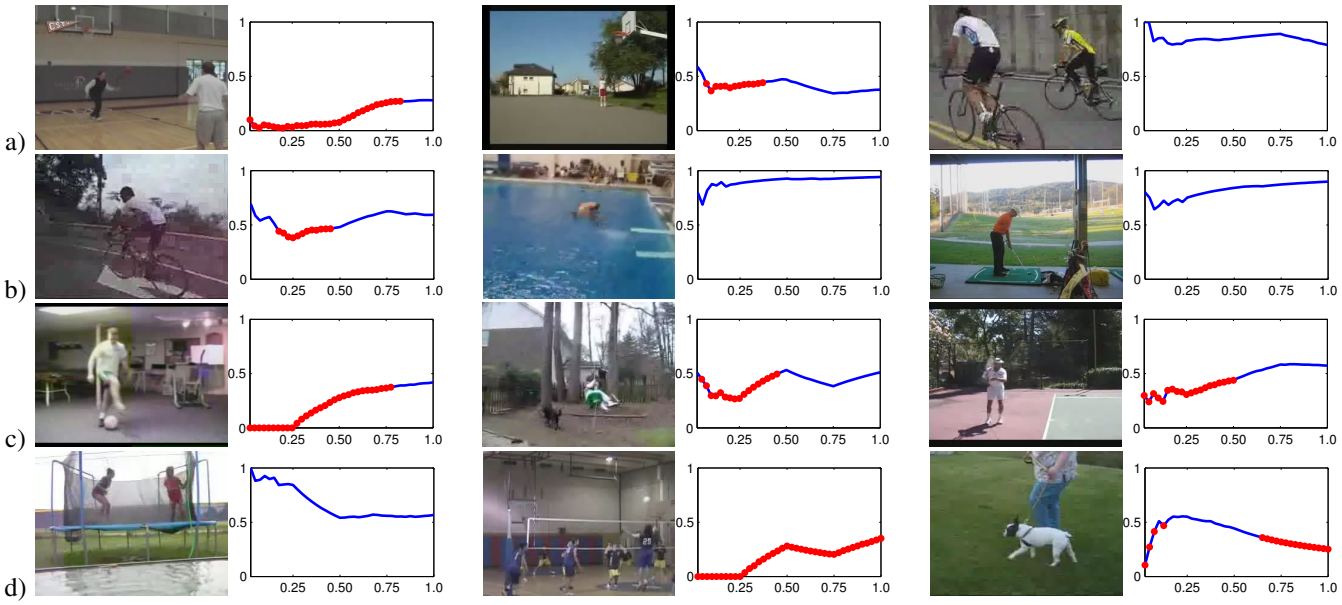

Fig. 9: This figure shows the performance of the proposed incremental activity modeling framework on individual test action clips of **UCF11** dataset. Above illustrated actions are as follows (left to right, top to bottom): a)basketball-1, basketball-2, biking; b) biking, diving, golf_swing; c) soccer_juggling, swing, tennis_swing; d) tramploline_swing, volleyball_spiking, walking. X-axis is the fraction of the examples presented so far to the incremental learning framework and Y-axis is the normalized confidence score $\mathcal{H}(\mathbf{x})$. Blue line means correct classification of the action, while red spots means missclassificaiton of the action at that particular instant. In most of the cases, our proposed incremental learning framework increases the confidence score of an action and can retain the correct classification; in some cases, updated model rectifies the missclassifications (red to blue). In some rare cases(volleyball_spiking and walking), our framework failed to perform well and missclassified an action even though it was correctly classified before (blue to red). Plots are best viewable in color.

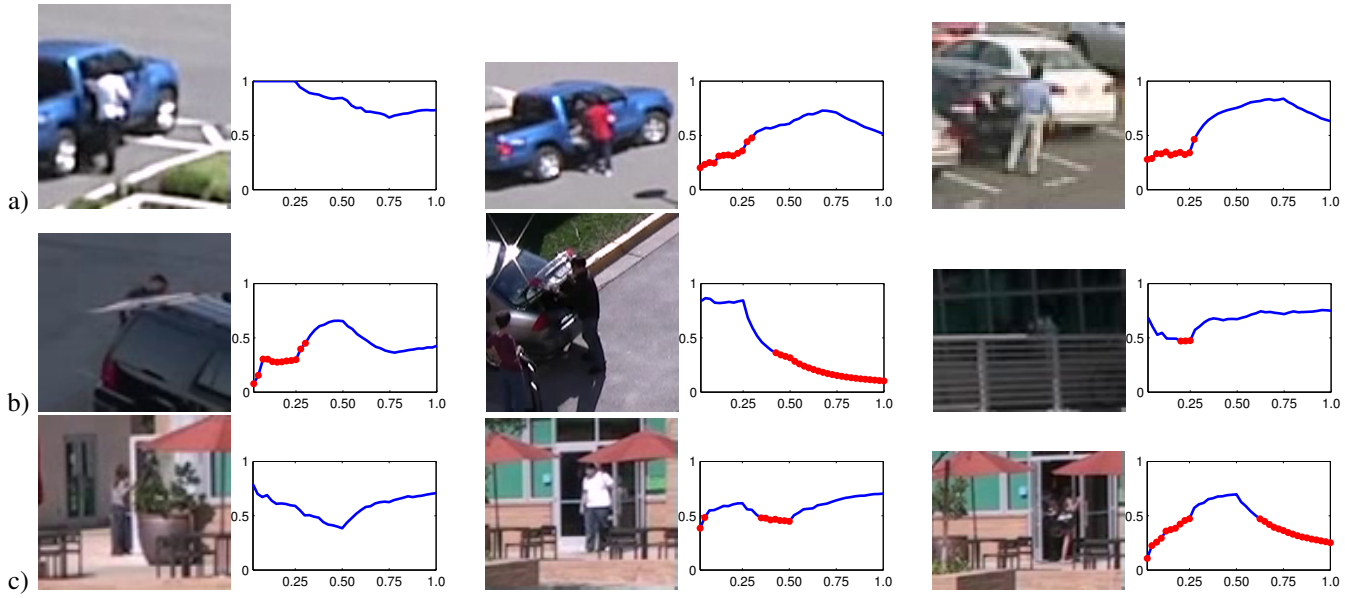

Fig. 10: This figure shows the performance of the proposed incremental activity modeling framework on individual test action clips of **VIRAT** dataset. Above illustrated actions are as follows (left to right, top to bottom): a) person unloading an object from a vehicle-1, person unloading an object from a vehicle-2, person openning a vehicle trunk-1; b) person openning a vehicle trunk-2, person closing a vehicle trunk, person entering a facility-1; c) person entering a facility-2, person exiting a facility-1, and person exiting a facility-2. X-axis is the fraction of the examples presented so far to the incremental learning framework and Y-axis is the normalized confidence score $\mathcal{H}(\mathbf{x})$. Blue line means correct classification of the action, while red spots means missclassificaiton of the action at that particular instant. In most of the cases, our proposed incremental learning framework increases the confidence score of an action and can retain the correct classification; in some cases, updated model rectifies the missclassifications (red to blue). In some rare cases(person closing a vehicle trunk and person exiting a facility-2), our framework failed to perform well and missclassified an action even though it was correctly classified before (blue to red). Plots are best viewable in color.