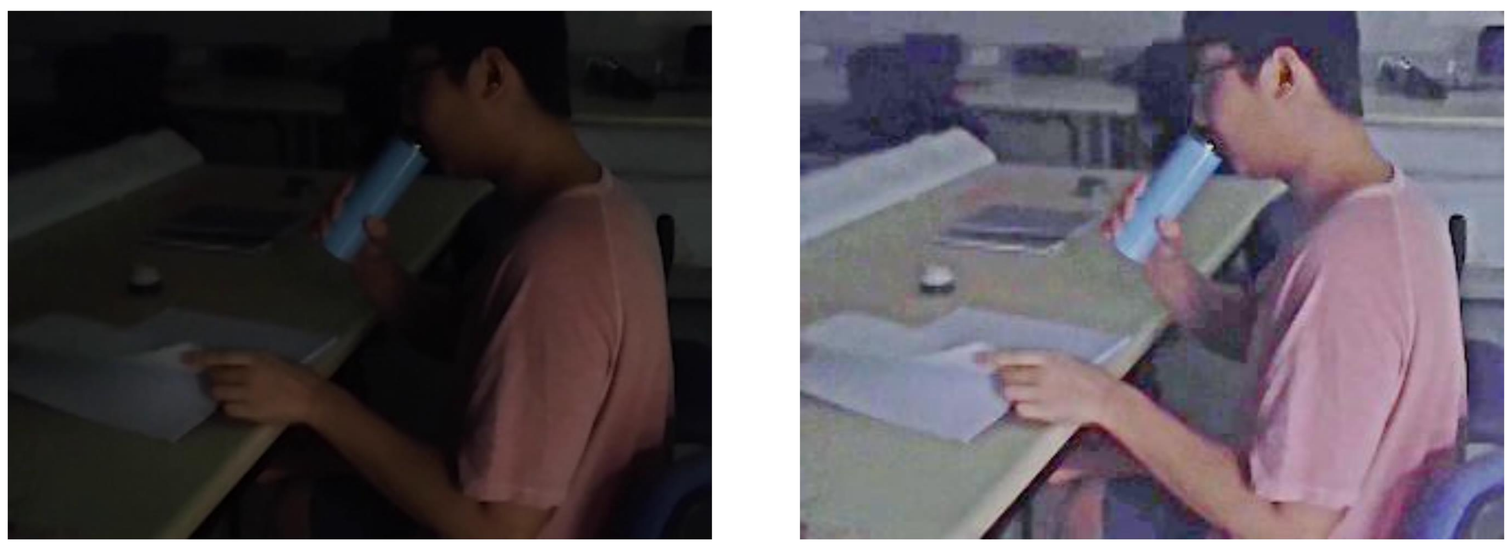
DL-KDD Dual-Lightness Knowledge Distillation for Action Recognition in the Dark

Chi-Jui Chang, Oscar Tai-Yuan Chen, Vincent S. Tseng*

Background

- **Action Recognition in the Dark (Dark HAR)** is a challenging task due to the degradation of the information in the videos. The three main challenges on Dark HAR are:
  - → **Information Completeness:** Ensuring the model to learn from enhanced video with essential information from the original video preserved.
  - → **Complexity Tradeoff:** Making full use of original video and considering enhanced features (extracted from enhanced video) without additional model complexity for input features.
  - → **Consistent Performance:** Improving the performance even if using only original video without enhancement as input during the inference phase.

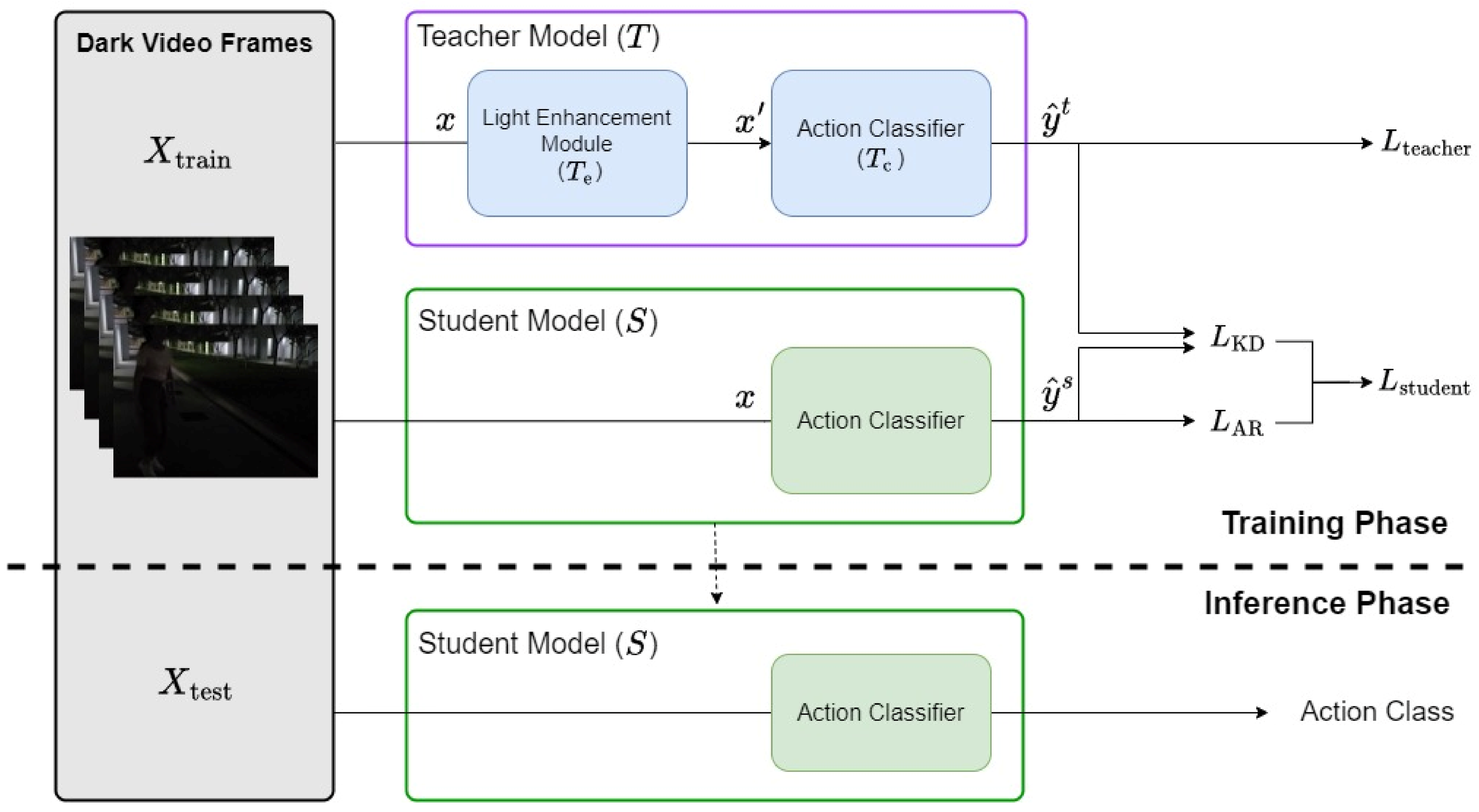


A comparison between the original frame and the enhanced frame.

Contributions

- To this end, we propose a novel knowledge-distillation-based architecture named ***Dual-Lightness Knowledge Distillation (DL-KDD)***. The main contributions of this work are threefold:
  - → By leveraging both the original video and the enhanced features during training, our approach facilitates more effective knowledge transfer.
  - → Our model can use only original video input without additional features or enhancement during inference, making it more efficient in real-world applications.
  - → We achieve state-of-the-art performance in dark video human action recognition across widely-used datasets with the model complexity maintained.

DL-KDD Framework



- The Dual-Lightness Knowledge Learning contains three steps:
  - → **Enhanced Frame Extraction**
    Training the teacher model with the video frames processed by the light enhancement module.
  - → **Original Representation Learning**
    Training the student model with the original video frames as inputs to learn the original representation of the videos.
  - → **Dual Knowledge Learning**
    The student model is also trained using the outputs from the teacher model as soft targets, which allows the student model to learn from the feature representation that has been extracted from enhanced frames by teacher model.

Results

Model	Method	Input Feature		Top-1 Accuracy (%)
		Training	Inference	
I3D-RGB	-	Original	Original	68.29
I3D Two-stream	Two-Stream	Original + Optical Flow	Original + Optical Flow	72.78
3D-ResNet-101	-	Original	Original	71.57
3D-ResNext-101	-	Original	Original	74.73
Video-Swin-B	-	Original	Original	89.79
DarkLight	Two-Stream	Original + GIC	Original + GIC	94.04
DTCM	-	ZeroDCE	ZeroDCE	96.36
R(2+1)D-GCN+BERT	-	ZeroDCE	ZeroDCE	96.60
DL-KDD (Ours)	KD	Original + ZeroDCE	**Original**	**97.64**

Table 1: Comparison to state-of-the-art (SOTA) methods on the ARID dataset with detailed method descriptions. The 'Method' column specifies the approach used to integrate multiple features, where a dash (-) indicates models that do not utilize multiple features. The 'Input Feature' section details the data used for training and inference, specifying any enhancement techniques. Finally, the 'Top-1 Accuracy' column presents the classification performance of each model.

Model	Top-1 Accuracy (%)
I3D-RGB	48.75
I3D Two-stream	51.24
DarkLight	84.13
R(2+1)D-GCN+BERT	86.93
DL-KDD (Ours)	**88.12**

Table 2: Comparison to state-of-the-art (SOTA) methods on the ARID V1.5 dataset. The 'Top-1 Accuracy' column presents the classification performance of each model.

Model	Top-1 (%)	Top-5 (%)
I3D-RGB	32.25	65.35
3D-ResNext-101	37.23	68.86
DarkLight	42.27	70.47
DTCM	46.68	75.92
DL-KDD (Ours)	**52.26**	**80.18**

Table 3: Comparison to state-of-the-art (SOTA) models on the Dark-48 dataset. The 'Top-1 Accuracy' and 'Top-5 Accuracy' columns represent the classification performance of each model.

Acknowledgements

This research was supported in part by National Science and Technology Council Taiwan under grant no. NSTC 113-2223-E-A49-008 and 113-2634-F-A49-005.

Contact

Feel free to contact me for more information through my mail: jerryyyyy708.cs12@nycu.edu.tw, or https://jerryyyyy708.github.io/ for other personal links.