

HW9: Human Action Recognition

Zhenye Na (zna2)

(top1_accuracy, top5_accuracy, top10_accuracy): Did the results improve after combining the outputs?

Single Frame Model

top1_accuracy=0.785091, top5_accuracy=0.949247, top10_accuracy=0.973302

Sequence Model

Combined output of the previous two models

Use the confusion matrices to get the 10 classes with the highest performance and the 10 classes with the lowest performance: Are there differences/similarities? Can anything be said about whether particular action classes are discriminated more by spatial information versus temporal information?

Single Frame Model

10 classes with highest performances:

TrampolineJumping, FloorGymnastics, BreastStroke, WritingOnBoard, HorseRace, Surfing, SumoWrestling, BasketballDunk, Rowing, SkyDiving

10 classes with lowest performances:

BodyWeightSquats, JumpRope, HandstandWalking, HighJump, YoYo, Shotput, Nunchucks, PullUps, CricketShot, PommelHorse

Sequence Model

Combined output of the previous two models

Use the confusion matrices to get the 10 most confused classes. That is, which off-diagonal elements of the confusion matrix are the largest: Are there any notable examples?

Single Frame Model

Sequence Model

Combined output of the previous two models