

Stage2

Multi-task training

based on

Stage1 Use auxiliary

top-K classes