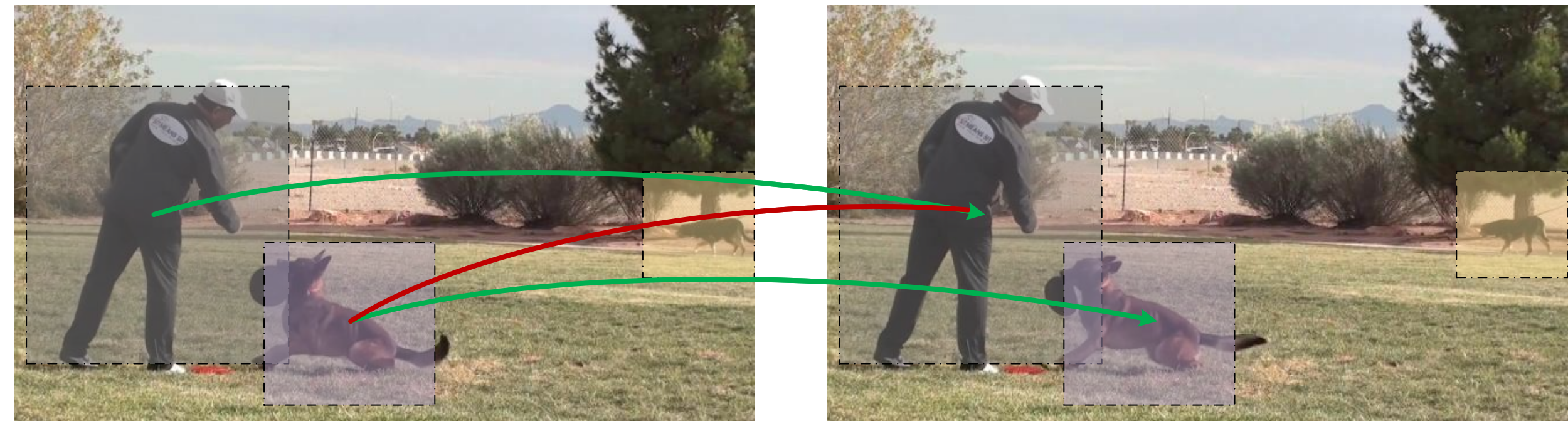


Target frame

Reference frame

Occlusion mask

**(a) Temporal Feature Learning**



Target frame

Reference frame

**(b) Spatial Feature Learning**