

Pose detection by Image Translation

Description:

Image translation is an important application of generative models in computer vision. A few real world problems, such as image segmentation [1], have been solved using image translation from real images to synthetic (segmentation labels). Besides, they have also been used extensively in order to close the domain gap for robust estimation on input data. There are several effective GAN models which solve image translation such as Cycle GANs [2] and its variants. Such models can be used to effectively translate images from one distribution to another. However, image translation in certain specific conditions can be done without using the fully probabilistic approach of GANs and instead using the 3D-2D relationship available in certain conditions. Robocup has one of these special conditions, where the flat known plane of the soccer field provides strict and constrained relationships for training a model for pose detection and image translation.

Goal

In the context of Robocup, where teams of NAO robots play soccer against each other in a small football field, image translation is potentially useful to solve many perception related tasks. Among many potential applications, one can easily obtain labeled synthetic images from various simulators. Similarly, unlabeled real data are available from previous recordings. The goal of the project is to identify the soccer field semantics using image translation by warping from real images to synthetic images.

[1] Shukla et al., “Extremely Weak Supervised Image-to-Image Translation for Semantic Segmentation”, *In ICCVW 2019*.

[2] Zhu et al., “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks”, *In ICCV 2017*.

Requirements:

Knowledge of deep learning and basic theory of generative adversarial networks.

Practical experience with deep learning libraries such as PyTorch.

Knowledge of camera projection and 3D vision is not necessary but appreciated.