



针对生成图像集合 G_i 其中 $i \in [0,9]$ ，对其进行关键点提取，提取的关键点就是 TG1 的经典 8 个关键点。为了实验准确度，本次采用手动比对法与 GT_i 的对应关键点像素点位置进行对比。结果如下

Frame	0	1	2	3	4	5	6	7	8	9
Error	0	0	0	0	0	0	0	0	<1	<1

Frame 是后续第几张图像，Error 是单个 Frame 中，8 个关键点的平均偏移几个像素，计算公式如下

$$Error_i = \sum_{j=1}^8 (GT_i^j - G_i^j) / 8$$

结论就是我们几乎生成了和 Ground Truth 完全一样的图像，几乎不存在误差，即使预测 10 帧之后的图像，8 个像素点也仅仅偏移了 1 个像素。