

# 实验部分

## 总览

### 一、数据集介绍

### 二、评估指标和实验设置

### 三、实验结果

#### 1. DCE 模块的消融实验

目标：2/3页

DCE先train好（当然Cstyle也要先train好，这个是用100Style数据集train的），frozen掉之后，取代MCM-LDM工作中的Style Remover，从头train HumanML3D数据集，轮次，学习率等参数与MCM-LDM保持一致，这样的话train完就可以和前面的工作做评估了。

Methods	FMD↓	CRA↑ (%)	SRA↑ (%)	TSI↓	FSF↓
Real Motions	—	99.24	100.00	—	—
1DConv+AdaIN [1]	42.68	31.18	57.00	0.22	2.05
STGCN+AdaIN [36]	129.44	<b>60.43</b>	17.66	<b>0.11</b>	<b>0.93</b>
Motion Puzzle [26]	113.31	26.31	46.33	0.22	2.43
Ours	<b>27.69</b>	35.75	<b>58.00</b>	0.40	1.28

Table 1. **Quantitative evaluation.** ‘↑’ (‘↓’) indicates that the value is better if the metric is larger (smaller); The **bold fonts** denote best performers. The results demonstrate that our MCM-LDM achieves balanced performance in all metrics.

DCE模块的训练过程（提一下学习率和epoch）：

- 1.训练一个Cstyle的分类器，做100Style数据集关于style的分类任务；
- 2.frozen Cstyle分类器，在100Style数据集上train DCE；

接着，Frozen掉DCE模块，用DCE模块替换掉本来MCM-LDM的style remover，从头train网络（数据集是HumanML3D）。

做完之后，即可跑评估标准，与上图进行一个定量的对比（可以写三行，不同的 $\lambda_{adv}$ 值调参的结果（ $\lambda_{adv}=0.2$ ， $\lambda_{adv}=0.8$ ， $\lambda_{adv}=0.5$ ））。

- （有时间，有最好）如果效果真的好很多，当然可以放结果对比图，这里的图是证明我们做风格迁移的效果更好。

## | 2.场景任务，两个模块的消融实验（FiLM和时间加权denoiser去噪模块）

目标（算上图）：2/3页

- 关于DCE模块的消融，看一下时间是否允许

**场景-动作一致性打分器 (Scene-Motion Consistency Scorer) - 定量（新指标）SMCS:**

- **概念:** 训练一个独立的模型（可以是基于Transformer的分类器或回归器），它的任务是接收一个动作序列和一个场景描述（文本嵌入），然后输出一个分数，表示这个动作和这个场景的匹配程度。

定量评估：只跟MCM-LDM比，我们的完整版/消融各个模块的版本，跟MCM-LDM只比SMCS指标。**比较让人信服的是有结果图的对比（有图），消融各个模块。**

## | 3.定性分析

这里也放一张和题图差不多的图，如果时间允许可以多放一些对比试验。

- User Study：选一些（15~20）生成结果的视频，
  - （1）选出哪个最符合某个场景？
  - （2）选出哪个最符合某个Style？
  - （3）内容保真度，轨迹保真度，
  - （4）给出若干场景，选择更贴近于哪个场景？