【项目描述】

CCF大数据与计算智能大赛基于飞桨实现花样滑冰选手骨骼点动作识别。

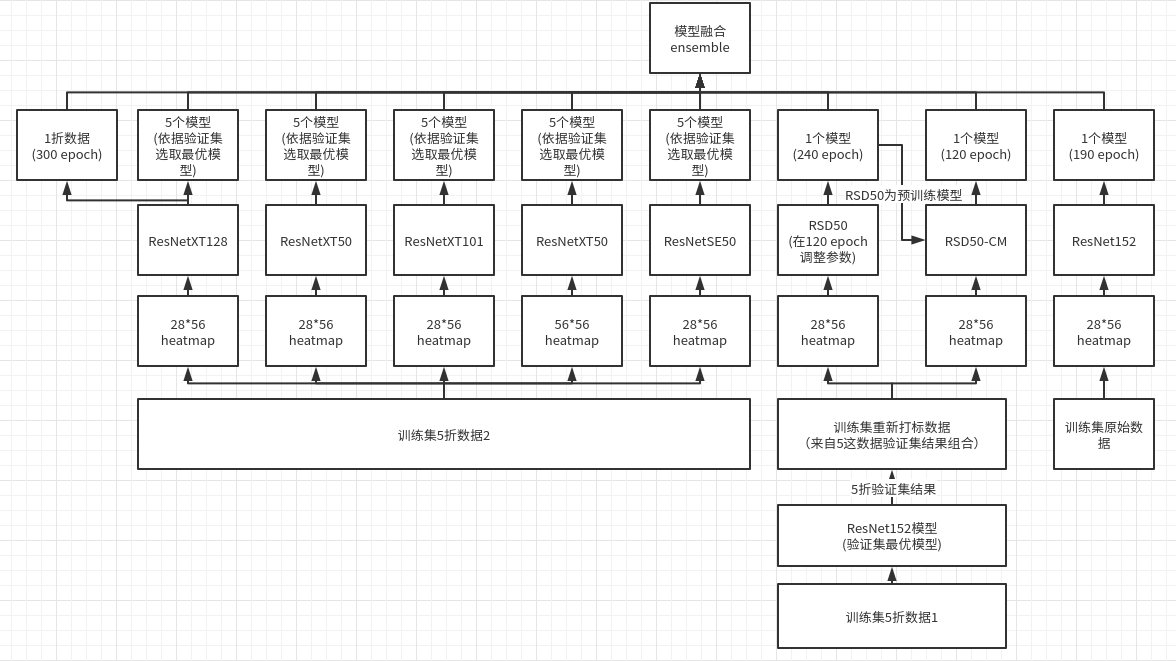
新东方人工智能研究院队伍

【硬件要求】

1. 由于模型基于3D卷积搭建，因此整体资源要求较高。
2. 由于模型较多，为保证快速复现num\_worker尽量不要设置为0，
3. 因为num\_worker不设置为0，训练过程中会占用大量的共享内存，因此共享内存目录（/dev/shm/）尽量分配多一些空间。若num\_worker=0，则占用共享内存较小，但是训练速度降低较多。

【项目整体架构】

本算法主要由三个系列的模型集合组成XtSe系列、RSD系列模型、ResNet152系列模型。每个系列模型的数据与训练方式均有不同，整体算法架构如下图所示，详细说明在下文。



【项目目录结构】

由于数据文件较多，首先需要解压数据，保证目录统一

```

cd data

unzip data118206/c\_data.zip

mv c\_data/\* ./

```

data解压后，项目的整体目录层级如下

| -- data

| -- XtSe

| -- Distill

| -- Res152

| -- test\_B\_data.npy

| -- work

| -- XtSe

| -- \*模型代码文件\*

| -- train.sh

| -- test.sh

| -- fast\_test.sh

| -- Distill

| -- \*模型代码文件\*

| -- train.sh

| -- test.sh

| -- Res152

| -- \*模型代码文件\*

| -- train.sh

| -- test.sh

| -- Kaggle-Ensemble-Guide-master

| -- requirements.txt

| -- 花样滑冰比赛文档.docx

在work/路径下，有三个主要的模型对应的文件夹，其中XtSe/文件夹里面的是XtSe系列的模型，Distill/文件夹里面是RSD系列模型，Res152/文件夹里面是ResNet152系列模型，最终结果基于以上所有的模型的集成得到。

【推理代码】

模型运行前需要安装必要环境

```

pip install -r requirements.txt -i <https://pypi.tuna.tsinghua.edu.cn/simple/>

```

直接利用本次比赛模型生成好的29个csv直接融合，生成B榜结果

```

cd work/Kaggle-Ensemble-Guide-master

tar -xvf testB.tar

python src/kaggle\_vote.py "testB/submissions/\*.csv" "./submission.csv"

```

利用模型推理29个模型csv结果

推理XtSe系列模型结果

```

cd work/XtSe/

sh fast\_test.sh

```

推理RSD系列模型结果

```

cd work/Distill/

sh test.sh

```

推理ResNet152系列模型结果

```

cd work/Res152/

sh test.sh

```

基于投票的方式进行模型融合，将所有模型得到的csv文件放到work/Kaggle-Ensemble-Guide-master/submissions下  
```  
cd work/Kaggle-Ensemble-Guide-master  
python src/kaggle\_vote.py "./submissions/\*.csv" "./submission.csv"  
```

【训练代码】

1. 数据

生成XtSe系列数据（AIStudio环境中已经给出我们分好的五折数据，尽量使用此数据进行复现，如果使用代码重新分组，不能保证数据分布与已给出的五折数据分布一致）。

```

#详见此目录下的README.md

python work/XtSe/tools/split\_crossval.py

```

生成RSD系列模型数据

```

#详见此目录下的README.md

cd work/Distill

python relabel\_5fold.py

```

RSD训练数据中npy文件为利用在**训练集5折数据**上训练的ResNet152模型在**训练集数据**上重新打标得到，可以详见work/RSD/README.md计算npy

Resnet152系列模型数据，就是原始训练数据，在如下位置

```

data/Res152

```

1. 模型训练

训练XtSe系列模型

```

cd work/XtSe/

#训练较为复杂，详见此目录下的README.md

```

训练RSD系列模型

```

cd work/Distill/

sh train.sh

```

训练Resnet152系列模型

```

cd work/Res152/

sh train.sh

```

【模型架构与思路】

模型主要参考PoseC3D模型(参考链接[https://github.com/open-mmlab/mmaction2](https://wx.qq.com/cgi-bin/mmwebwx-bin/webwxcheckurl?requrl=https://github.com/open-mmlab/mmaction2&skey=@crypt_b586cb1c_eaf2c15150eb6f2a4a073fcb0e1b50bb&deviceid=e671115909095331&pass_ticket=fYoV1PdKkoI01oYcUsJV90m3vzG%2BFSf60Q0TDBeGG81Ov8Na%2FO4VF8gVZljxtG0n&opcode=2&scene=1&username=@ab95e0d7352ffa7a69f759830a1869cf2f4398ed96de1407f62b411cc94729b0" \t "/home/tonychen/Documents\\x/_blank)

)。并在这个模型基础上完成了XtSe系列、RSD系列模型、Resnet152系列三个系列模型。我们在模型训练方式与结构上进行了多样的优化，利用模型与数据的分布不同，从不同角度学习了不同类别样本的分布，最大限度地减少了数据过拟合造成的影响。

1. 训练方式优化（RSD系列）
2. 自蒸馏（RSD50模型）：

在分析数据的过程中，我们发现一些类别很难区分，且样本量较少，针对这种问题，我们借鉴了自蒸馏的方法（参考来源：https://www.kaggle.com/c/plant-pathology-2020-fgvc7/discussion/154056）来解决该问题：我们首先使用5折数据训练了一个resnet152模型，之后将该模型在五折的验证集上预测的结果作为第二标签，在训练时，和真实标签按照7：3的比例进行混合。

1. 代价敏感矩阵（RSD50-CM模型）：

针对数据中存在的数据不均衡问题，同时为了改善模型训练中存在的易混淆类别，我们使用了基于混淆矩阵的代价敏感矩阵作为损失函数的权重。参考方法来自魏秀参的《解析卷积神经网络》，在实际使用时由于混淆矩阵数值过大，造成loss过大，因此使用sigmiod函数对其进行了处理。RSD50-CM模型在训练前利用RSD50模型240epoch训练权重作为预训练模型。

1. 模型结构优化（XtSe系列）

我们基于PoseC3D模型的进行了改进，主要针对主干网络进行了修改。我们将原始的ResNet主干网络分别优化成为基于3D卷积的ResNetSE、ResNetXT（修改方式与基于普通2D卷积的ResNetSE、ResNetXT修改方式一致）。

1. 利用模型与数据分布不同进行模型融合

竞赛的最大难点是数据有限，数据复杂，单一模型训练大概率会导致过拟合现象的发生。因此我们利用数据差异性、模型差异性、训练差异性习得数据样本的不同空间分布，对抗过拟合产生的影响。

（1）利用原始整体训练数据，训练ResNet152（Resnet152系列）

这里直接利用原始整体数据集，训练以ResNet152为主干网络的模型。不过多赘述。

1. 利用重新打标的训练数据，训练自蒸馏ResNet50（RSD系列，图中RSD50与RSD50-CM）

主要参考前文”训练方式优化”

1. 利用5折分折数据，训练基于ResNetSE、基于ResNetXT的主干网络（XtSe系列）

为了更好地验证模型训练过程中的准确性，我们将数据集均匀地分为5折（保证每一折中各个类别分布均衡）。本模型前处理是将人体坐标点转化为heatmap形式，为了保证模型能学到类别的不同分布，我们生成了两种尺度的heatmap，分别是28×56与56×56。根据以上论述，我们训练了ResNetXT128（输入heatmap尺寸28×56）、ResNetXT50（输入heatmap尺寸28×56）、ResNetXT101（输入heatmap尺寸28×56）、ResNetXT50（输入heatmap尺寸56×56）、ResNetSE50（输入heatmap尺寸28×56），5个模型，结合5折数据，一共25个模型。在这些模型中，选取每一折数据在验证集上表现最好的，作为推理模型。

另外根据数据分析，我们认为第一折数据噪声较少，样本分布性较宽，可以适当增加第一折数据模型结果在最终模型融合算法中的权重。因此我们另外选取ResNetXT128在第一折数据上最后一轮的结果，参与模型融合，以增加最终准确率。

1. 模型融合

基于以上模型，一共得到29个模型结果，在这29个模型基础上，做模型融合，得到最终结果。