

基于联邦学习的医学影像分类研究



合肥工业大学

苏伊阳

2022.4.29

目录

01

论文分享

02

个人实验设计

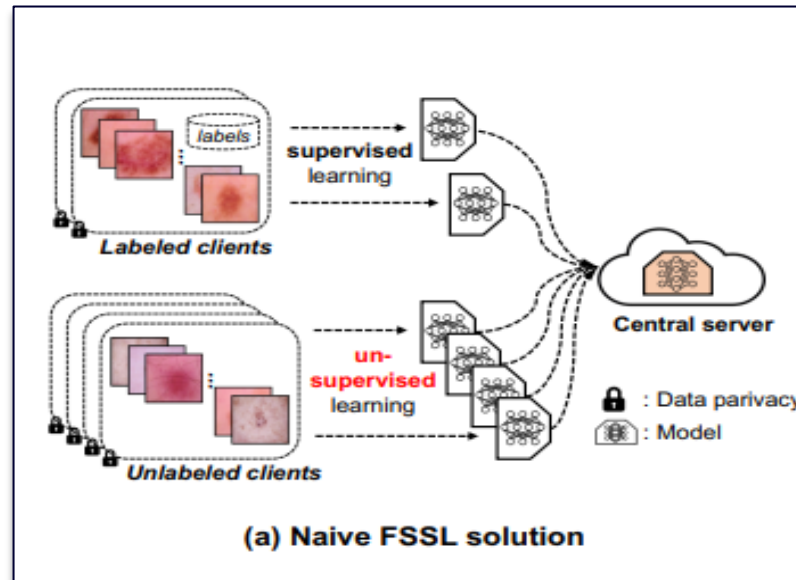
论文分享

Federated Semi-supervised Medical Image Classification via Inter-client Relation Matching

基于客户关系匹配的联邦半监督医学图像分类

研究场景：只有几个客户端有标签数据，而大多数客户端都是无标签数据。

半监督学习：首先在标签数据上训练模型，然后使用经过训练的模型来预测无标签数据的标签，从而创建伪标签。然后，将标签数据和新生成的伪标签数据结合起来作为新的训练数据。



在联邦学习背景下，标签数据
与无标签数据无法直接交互

论文分享

FSSL框架

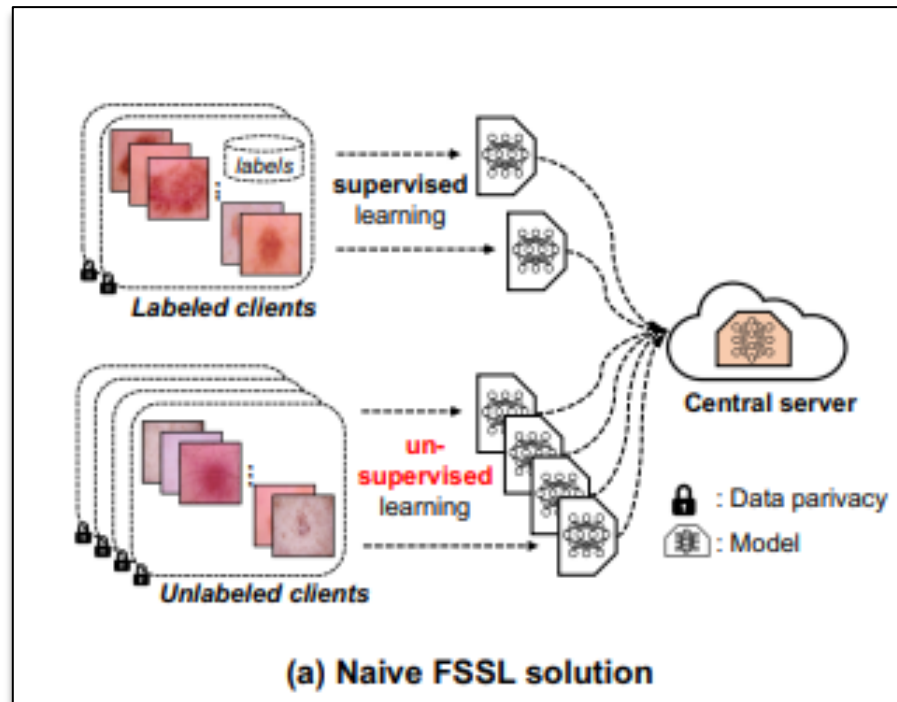
$$\mathcal{D}_L = \{\mathcal{D}^1, \mathcal{D}^2, \dots, \mathcal{D}^m\}, \quad \mathcal{D}^l = \{(x_i^l, y_i^l)\}_{i=1}^{N^l}$$

$$\mathcal{D}_U = \{\mathcal{D}^{m+1}, \mathcal{D}^{m+2}, \dots, \mathcal{D}^{m+n}\}, \quad \mathcal{D}^u = \{(x_i^u)\}_{i=1}^{N^u}$$

labeled clients: cross entropy loss

unlabeled clients: $\mathcal{L}_c(\mathcal{D}^u, \theta^u) = \sum_{i=1}^{N^u} \mathbb{E}_{\xi, \xi'} \|f_{\theta^u}(x_i^u, \xi), f_{\theta^u}(x_i^u, \xi')\|_2^2$

优化算法: FedAvg



横向联邦学习

FedAvg

Algorithm 1 FederatedAveraging. The K clients are indexed by k ; B is the local minibatch size, E is the number of local epochs, and η is the learning rate.

Server executes:

```
initialize  $w_0$ 
for each round  $t = 1, 2, \dots$  do
   $m \leftarrow \max(C \cdot K, 1)$ 
   $S_t \leftarrow$  (random set of  $m$  clients)
  for each client  $k \in S_t$  in parallel do
     $w_{t+1}^k \leftarrow \text{ClientUpdate}(k, w_t)$ 
   $w_{t+1} \leftarrow \sum_{k=1}^K \frac{n_k}{n} w_{t+1}^k$ 
```

```
ClientUpdate( $k, w$ ): // Run on client  $k$ 
 $\mathcal{B} \leftarrow$  (split  $\mathcal{P}_k$  into batches of size  $B$ )
for each local epoch  $i$  from 1 to  $E$  do
  for batch  $b \in \mathcal{B}$  do
     $w \leftarrow w - \eta \nabla \ell(w; b)$ 
return  $w$  to server
```

C是随机分数，用来随机挑选客户端的数量

K是总共的客户端数量

$$\min_{w \in \mathbb{R}^d} f(w) \quad f(w) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n f_i(w)$$

$$f_i(w) = l(x_i, y_i; w)$$

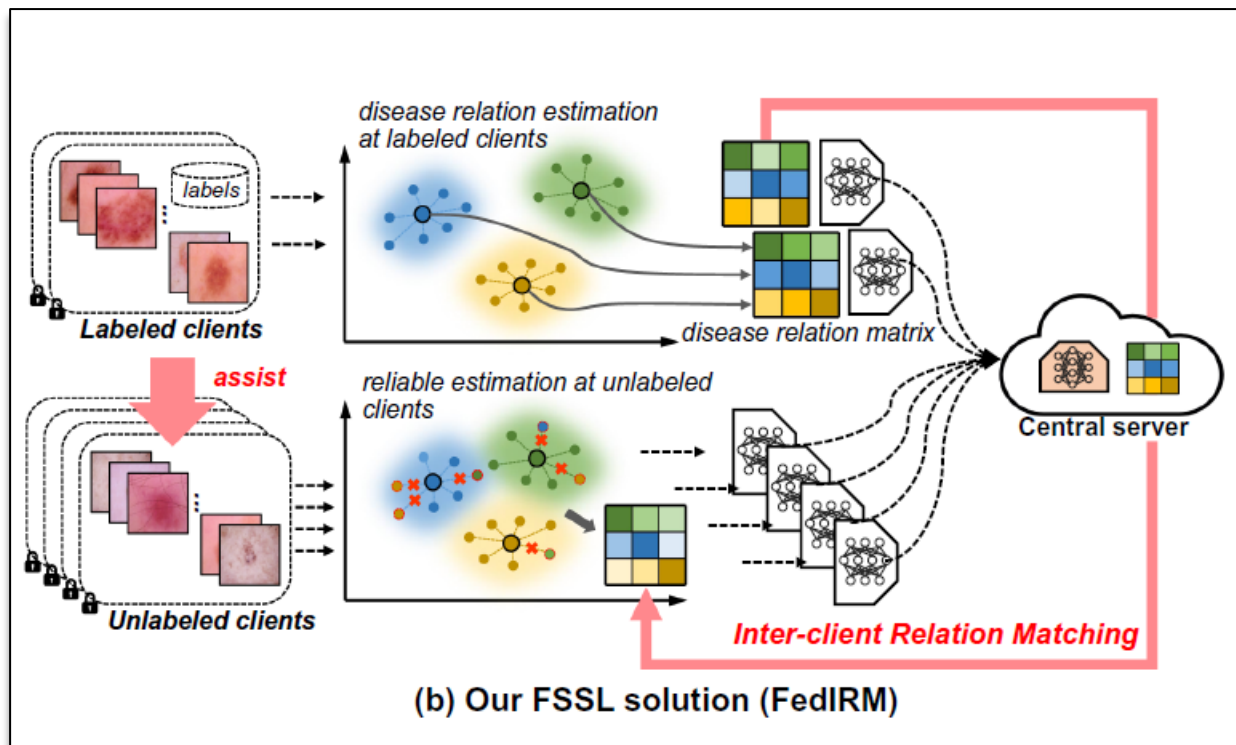
$$f(w) = \sum_{k=1}^K \frac{n_k}{n} F_k(w) \quad F_k(w) = \frac{1}{n_k} \sum_{i \in \mathcal{P}_k} f_i(w)$$

$E_{\mathcal{P}_k}(F_k(w)) = f(w)$ 成立时，是独立同分布的数据

论文分享

主要思想:

- ◆ 疾病类别之间存在一定的相关关系，反映了医学图像分类背景下的**结构性任务知识**
- ◆ 从有标签的数据中提取不变的疾病关系信息，来监督无标签数据的学习，从而减少在无标签数据学习的过程中产生任务知识丢失，更加有效地利用无标签数据样本。
- ◆ 半监督学习方法：一致性正则化机制



论文分享

标签客户端

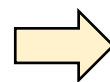
对于每个标签客户端 D^l ，计算每个类别的平均特征向量 \mathbf{v}_c^l

$$\mathbf{v}_c^l = \frac{1}{N_c^l} \sum_{i=1}^{N_c^l} \boxed{1[y_i^l = c]} \hat{f}_{\theta^l}(x_i^l)$$

N_c^l 表示有标签客户端 l 中第 c 类样本的数量， \hat{f} 表示没有最后softmax层的模型

$$1[y_i^l = c] \iff 1[y_i^l = c] = \begin{cases} 1 & y_i^l = c \\ 0 & y_i^l \neq c \end{cases}$$

$$\mathbf{s}_c^l = \text{softmax}(\mathbf{v}_c^l / \tau), \quad \tau > 1$$



$$\mathcal{M}^l = [\mathbf{s}_1^l, \dots, \mathbf{s}_c^l]$$



标签客户端疾病关系矩阵-表示不同疾病类别间的类间关系

传达了特定类别样本的网络预测通常在所有类别上是如何分布

论文分享

无标签客户端

对mini batch x^u (B个样本) 随机dropout T次得到预测概率向量集合 $\{q_t^u\}_{t=1}^T$

$$\text{不确定性 } w^u = -\sum_{c=1}^C \bar{q}_{(c)}^u \log(\bar{q}_{(c)}^u), \text{ with } \bar{q}_{(c)}^u = \frac{1}{T} \sum_{t=1}^T q_{t(c)}^u$$

$q_{t(c)}^u$ 表示 q_t^u 中第 c 类的值

过滤不确定性大于某个值的标签后, 为保留的数据计算伪标签:

$$y^u = \operatorname{argmax}(p^u) \quad (p^u \text{ 经过softmax层})$$

$$v_c^u = \frac{\sum_{i=1}^B 1_{[(y_i=c) \cdot (w_i^u < h)]} \cdot p_i^u}{\sum_{i=1}^B 1_{[(y_i=c) \cdot (w_i^u < h)]}} \quad (p_i^u \text{ 没有经过softmax层, 已经过滤的})$$

$$s_c^u = \operatorname{softmax}(v_c^u / \tau), \quad \tau > 1$$

$$\mathcal{M}^u = [s_1^u, s_2^u, \dots, s_c^u]$$



无标签客户端疾病关系矩阵-表示不同疾病类别间的类间关系

论文分享

客户关系匹配

目标：使无标签客户端产生的疾病关系矩阵与标签客户端产生的疾病关系矩阵相似

每一轮中心服务器收集标签客户端的 \mathcal{M}^l

计算 $\mathcal{M} = \frac{1}{m} \sum_{l=1}^m \mathcal{M}^l$ ，将 \mathcal{M} 传给无标签客户端

有标签客户端先训练20轮，然后开始训练无标签客户端

客户关系匹配损失函数 $\mathcal{L}_{\text{IRM}} = \frac{1}{C} \sum_{c=1}^C (\mathcal{L}_{\text{KL}}(\mathcal{M}_c \parallel \mathcal{M}_c^u) + \mathcal{L}_{\text{KL}}(\mathcal{M}_c^u \parallel \mathcal{M}_c))$

$$\mathcal{L}_{\text{KL}}(\mathcal{M}_c \parallel \mathcal{M}_c^u) = \sum_j \mathcal{M}_{c(j)} \log \frac{\mathcal{M}_{c(j)}}{\mathcal{M}_{c(j)}^u} \quad \mathcal{M}_c = s_c$$

$$\mathcal{L}^l = \mathcal{L}_{ce}(\mathcal{D}^l, \theta^l) \text{ and } \mathcal{L}^u = \lambda(\omega)(\mathcal{L}_c + \mathcal{L}_{\text{IRM}})$$

$$\lambda(\omega) = 1 * e^{(-5(1-\omega/\Omega))}$$

$\lambda(\omega)$ 有助于减少在早期一轮一轮训练欠拟合时对无标签客户端的学习影响

论文分享

实验结果

数据集：颅内出血的脑部CT，皮肤病变图像

颅内出血的脑部CT：5个亚型 训练集、验证集和测试集7:1:2

皮肤病变图像：7种类型 训练集、验证集和测试集7:1:2

评估指标：AUC、Sensitivity、Specificity、Accuracy和F1 score

模型：DenseNet121

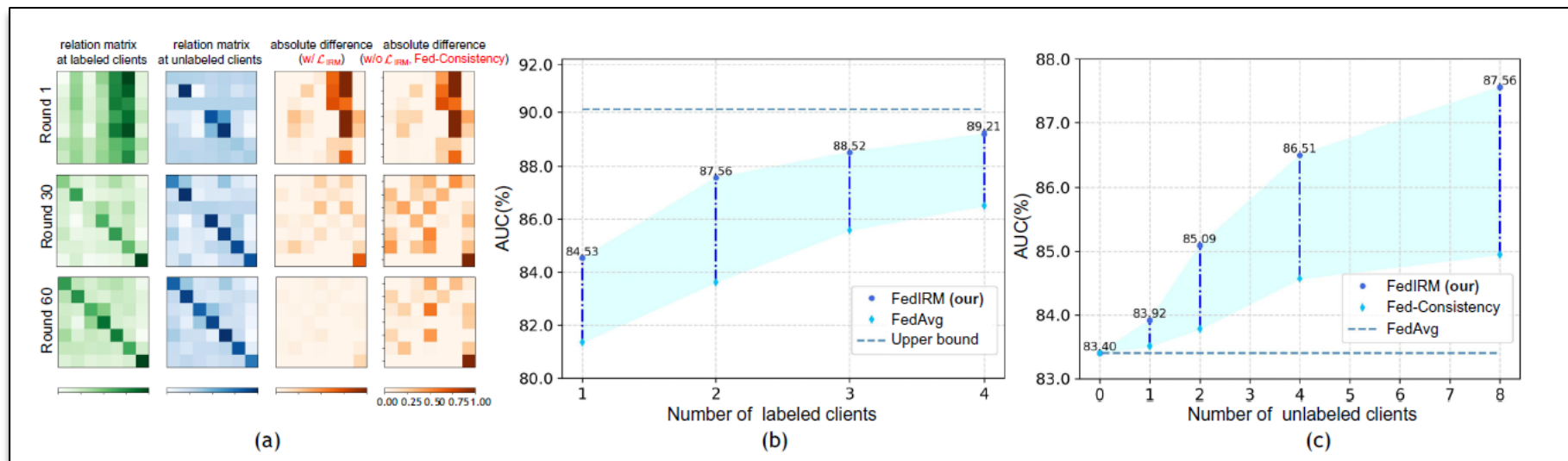
扰动：对输入数据的随机变换(旋转、平移、翻转)和dropout层

Table 1. Quantitative comparisons with state-of-the-arts on two different tasks.

Method	Client num		Metrics				
	Label	Unlabel	AUC	Sensitivity	Specificity	Accuracy	F1
Task 1: Intracranial hemorrhage diagnosis							
FedAvg [20]	10	0	90.48 \pm 0.31	64.33 \pm 1.13	92.68 \pm 0.43	89.94 \pm 0.92	63.94 \pm 1.20
FedAvg [20]	2	0	83.40 \pm 0.87	57.88 \pm 1.68	90.48 \pm 0.79	87.45 \pm 1.08	57.10 \pm 1.29
Fed-SelfTraining [33]	2	8	84.32 \pm 0.82	57.94 \pm 1.66	90.22 \pm 0.74	87.90 \pm 1.81	57.48 \pm 1.14
Fed-Consistency [31]	2	8	84.83 \pm 0.79	57.26 \pm 1.93	90.87 \pm 0.62	88.35 \pm 1.32	57.61 \pm 1.08
FedIRM (ours)	2	8	87.56 \pm 0.56	59.57 \pm 1.57	91.53 \pm 0.81	88.89 \pm 1.29	59.86 \pm 1.65
Task 2: Skin Lesion Diagnosis							
FedAvg [20]	10	0	94.82 \pm 0.32	75.11 \pm 1.82	94.87 \pm 0.35	95.24 \pm 0.21	70.16 \pm 1.21
FedAvg [20]	2	0	90.65 \pm 1.23	65.53 \pm 1.76	91.76 \pm 0.48	92.53 \pm 0.67	52.59 \pm 1.42
Fed-SelfTraining [33]	2	8	90.82 \pm 0.56	67.03 \pm 1.93	93.61 \pm 0.21	92.47 \pm 0.34	53.44 \pm 1.85
Fed-Consistency [31]	2	8	91.13 \pm 0.62	68.55 \pm 1.29	93.45 \pm 0.94	92.67 \pm 0.39	54.25 \pm 1.31
FedIRM (ours)	2	8	92.46 \pm 0.45	69.05 \pm 1.71	93.29 \pm 0.59	92.89 \pm 0.25	55.81 \pm 1.49

论文分享

实验结果



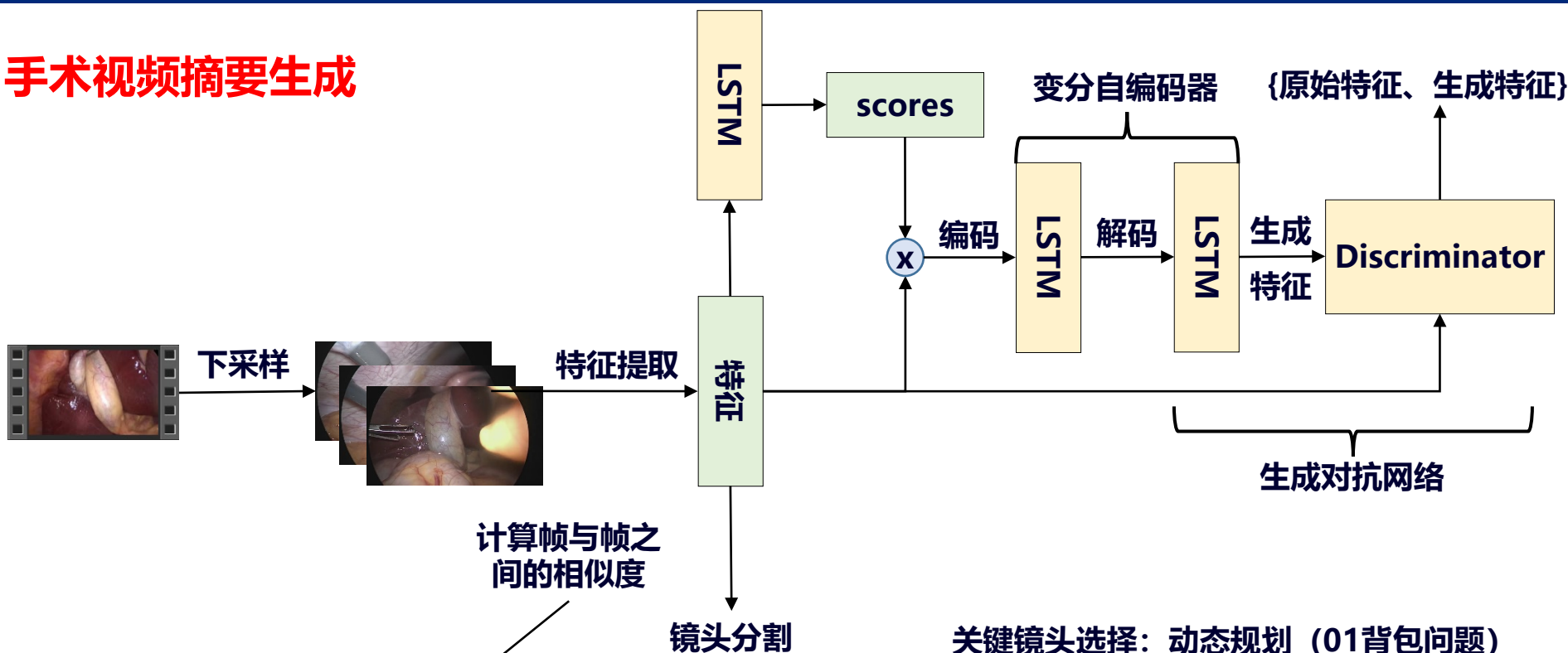
(a) 标签客户端、无标签客户端的疾病关系矩阵以及没有客户关系匹配损失的可视化

(b) FedIRM和FedAvg在不同标签客户端数量下的性能

(c) 无标签客户端数量对模型性能的影响

个人实验设计

手术视频摘要生成



每十秒分一个片段（或者滑动窗口）【片段的第一帧和最后一帧计算相似度，大于这个相似度的百分之多少设为阈值，从里面选一个最大的】【全部视频设置一个阈值，大于这个数设为镜头边界】，计算下采样后帧与帧之间的不相似度，从一个片段里选出一个不相似度最大的，选出这个片段里大于阈值的不相似度，从中选一个最大的，作为镜头边界。

个人实验设计

关键镜头选择：动态规划（01背包问题）

selected = knapSack(摘要的长度, 每个镜头的长度, 每个镜头的scores, 镜头的个数)

评价生成摘要的好坏在于摘要质量,最好是对视频内容的代表,同时多样化,最好能覆盖整个内容。

评价指标：摘要视频里的三元组类型数量/整个视频的三元组类型数量

Q&A



合肥工業大學