2019 TJMSC Tech. Courses

Re-ID

Bingchen Zhao
zhaobc@tongji.edu.cn
Tongji Microsoft Student Club

May, 28th
Room 516, Ji Shi Building
Tongji Univ

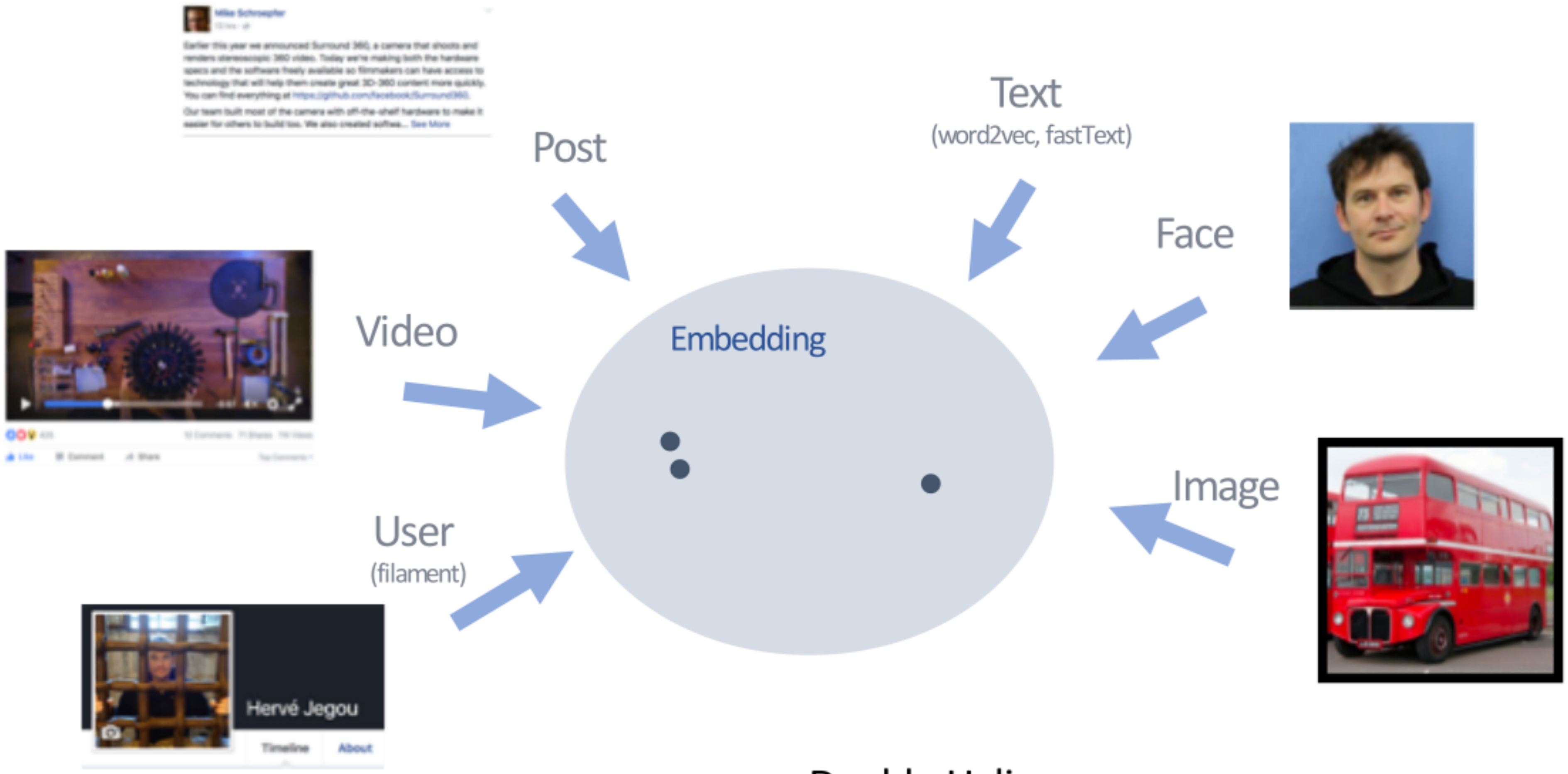
Facial Recognition

同济微软俱乐部
Tongji Microsoft Student Club

CelebA is a large scale face dataset contains of famous people's faces

Facial Recognition

Embeddings



Double Helix
DeepText

Embeddings

同济微软俱乐部
Tongji Microsoft Student Club

麻辣烫 ——

- 麻辣拌, 0.884719466741
- 冒菜, 0.878425403476
- 麻辣汤, 0.872023557375
- 杨国福, 0.863779634807
- 云南米线, 0.855127326166
- 张亮, 0.833823338064
- 过桥米线, 0.833199404157
- 云南过桥米线, 0.825620560846
- 羊肉米线, 0.824504212524
- 重庆小面, 0.823935462376

肯德基 ——

- KFC, 0.982491910803
- kfc, 0.970919196284
- 麦当劳, 0.94600897271
- 德克士, 0.945771247431
- 华莱士, 0.94466343164
- 汉堡王, 0.938037115861
- 正新鸡排, 0.920964315301
- 派乐, 0.91322924147
- 豪大大, 0.903992194302
- 正新, 0.903950289626

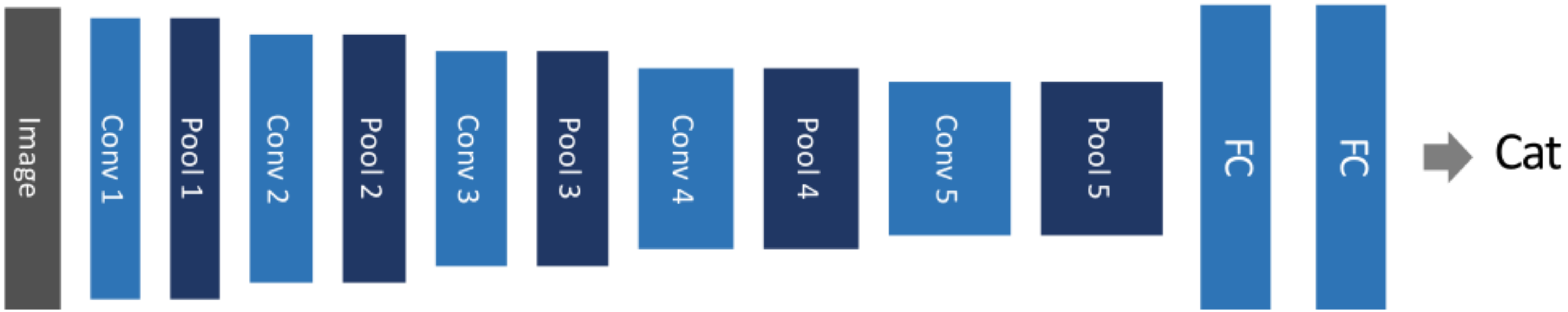
黄焖鸡 ——

- 黄焖鸡米饭, 0.952911757759
- 鸡公煲, 0.91343873243
- 鸡米饭, 0.90181777612
- 大盘鸡, 0.876224767056
- 黄闷鸡, 0.875185069057
- 麻辣香锅, 0.869952513231
- 香锅, 0.857016354508
- 鸡汤泡饭, 0.821301219974
- 水煮肉, 0.81307743024
- 炒菜, 0.808240640778

馄饨 ————

- 混沌, 0.971946557298
- 馄炖, 0.959630998451
- 饺子, 0.942945794257
- 水饺, 0.926106932097
- 锅贴, 0.923119816702
- 小混沌, 0.914481422287
- 生煎, 0.900458398145
- 汤包, 0.9000372527
- 小馄饨, 0.897959492742
- 馄钝, 0.896437234474

Image Embeddings



FaceNet and triplet loss

FaceNet: A Unified Embedding for Face Recognition and Clustering

Florian Schroff
fschroff@google.com
Google Inc.

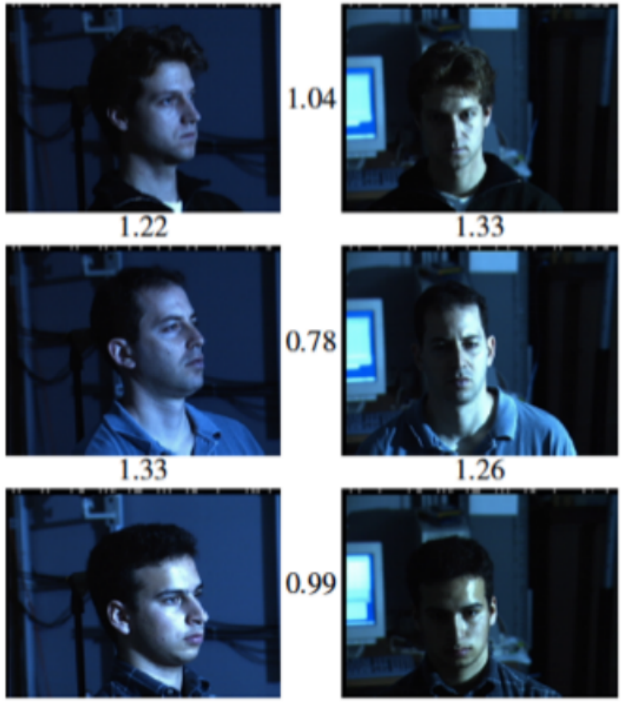
Dmitry Kalenichenko
dkalenichenko@google.com
Google Inc.

James Philbin
jphilbin@google.com
Google Inc.

Abstract

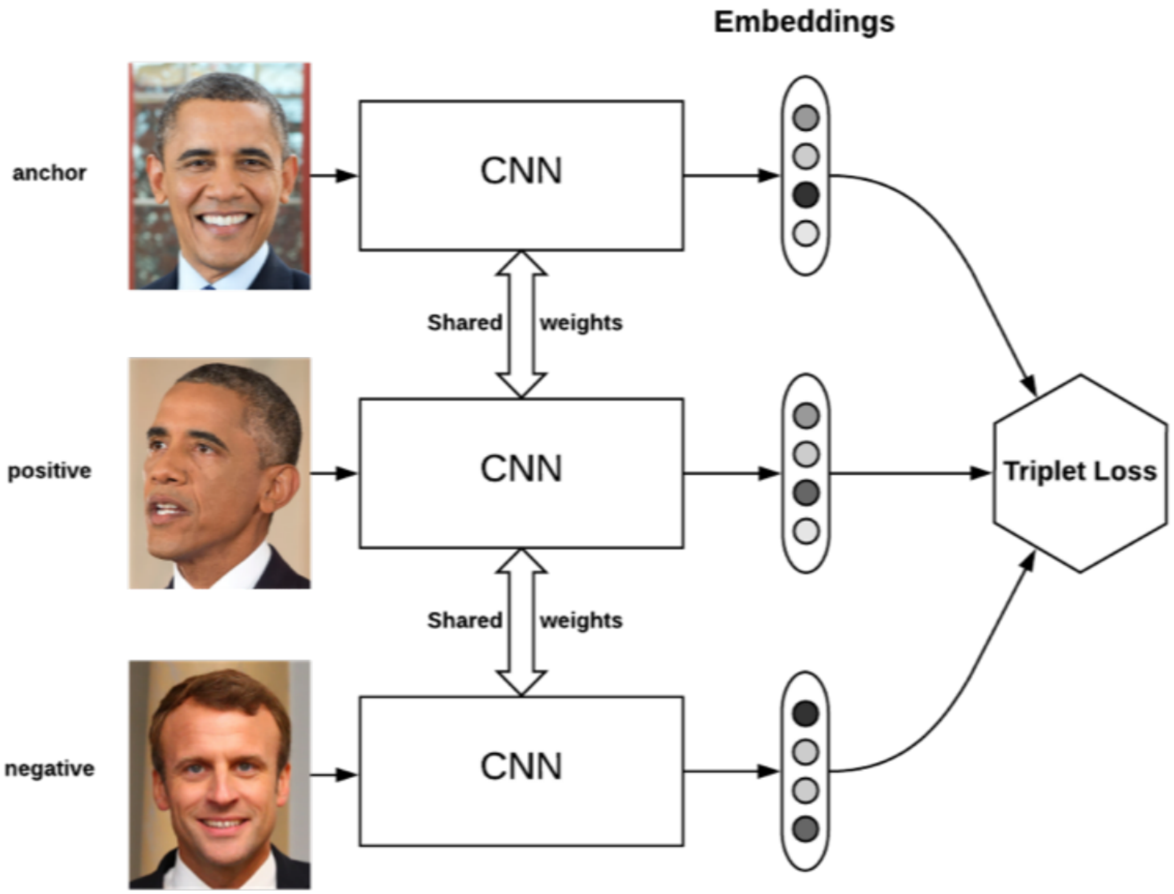
Despite significant recent advances in the field of face recognition [10, 14, 15, 17], implementing face verification and recognition efficiently at scale presents serious challenges to current approaches. In this paper we present a system, called FaceNet, that directly learns a mapping from face images to a compact Euclidean space where distances directly correspond to a measure of face similarity. Once this space has been produced, tasks such as face recognition, verification and clustering can be easily implemented using standard techniques with FaceNet embeddings as feature vectors.

Our method uses a deep convolutional network trained to directly optimize the embedding itself, rather than an intermediate bottleneck layer as in previous deep learning approaches. To train, we use triplets of roughly aligned matching / non-matching face patches generated using a novel online triplet mining method. The benefit of our approach is much greater representational efficiency: we achieve state-of-the-art face recognition performance using



FaceNet and triplet loss

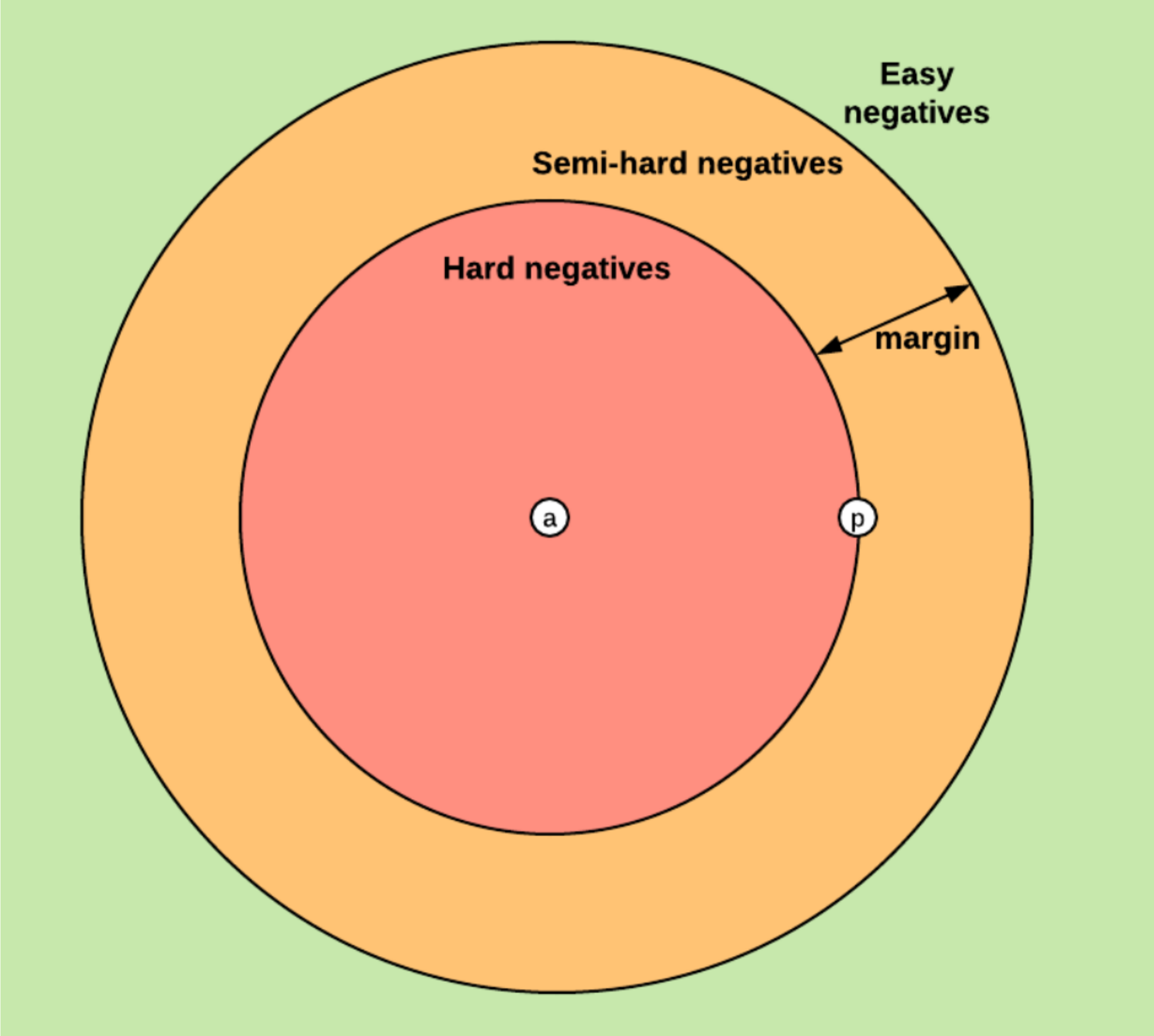
L=max(d(a,p)-d(a,n)+margin,0)



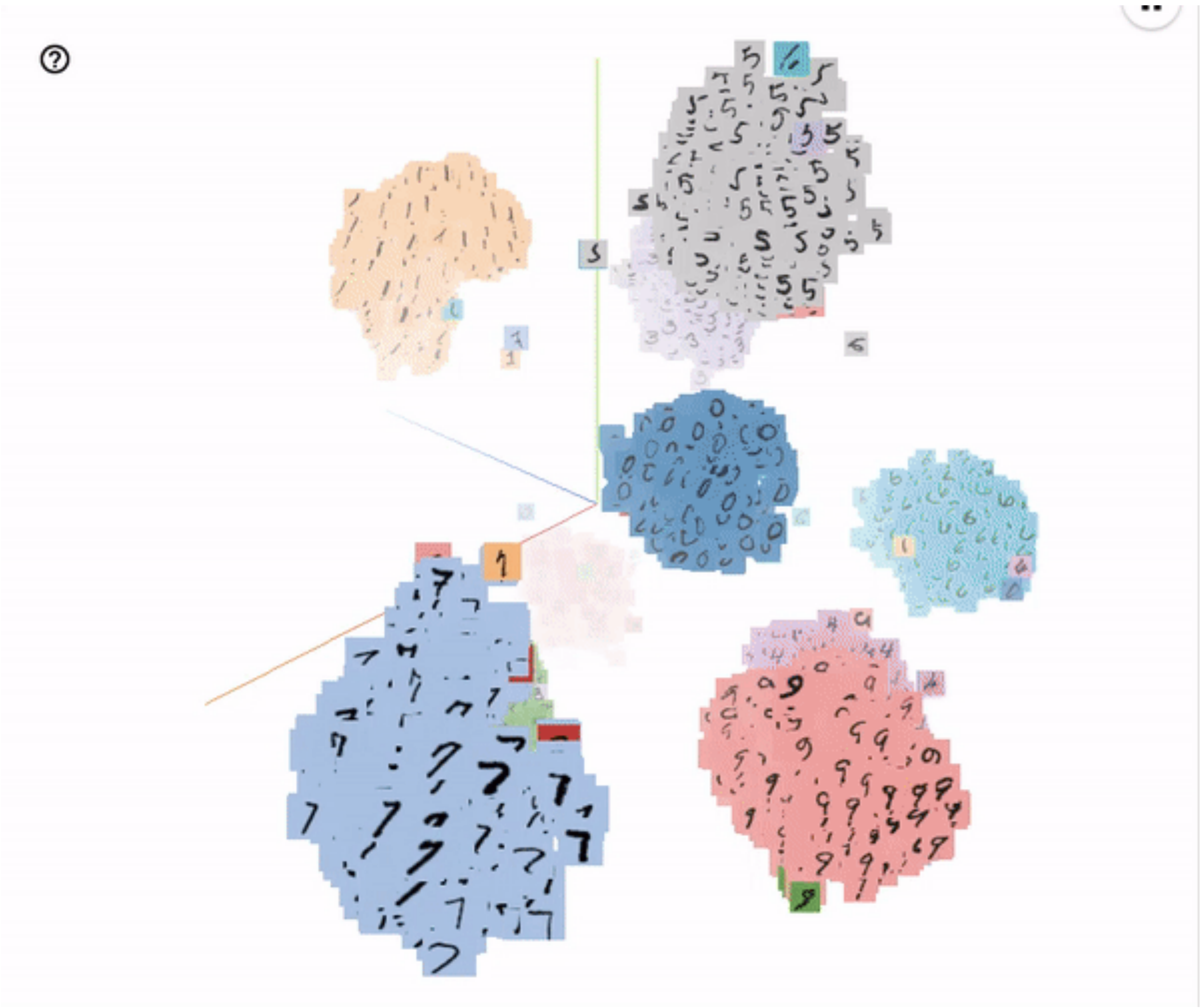
FaceNet and triplet loss

- $L=\max(d(a,p)-d(a,n)+margin,0)$
- **easy triplets**: triplets which have a loss of 0, because $d(a,p)+margin<d(a,n)$
- **hard triplets**: triplets where the negative is closer to the anchor than the positive, i.e. $d(a,n)<d(a,p)$
- **semi-hard triplets**: triplets where the negative is not closer to the anchor than the positive, but which still have positive loss: $d(a,p)<d(a,n)<d(a,p)+margin$

FaceNet and triplet loss



FaceNet and triplet loss



Batch Hard for triplet mining

Sample P identities and for each id sample K images to construct a batch.
Calculate the hardest positive and negative within a batch.

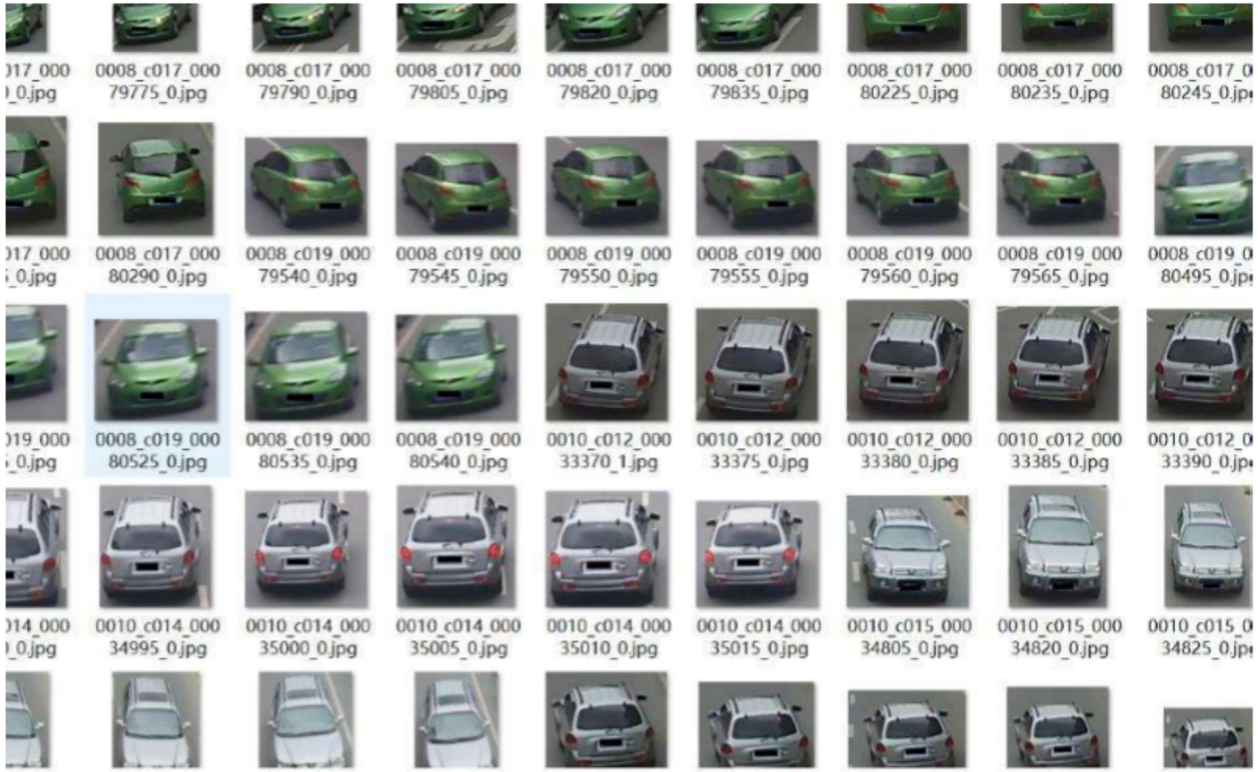
$$\mathcal{L}_{\mathrm{BH}}(\theta; X) = \overbrace{\sum_{i=1}^{P} \sum_{a=1}^{K}}^{\text{all anchors}} \Big[m + \overbrace{\max_{p=1\dots K} D\left(f_\theta(x_a^i), f_\theta(x_p^i)\right)}^{\text{hardest positive}} \quad (5)$$

$$- \underbrace{\min_{\substack{j=1\dots P \\ n=1\dots K \\ j \neq i}} D\left(f_\theta(x_a^i), f_\theta(x_n^j)\right)}_{\text{hardest negative}} \Big]_+ ,$$

The VeRi776 dataset

Dataset statistics:

subset	# ids	# images	# cameras
train	576	37778	20
query	200	1678	19
gallery	200	11579	19

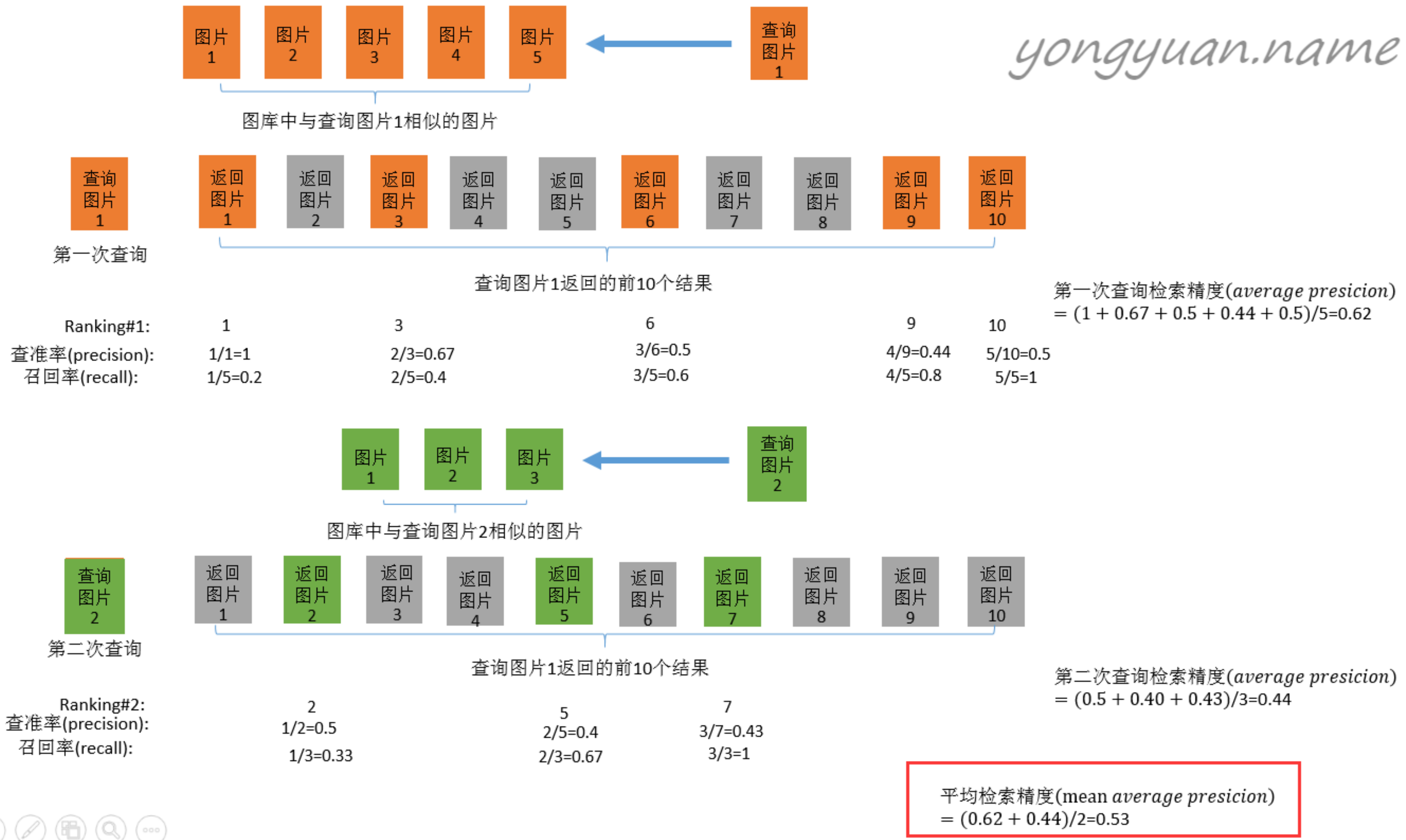


同济微软俱乐部
Tongji Microsoft Student Club

The VeRi776 dataset

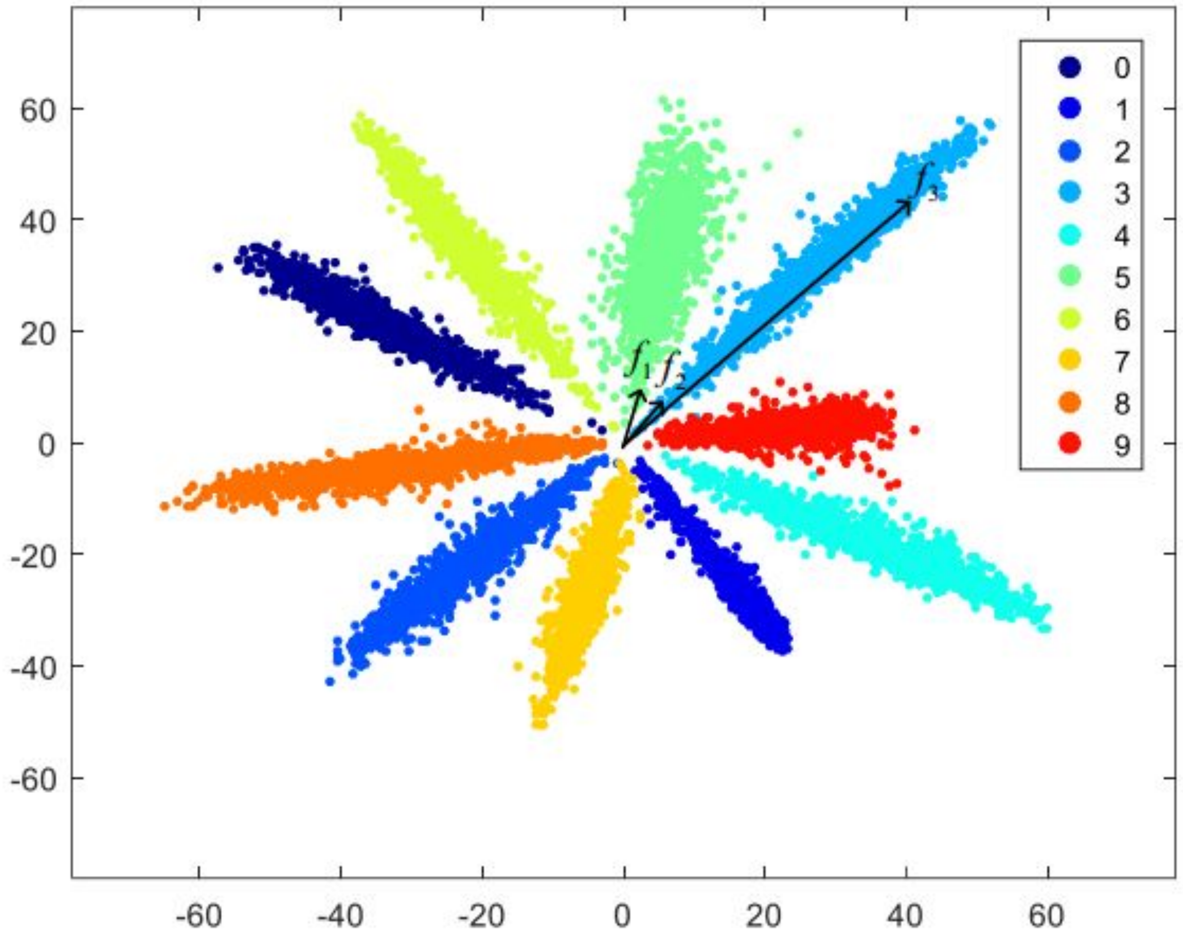
Given a query image, find corresponding image in the gallery

Mean average precision

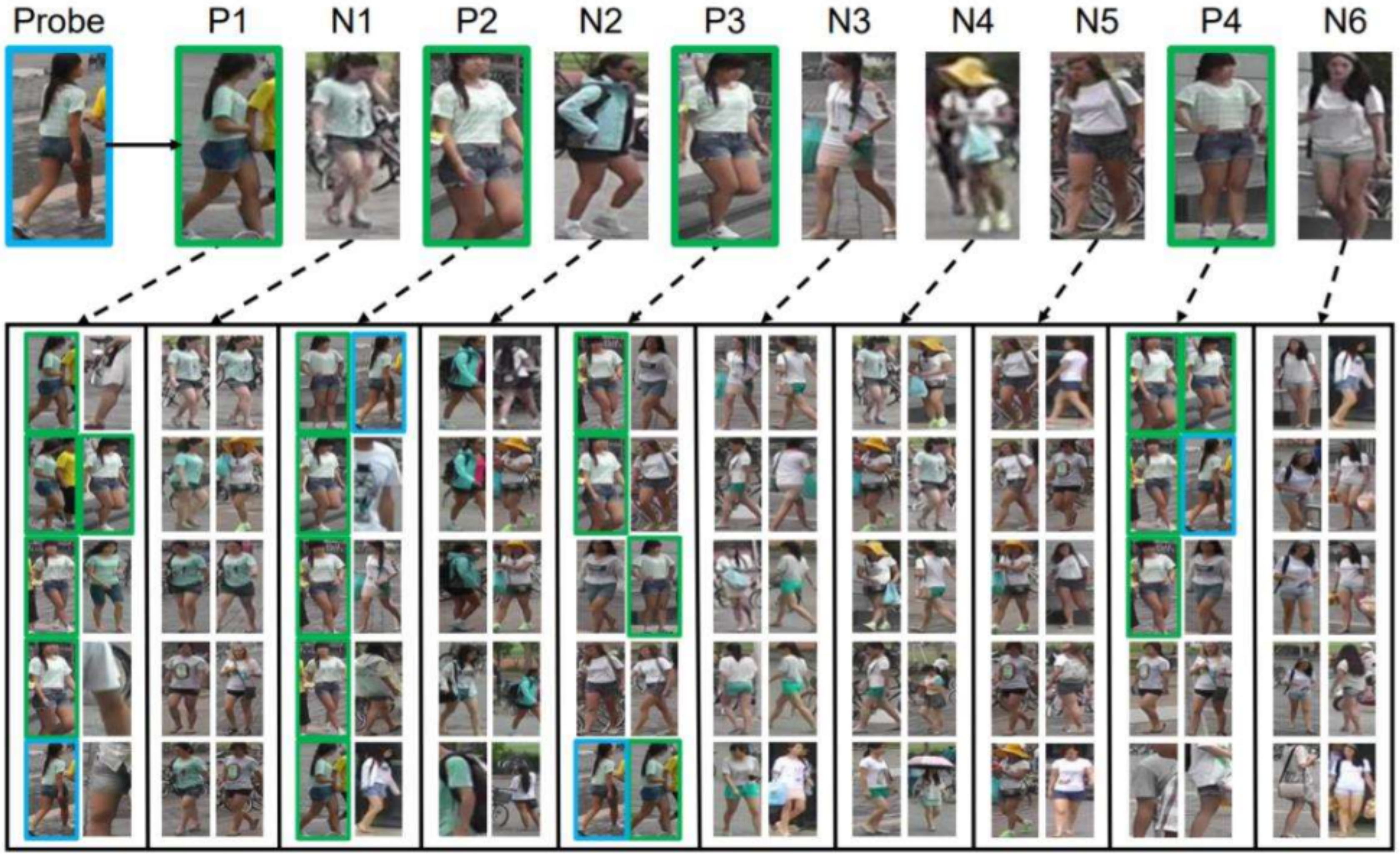


Joint training of triplet and softmax

同济微软俱乐部
Tongji Microsoft Student Club



Re-Ranking



Re-Ranking

if a gallery image is similar to the probe in the k-reciprocal nearest neighbors, it
is more likely to be a true match.

$$\mathcal{R}(p,k) = \{g_i \mid (g_i \in N(p,k) \wedge (p \in N(g_i,k)))\}$$

$$\mathcal{R}^*(p,k) \leftarrow \mathcal{R}(p,k) \cup \mathcal{R}(q, \frac{1}{2}k)$$

$$s.t. \left|\mathcal{R}(p,k) \cap \mathcal{R}(q, \frac{1}{2}k)\right| \geq \frac{2}{3}\left|\mathcal{R}(q, \frac{1}{2}k)\right|,$$

$$\forall q \in \mathcal{R}(p,k)$$

$$d_j(p,g_i) = 1 - \frac{|\mathcal{R}^*(p,k) \cap \mathcal{R}^*(g_i,k)|}{|\mathcal{R}^*(p,k) \cup \mathcal{R}^*(g_i,k)|}$$

$$D(p,g_i) = d(p,g_i) + d_j(p,g_j)$$

Microsoft

同 济 微 软 俱 乐 部
Tongji Microsoft Student Club

加入TJMSC QQ群

关注TJMSC微信公众号