



北京邮电大学

---

# 科研路上的 心得与体会

分享人：王玫

时间：2021年12月15日



**科研经历**



**科研方向**



**心得体会**

## 从进入实验室到一篇算法论文的诞生：





**科研经历**



**科研方向**



**心得体会**

# 人脸识别的种族偏见

## Nature 封面 2020/11/19: 对抗生物特征识别的偏见



### BEATING BIOMETRIC BIAS

Facial recognition is improving — but the bigger issue is how it's used. By Davide Castelvecchi

When London's Metropolitan Police tested real-time facial-recognition technology between 2016 and 2018, they invited Sarah Martin to monitor some of the data from a camera feed mounted inside a police van. "It's like you see in the movies," says Martin, a legal scholar at the University of Essex in Colchester, UK. As cameras scanned passers-by in shopping centres or public squares, they fed images to a computer inside the van. Martin and police officers saw the software draw rectangles around faces as it identified them in the live feed. It then extracted key features and compared them to a match, it pulls an image from the live feed, together with the image from the watch list. Officers then reviewed the match and decided whether to reach out to stop the 'suspect' and, occasionally, arrest them.

Fears and Martin listed a number of ethics and privacy concerns with the design, and questioned whether it was legal at all. And they questioned the accuracy of the system, which is sold by Tokyo-based technology giant NEC. The software flagged 42 people over a trial, but the researchers and police officers dismissed 16 as false 'non-crime' but reached out to stop the others. They lost a people in the crowd, but still stopped 22, only 10 turned out to be correct matches.

The police saw the issue differently. They said the system's number of false positives was tiny, considering the many thousands of faces that had been scanned. (They didn't reply to Nature's requests for comment for this article.)

The accuracy of facial recognition has improved drastically since 'deep learning' techniques were introduced into the field about a decade ago. But whether that means it's good enough to be used on lower-quality, in-the-wild images is a highly controversial issue. And questions remain about how to transparently evaluate facial-recognition systems.

In 2018, a seminal paper by computer scientists Timnit Gebru, then at Microsoft Research in New York City and now at Google in Mountain View, California, and Jue Han, then at the Massachusetts Institute of Technology in Cambridge, found that leading facial-recognition software packages performed much worse at identifying the gender of women and people of colour than a classifying male, white faces.

# MIT Technology Review



### Intelligent Machines

## Are Face Recognition Systems Accurate? Depends on Your Race.



99%



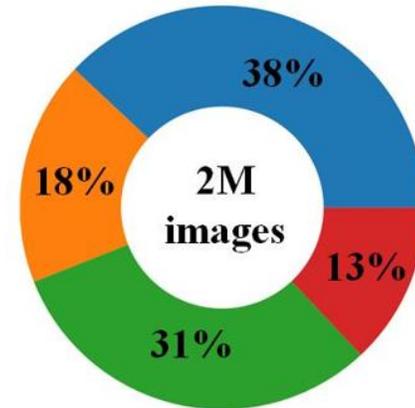
70%



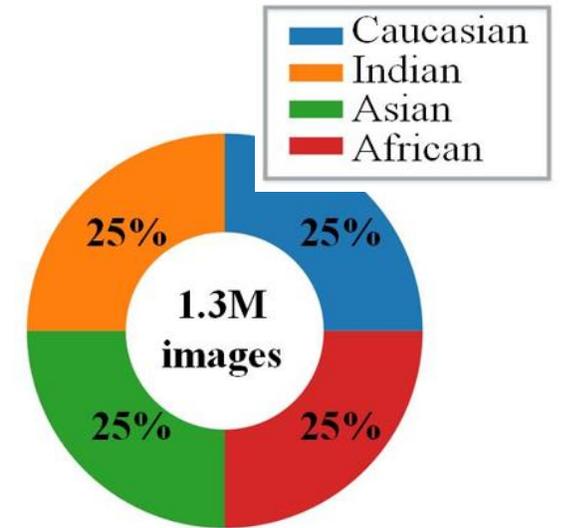
人脸识别在开放环境下的准确率仍不足，且安全性堪忧，对有色人种存在严重偏见

## 成果一：种族偏见数据集

创新点①：对海量图像进行**种族标注**，构建**种族平衡**训练集和测试集，弥补了人脸识别中多种族数据集的缺乏。



(b) BUPT-Globalface



(c) BUPT-Balancedface

所公开数据集BUPT-Xface和RFW已被来自哈佛大学、麻省理工学院等近**600**名科研人员申请使用。

Mei Wang, Weihong Deng, et al. Racial faces in the wild: Reducing racial bias by information maximization adaptation network. *IEEE ICCV* 2019. 谷歌引用: 137

## 成果一：种族偏见数据集

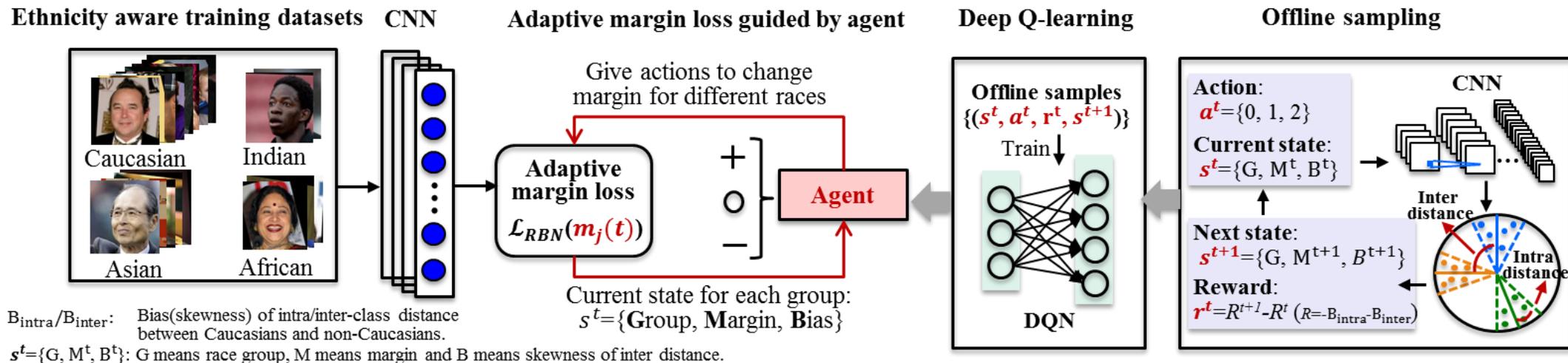
创新点①：证明现有深度人脸识别算法均存在偏见，黑人的平均错误率是白人的两倍

| 模型    | Model       | RFW          |              |              |              |             |
|-------|-------------|--------------|--------------|--------------|--------------|-------------|
|       |             | 白人           | 印度人          | 亚洲人          | 黑人           | 方差          |
| 算法    | Center-loss | 87.18        | 81.92        | 79.32        | 78.00        | 4.07        |
|       | SphereFace  | 90.80        | 87.02        | 82.95        | 82.28        | 3.96        |
|       | ArcFace     | 92.15        | 88.00        | 83.98        | 84.93        | 3.68        |
|       | VGGFace2    | 89.90        | 86.13        | 84.93        | 83.38        | 2.78        |
|       | <b>Mean</b> | <b>90.01</b> | <b>85.77</b> | <b>82.80</b> | <b>82.15</b> | <b>3.58</b> |
| 商业API | Face++      | 93.90        | 88.55        | 92.47        | 87.50        | 3.07        |
|       | Baidu       | 89.13        | 86.53        | 90.27        | 77.97        | 5.56        |
|       | Amazon      | 90.45        | 87.20        | 84.87        | 86.27        | 2.37        |
|       | Microsoft   | 87.60        | 82.83        | 79.67        | 75.83        | 4.98        |
|       | <b>Mean</b> | <b>90.27</b> | <b>86.28</b> | <b>86.82</b> | <b>81.89</b> | <b>3.44</b> |

表格1 RFW数据集上四种人脸算法和四种商业API的人脸验证识别率

# 成果二：动态间距公平性学习-强化学习

创新点②：利用**强化学习**算法，提出**自主调参智能体**，自动为有色人种学习合适的分类间距，提升种族识别率的公平性。

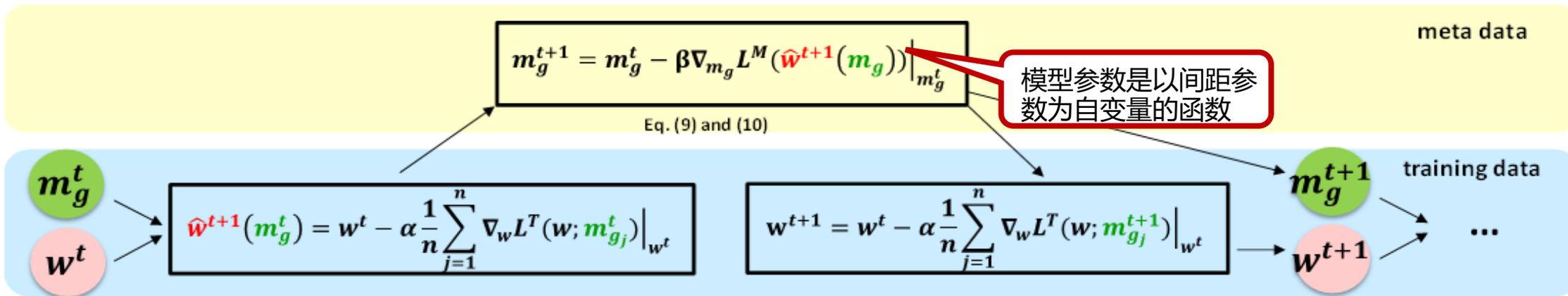


| BUPT-Globalface |              |              |              |              |              |
|-----------------|--------------|--------------|--------------|--------------|--------------|
| 方法              | 白人           | 印度人          | 亚洲人          | 黑人           | 方差           |
| arcface         | 97.37        | 95.68        | 94.55        | 93.87        | 1.53         |
| RL-RBN          | <b>97.08</b> | <b>95.63</b> | <b>95.57</b> | <b>94.87</b> | <b>0.93↓</b> |

| BUPT-Balanced |              |              |              |              |              |
|---------------|--------------|--------------|--------------|--------------|--------------|
| 方法            | 白人           | 印度人          | 亚洲人          | 黑人           | 方差           |
| arcface       | 96.18        | 94.67        | 93.72        | 93.98        | 1.11         |
| RL-RBN        | <b>96.27</b> | <b>94.68</b> | <b>94.82</b> | <b>95.00</b> | <b>0.73↓</b> |

# 成果三：动态间距公平性学习-元学习

创新点③：不同于强化学习的离散搜索，引入元学习，基于元数据集梯度，连续地为人种搜索间距，进一步提升公平性。

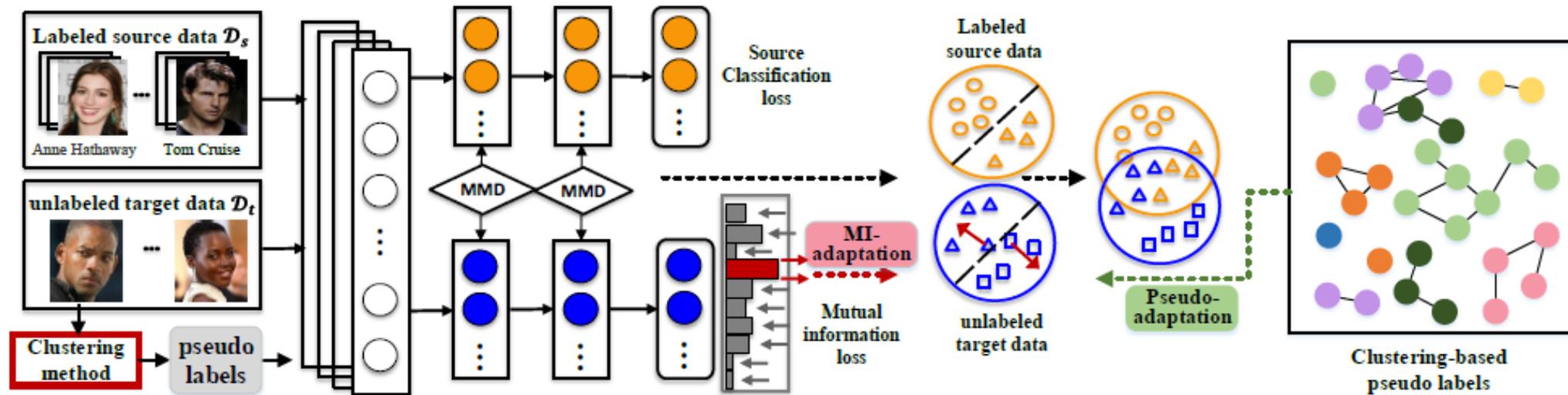


| BUPT-Globalface |       |       |       |       |       |
|-----------------|-------|-------|-------|-------|-------|
| 方法              | 白人    | 印度人   | 亚洲人   | 黑人    | 方差    |
| arcface         | 97.37 | 95.68 | 94.55 | 93.87 | 1.53  |
| MBN             | 96.87 | 96.20 | 95.63 | 95.00 | 0.80↓ |

| BUPT-Balanced |       |       |       |       |       |
|---------------|-------|-------|-------|-------|-------|
| 方法            | 白人    | 印度人   | 亚洲人   | 黑人    | 方差    |
| arcface       | 96.18 | 94.67 | 93.72 | 93.98 | 1.11  |
| MBN           | 96.25 | 95.32 | 94.85 | 95.38 | 0.58↓ |

# 成果四：跨种族自适应学习

创新点④：提出了种族迁移互信息网络，利用聚类算法标注伪标签，提出互信息损失进行迁移，提升了网络的泛化能力。



| 方法      | 白人    | 印度人           | 亚洲人           | 黑人            |
|---------|-------|---------------|---------------|---------------|
| softmax | 94.12 | 88.33         | 84.60         | 83.47         |
| IMAN-S  | -     | <b>91.08↑</b> | <b>89.88↑</b> | <b>89.13↑</b> |

| 方法      | 白人    | 印度人           | 亚洲人           | 黑人            |
|---------|-------|---------------|---------------|---------------|
| arcface | 94.78 | 90.48         | 86.27         | 85.13         |
| IMAN-A  | -     | <b>94.15↑</b> | <b>91.15↑</b> | <b>91.42↑</b> |

Mei Wang, Weihong Deng, et al. Racial faces in the wild: Reducing racial bias by information maximization adaptation network. ICCV 2019.



**科研经历**



**科研方向**



**心得体会**

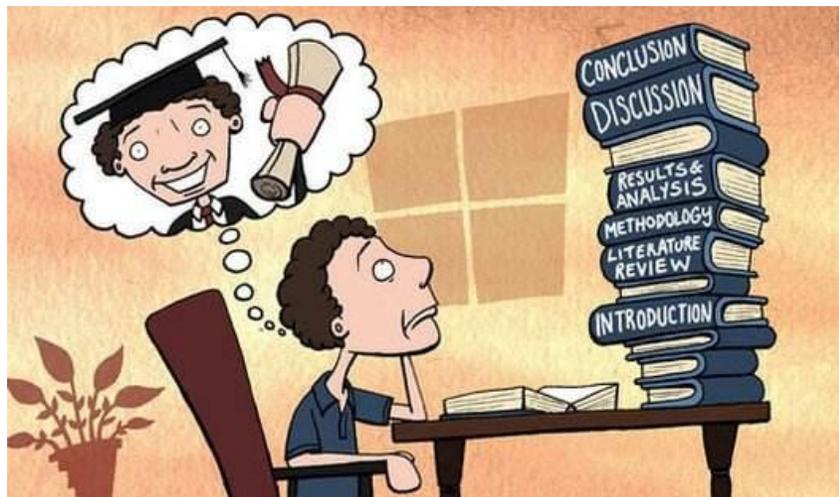


# 心得体会



## 如何调研

积累知识基础、确定idea新颖性、多领域能够提供灵感



师兄师姐推荐/综述

Google scholar/Arxiv搜索关键词

刷本领域大牛的主页

文献的参考文献/related work



## 如何读论文

**中文or英文**

**逐句翻译→熟悉专业词汇**

**理清逻辑→建立阅读习惯**

**整理笔记→增强记忆**

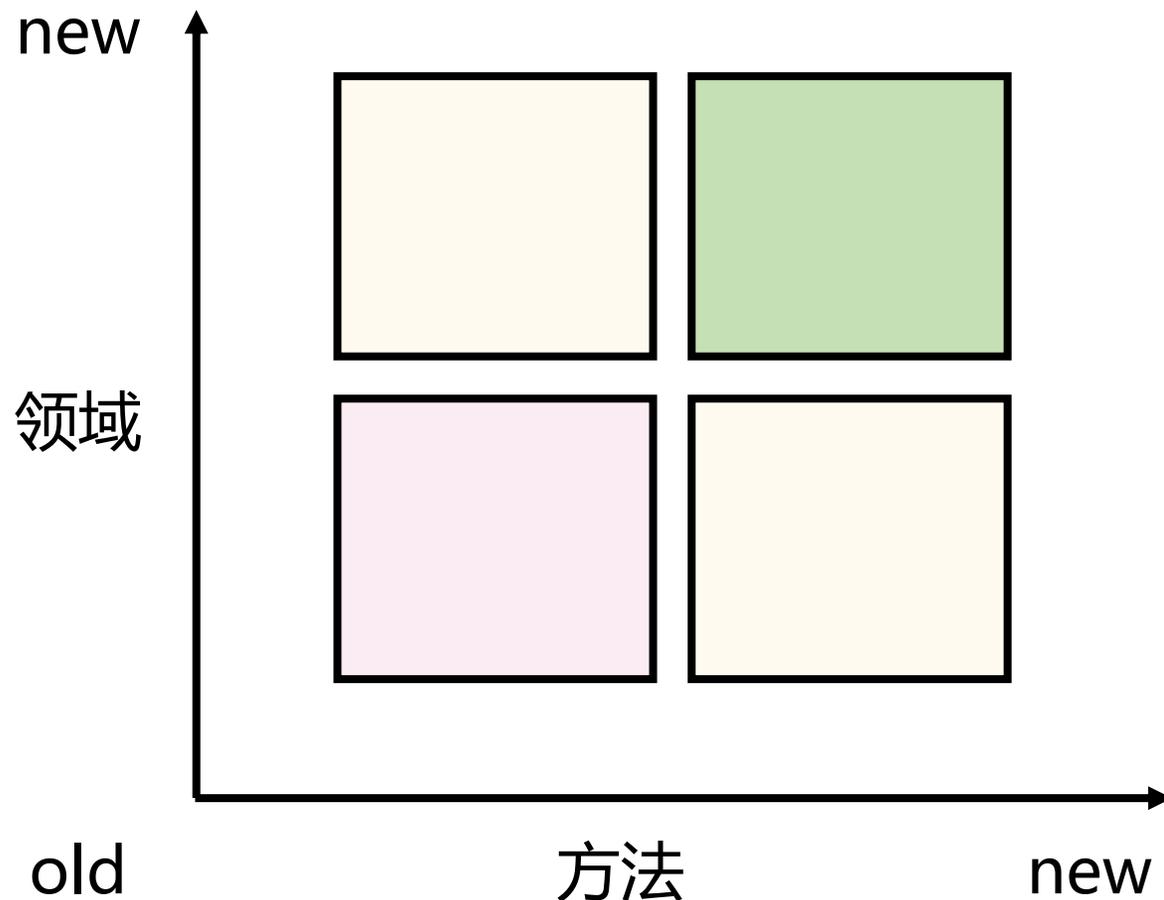
真正看懂看透彻20篇论文足以。  
文献看3遍才能体会深意。  
保持阅读习惯，接受新思想。



# 心得体会



## 创新点





## 如何写论文



Latex/Word

**论文写作对录用影响非常大**

**简洁且有逻辑性的阐述自己的idea**

**明确且包装自己的创新点**

**平时摘抄好的英文句子/短语/连接词**



## Rebuttal重要性

下图是CVPR2019总计超过15000条review在rebuttal前后的状态对比统计

| 最终<br>首轮      | Strong<br>Accept | Accept | Borderline<br>Accept | Borderline<br>Reject | Reject |
|---------------|------------------|--------|----------------------|----------------------|--------|
| Strong Accept | 45.9%            | 30.3%  | 7.4%                 | 4.0%                 | 1.2%   |
| Weak Accept   | 2.0%             | 39.7%  | 32.2%                | 13.0%                | 2.5%   |
| Borderline    | 0.3%             | 8.0%   | 25.3%                | 45.1%                | 10.9%  |
| Weak Reject   | 0.0%             | 0.9%   | 6.0%                 | 35.9%                | 47.1%  |
| Reject        | 0.1%             | 0.0%   | 0.7%                 | 4.3%                 | 84.0%  |

- strong reject → positive: 0.8%;
- **weak reject → positive: 6.9%;**
- **borderline → positive: 33.6%;**
- borderline → negative: 56%;
- weak accept → negative: 15.5%;
- strong accept → negative: 5.2%。



## 如何rebuttal

### Rebuttal原则

- 意见逐条回复，不要回避忽略，尽量揣摩意思
- Positive尽量迎合
- Negative据理力争
- 不要节外生枝，画蛇添足
- 礼貌用语

### 常见审稿意见

- Novelty不足（低分原因）
- 描述错误：假设不合理，语言表达不合理，方法有缺陷等
- 效果不明显（提升有限）
- 实验不充分（补充实验）
- 语法、结构、参考文献遗漏等



北京邮电大学

---

感谢聆听

祝大家科研顺利