# 测试&测试集

# 一，Benchmark

## 1.MIFA自采A柱：

### 统计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **版本** | **人数** | **嘘** | **V** | **无动作** |
| **Version\_20220103** | 6 | 51 | 44 | 85 |
| **Version\_20220125** | 6 | 38 | 28 | 84 |

测试段地址：oss://alios-cv/qiuchao.qiu/static\_hand\_testset/inCar/MIFA/mifa\_segment

*\*\*\*子目录名称为对应的标签ID*

### 样本示例:

[MISSING IMAGE: brx\_real\_38ID.mp4 (6.36MB), brx\_real\_38ID.mp4 (6.36MB) ][MISSING IMAGE: vs\_20211230-151455\_A\_hand\_zhouye\_882ID.mp4 (5.25MB), vs\_20211230-151455\_A\_hand\_zhouye\_882ID.mp4 (5.25MB) ][MISSING IMAGE: vs\_20211230-160250\_A\_hand\_liyujun\_236ID.mp4 (4.08MB), vs\_20211230-160250\_A\_hand\_liyujun\_236ID.mp4 (4.08MB) ]

## 2.ADS采集：

### 统计

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **版本** | **人数** | **干扰** | **无动作** | **嘘** | **666** | **V** | **比心** | **ok** | **赞** | **握拳** | **比5** | **左** | **右** |
| **Version\_A\_20220103** | 30 | 622 | 199 | 143 | ​ | 286 | ​ | ​ | ​ | ​ | ​ | ​ | ​ |
| **Version\_A\_20220125** | 30 | 200 | 199 | 143 | ​ | 286 | ​ | ​ | ​ | ​ | ​ | ​ | ​ |

测试段地址：oss://alios-cv/qiuchao.qiu/static\_hand\_testset/inCar/ADS/ads\_segment

*\*\*\*子目录名称为对应的标签ID*

### 样本示例:

[MISSING IMAGE: StaticGesture\_A\_IR\_Normal\_137\_202112111632\_1699ID.mp4 (4.76MB), StaticGesture\_A\_IR\_Normal\_137\_202112111632\_1699ID.mp4 (4.76MB) ][MISSING IMAGE: StaticGesture\_A\_IR\_Normal\_001\_202110291545\_149ID.mp4 (1.47MB), StaticGesture\_A\_IR\_Normal\_001\_202110291545\_149ID.mp4 (1.47MB) ]

### 干扰样本示例：

|  |  |  |  |
| --- | --- | --- | --- |
| [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] |
| [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] |

# 二，测试方法

**测试样本为~=2s的视频片段（当前可能包含更短的视频片段），一个片段代表一个手势事件**

**正样本：1.抬手-稳定，2.稳定，3稳定-放下**

**负样本：1.无动作，2.有干扰动作**

## 精度/召回计算：

**1.算法输出视频片段每帧结果（多帧平滑）**

**2.结果中出现的第一个有效手势即为此片段预测结果**

## 算法延时计算：

**1.第一个有效手势帧的帧号和标注稳定帧号求差即为延时**

**2.即使算法输出的有效手势不是正确结果也参与计算延时**

**3.如果没有识别出任何结果,则此段测试样本不参与计算延时**

# 三，测试结果

模型更新的详细记录查看对应单模块记录：[检测器](https://yuque.antfin-inc.com/zppdqz/tg1g0l/kyby1t)；[分类器](https://yuque.antfin-inc.com/zppdqz/tg1g0l/tn6ka4)

MIFA项目交付：<https://yuque.antfin-inc.com/zppdqz/arh88c/gk37uk>

## 更新

### 0207：

检测器优化------>slim\_unsim\_20220129：提升检测框召回

分类器优化------>base\_clean\_unsim\_20220207：抑制干扰数据召回

工程代码：

1.修改分类器输入数据标准化参数：mean=0,std=1

### 0209:

工程代码：

1.修正之前平滑窗口实际大小比设定值大1问题

2.删除利用中心点判断误检测代码--->有逻辑bug

3.增加判断框过大>3/4xImg\_H or 3/8xImg\_W，或过小<Img\_H/10 or Img\_W/20

### 0217：

分类器优化------>base\_clean\_unsim\_20220216：提升召回

### 0219:

1.优化后处理代码：去除过大/过小，长宽比异常的框

2.去除跟踪代码，采用隔帧检测方法。隔帧可以不做前处理，省去前处理时间。

### 0222:

1.优化后处理过滤逻辑

2.增加各个模块详细结果显示

### 0225:

1.为分类器增加softmax函数和得分阈值判断：嘘手势只有大于一定阈值（0.8）才能输出

2.增加检测器阈值=0.3

3.分类器优化------>base\_clean\_unsim\_20220224：抑制背景误识别

### 0225\_2:

1.分类器优化------>base\_clean\_unsim\_20220224\_3：抑制背景误识别

## 综合：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***全数据（10fps）*** | **嘘+V** | | | | **嘘** | | | | | **V** | | | | | **负样本** | | | |
| **命中** | **预测** | **真实** | **P/R** | **命中** | **预测** | **真实** | **P/R** | **延时(s)** | **命中** | **预测** | **真实** | **P/R** | **延时(s)** | **命中** | **预测** | **真实** | **P/R** |
| **1.25** | 389 | 428 | 495 | 0.909/0.786 | 125 | 142 | 181 | 0.880/0.691 | 0.585 | 264 | 286 | 314 | 0.923/0.841 | 0.659 | 446 | 550 | 483 | 0.811/0.923 |
| **1.25--fix bugs** | 392 | 435 | 495 | 0.901/0.792 | 124 | 143 | 181 | 0.867/0.685 | 0.581 | 268 | 292 | 314 | 0.918/0.854 | 0.646 | 444 | 543 | 483 | 0.818/0.919 |
| **1.26** | 394 | 430 | 495 | 0.916/0.795 | 143 | 155 | 181 | 0.923/0.790 | 0.66 | 251 | 275 | 314 | 0.913/0.799 | 0.622 | 449 | 548 | 483 | 0.819/0.930 |
| **2.07** | 435 | 470 | 495 | 0.926/0.879 | 163 | 179 | 181 | 0.911/0.901 | 0.543 | 272 | 291 | 314 | 0.935/0.866 | 0.623 | 451 | 508 | 483 | 0.888/0.934 |
| **2.09** | 443 | 487 | 495 | 0.910/0.895 | 166 | 187 | 181 | 0.888/0.917 | 0.424 | 277 | 300 | 314 | 0.923/0.882 | 0.476 | 444 | 491 | 483 | 0.904/0.919 |
| **2.09\_dsp** | 418 | 462 | 495 | 0.905/0.844 | 155 | 175 | 181 | 0.886/0.856 | 0.473 | 263 | 287 | 314 | 0.916/0.838 | 0.580 | 446 | 516 | 483 | 0.864/0.923 |
| **2.09\_dsp--fix bugs** | 437 | 476 | 495 | 0.918/0.883 | 162 | 177 | 181 | 0.915/0.895 | 0.452 | 275 | 299 | 314 | 0.920/0.876 | 0.531 | 449 | 502 | 483 | 0.894/0.930 |
| **2.17** | 448 | 477 | 495 | 0.939/0.905 | 158 | 167 | 181 | 0.946/0.873 | 0.431 | 290 | 310 | 314 | 0.935/0.924 | 0.523 | 455 | 501 | 483 | 0.908/0.942 |
| **2.19** | 451 | 479 | 495 | 0.942/0.911 | 157 | 166 | 181 | 0.946/0.867 | 0.182 | 294 | 313 | 314 | 0.939/0.936 | 0.214 | 456 | 499 | 483 | 0.914/0.944 |
| **2.22** | 446 | 474 | 495 | 0.941/0.901 | 156 | 165 | 181 | 0.945/0.862 | 0.175 | 290 | 309 | 314 | 0.939/0.924 | 0.214 | 456 | 504 | 483 | 0.905/0.944 |
| **2.22检测阈值=0.3** | 418 | 429 | 495 | 0.974/0.844 | 154 | 156 | 181 | 0.987/0.851 | 0.186 | 264 | 273 | 314 | 0.967/0.841 | 0.189 | 472 | 549 | 483 | 0.844/0.974 |
| **2.25** | 353 | 360 | 495 | 0.981/0.713 | 97 | 97 | 181 | 1.000/0.536 | 0.38 | 256 | 263 | 214 | 0.973/0.815 | 0.26 | 476 | 618 | 483 | 0.770/0.986 |
| **2.25\_2** | 382 | 393 | 495 | 0.972/0.772 | 122 | 123 | 181 | 0.992/0.674 | 0.312 | 260 | 270 | 314 | 0.963/0.828 | 0.189 | 472 | 585 | 483 | 0.807/0.977 |

## 长时间误报率测试：

**测试方法：**

1.以20帧（2s）为一个测试段。

2.只要出现嘘或V则算一次误报。一个测试段至多有一次误报。

**1.数据（A柱-ADS）：**

1.oss://adsdata/xin\_BanMa\_Video\_202111/20211120/LipDetect/

2.oss://adsdata/xin\_BanMa\_Video\_202111/20211123/LipDetect/005~013

77个视频，约1h。

|  |  |
| --- | --- |
| **唇动数据（10fps）** | **误报次数/h** |
| **2.09** | 30 |
| **2.17** | 35 |
| **2.19** | 16 |
| **2.22** | 14 |
| **2.22检测阈值=0.3** | 0 |
| **2.25** | 0 |
| **2.25\_3** | 0 |

**2.数据（MIFA实车采集）:**

9个视频，约25min

|  |  |
| --- | --- |
| **MIFA（10fps）** | **误报次数/25min** |
| **2.22** | 183 |
| **2.22检测阈值=0.3** | 65 |
| **2.25** | 3 |
| **2.25\_3** | 0 |

## MIFA

**Test set: Version\_20220103(*10fps*)**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***MIFA实车（10fps）*** | **嘘** | | | | | **V** | | | | | **负样本** | | | |
| **命中** | **预测** | **真实** | **P/R** | **延时(s)** | **命中** | **预测** | **真实** | **P/R** | **延时(s)** | **命中** | **预测** | **真实** | **P/R** |
| **1.25** | 34 | 37 | 38 | 0.919/0.895 | 0.585 | 26 | 26 | 28 | 1.000/0.929 | 0.659 | 81 | 87 | 84 | 0.931/0.964 |
| **1.26** | 31 | 32 | 38 | 0.969/0.816 | 0.65 | 19 | 19 | 28 | 1.000/0.679 | 0.62 | 83 | 99 | 84 | 0.838/0.988 |
| **2.07** | 32 | 39 | 38 | 0.821/0.842 | 0.543 | 25 | 29 | 28 | 0.862/0.893 | 0.623 | 76 | 82 | 84 | 0.927/0.905 |
| **2.09** | 33 | 41 | 38 | 0.805/0.868 | 0.424 | 25 | 29 | 28 | 0.862/0.893 | 0.424 | 75 | 80 | 84 | 0.937/0.893 |
| **2.17** | 33 | 35 | 38 | 0.943/0.868 | 0.431 | 26 | 28 | 28 | 0.929/0.929 | 0.523 | 80 | 87 | 84 | 0.920/0.952 |
| **2.19** | 31 | 36 | 38 | 0.861/0.816 | 0.182 | 26 | 28 | 28 | 0.929/0.929 | 0.214 | 78 | 86 | 84 | 0.907/0.929 |

## ADS

**Test set: Version\_20220117(*10fps*)**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***ADS车内A柱(10fps)*** | **嘘** | | | | | **V** | | | | | **负样本** | | | |
| **命中** | **预测** | **真实** | **P/R** | **延时(s)** | **命中** | **预测** | **真实** | **P/R** | **延时(frames)** | **命中** | **预测** | **真实** | **P/R** |
| **1.25** | 91 | 105 | 143 | 0.867/0.636 | - | 238 | 260 | 286 | 0.915/0.832 | - | 365 | 463 | 399 | 0.788/0.915 |
| **2.07** | 131 | 140 | 143 | 0.936/0.916 |  | 247 | 262 | 286 | 0.943/0.864 |  | 375 | 426 | 399 | 0.880/0.940 |
| **2.09** | 133 | 146 | 143 | 0.911/0.930 |  | 252 | 271 | 286 | 0.930/0.881 |  | 369 | 411 | 399 | 0.898/0.925 |
| **2.17** | 125 | 132 | 143 | 0.947/0.874 |  | 264 | 282 | 286 | 0.936/0.923 |  | 375 | 414 | 399 | 0.906/0.940 |
| **2.19** | 126 | 130 | 143 | 0.969/0.881 |  | 268 | 285 | 286 | 0.940/0.937 |  | 378 | 413 | 399 | 0.915/0.947 |

## Badcase

1.嘘和V手势的漏召回

|  |  |
| --- | --- |
| [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] |
| [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] |

2.背景/干扰动作误识别

|  |  |  |
| --- | --- | --- |
| [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] |
| [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] | [MISSING IMAGE: image.png, image.png ] |

#### Badcase分布情况：

**1h误报测试集：**

人脸：1

部分嘴巴+背景：8 （连续的误识别）

纯背景：2

手部：3

**指标测试集：**

部分嘴巴+背景：6

强干扰：18

纯背景：3