







### 主要解决的四个问题





#### Sources & Works



机器学习算法 支持向量机 包华清提供



实时摄**像人**脸**数据搜集** 基于OPENCV

https://www.cnblogs.com/xp12345/ p/9818435.html



数据集 IMDB-WIKI

https://data.vision.ee.ethz.ch/cvl/rrot he/imdb-wiki/





### 选择合适的数据集 (包含性别和年龄的标签)

数据集: IMDB-WIKI- 500k+ face images with age and gender labels ttps://data.vision.ee.ethz.ch/cvl/r

网址:

**Second Step** 



## 图像处理

- 1、截取人脸灰度图片
- 2、调整为同一大小(255,255)
- 3、小波分解

## 模型建立与检验 •

- 1、将经过处理的数据集分为训练集+检验集
- 2、将训练集导入到机器学习算法中,建立模型A
- 3、利用检验集检验模型的准确度

#### Fourth Step

#### Third Step

### 模型应用

- 1、通过摄像头实时采集包含人脸的图像或使用其他包含人脸的图像
- 2、对图像进行像素和人脸的截取
- 3、将经过处理的新图像导入模型,自动生成图像中人物的性别信息

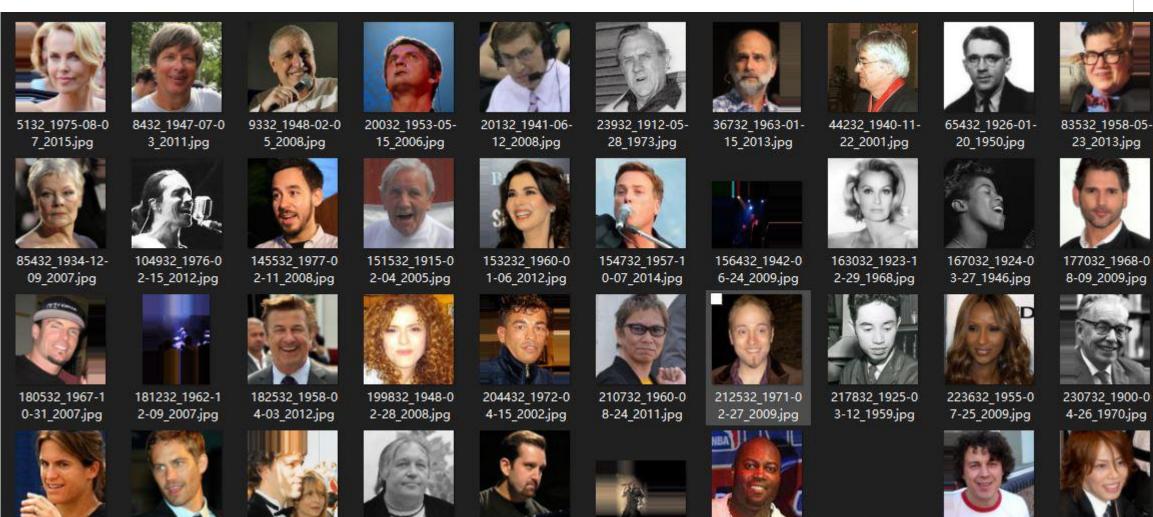






### 数据集

# 数据集: IMDB-**WIKI-500k**+ face images **with age and gender labels** 网址:https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/





## 图像处理



## 截取人脸灰度图片

调整为同一大小







### 模型的生成与检验





## 优化数据集

# 去除噪声

















### 优化前的数据集



5132\_1975-08-0 7\_2015.jpg



85432 1934-12-09 2007.jpg



180532 1967-1 0-31\_2007.jpg





8432 1947-07-0 3 2011.jpg

104932 1976-0

2-15\_2012.jpg

181232 1962-1

2-09\_2007.jpg



145532 1977-0 2-11\_2008.jpg

9332 1948-02-0

5\_2008.jpg



182532 1958-0 4-03\_2012.jpg





20032 1953-05-15\_2006.jpg

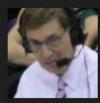


151532 1915-0 2-04\_2005.jpg



199832 1948-0 2-28\_2008.jpg





20132 1941-06-12 2008.jpg



153232 1960-0 1-06\_2012.jpg



204432 1972-0 4-15\_2002.jpg





28\_1973.jpg

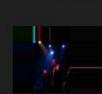
154732 1957-1

0-07\_2014.jpg

210732 1960-0

8-24\_2011.jpg

36732 1963-01-15\_2013.jpg



156432 1942-0 6-24 2009.jpg



212532 1971-0 2-27\_2009.jpg





44232 1940-11-22\_2001.jpg



163032 1923-1 2-29\_1968.jpg



217832 1925-0 3-12\_1959.jpg



65432 1926-01-20\_1950.jpg



83532 1958-05-23\_2013.jpg



177032\_1968-0 8-09\_2009.jpg



167032 1924-0

223632 1955-0



230732 1900-0 4-26\_1970.jpg





### 优化后的数据集





11.jpg

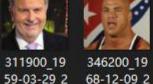


196900 18 84-05-20 1 940.jpg



305500 19 311200 19 40-09-18 1 43-09-21 2 013.jpg







64-09-14 1

994.jpg











65-07-01 2

010.jpg











963.jpg







010.jpg



005.jpg











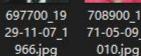






615800 19 70-06-07 2 010.jpg





71-05-09 2 67-02-20 1 993.jpg

756400 19 83-11-15 2 008.jpg

764000 19 54-03-17 2 008.jpg

784200 19 62-06-27 2 006.jpg

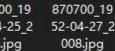
787300 19 67-03-16 2 014.jpg

797100 19 17-11-02 1 940.jpg

805000 19 60-07-30 2 015.jpg

855800 19 43-08-18 1 976.jpg

868000 19 77-04-25 2 009.jpg





20-05-20 2

001.jpg





886500 19

05-07-22 1



944600 19

68-02-05 2

006.jpg



947500 19

76-03-13 2



42-04-16 2

011.jpg



1035700 1

945-11-27

2013.jpg









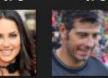
1285200 1

975-01-05









882800 19 67-12-07 2













976-07-31

2007.jpg

981-08-18

2012.jpg

1638200 1

2009.jpg

1653100 1



1315300 1

965-09-09

2015.jpg

1687000 1

1335100 1

954-09-18

2007.jpg



1363900 1

978-02-02

2010.jpg



1698400 1 958-12-10 2008 ina

1372000 1

971-12-02

2009.jpg

1422400 1 958-01-07 1987 ina

1427800 1 932-05-08 2005 ina



1536200 1 982-06-25 2014 ina

1554700 1 965-12-07 2013 ina

1576600 1 970-12-02 2012 ina

1959 ina

1578500 1 919-06-14

1591900 1 970-04-13 2007 ina

931-01-30 2007 ina

980-01-25 2008 ina

1665600 1 957-03-20 2007 ina

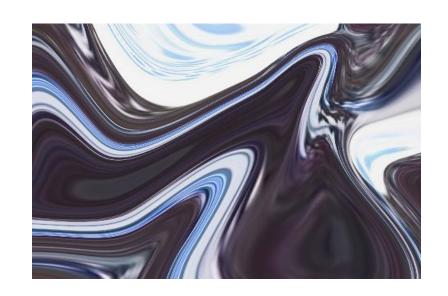
932-12-12 1962 ina

1694300 1 928-03-16



### 新模型的生成与检验

人工优化了15个文件夹,包含大约5000张人脸图片,随机抽取1500张图片, 其中70%作为训练集进行新的模型训练,30%作为测验集测算准确度。



● 重复前文描述的2-3的过程生成模型B



### 两个模型的比较

模型A

数据集大小: 1500张图像



模型B

数据库大小: 1500张图像

性别准确度: 60%

性别准确度: 70%





### 人脸的性别与年龄的检测

通过摄像头实时进行人脸数据搜集



选取包含人脸的图片作为展示样例



### Showtime

















进一步提高性别的准确度



年龄模型的建立

