

# 第三版预处理安装与配置

## 版本要求

- basic: 1.5.8

## 预处理服务安装

以 131.16.42.142 为例：

● 1 省院-测试 ( 131.16.42.142... × +

```
[root@localhost scandocprocess]# pwd
/home/visionread/scandocprocess
[root@localhost scandocprocess]# ll
total 2897712
-rw-r--r-- 1 visionread visionread      591 Mar 14 19:28 docker-compose.yml
-rw-rw-r-- 1 visionread visionread 2967250432 Mar 14 18:48 yuchuli-3.3.tar.gz
[root@localhost scandocprocess]#
```

- yuchuli-3.3.tar.gz：预处理镜像

### 1、加载镜像

```
docker load -i yuchuli-3.3.tar.gz
```

### 2、查看镜像

```
docker images
```

```
[root@localhost scandocprocess]# docker images
REPOSITORY                                TAG                IMAGE ID           CREATED           SIZE
whrp.biz.case-info-service                1.0               a7ad9e0ffbeb      25 hours ago     908MB
whrp.biz.basic-service                    1.0               8b2dd814d9dd      25 hours ago     1.02GB
whrp.biz.bms-service                      1.0               ff4f5c042151      2 days ago       931MB
registry.internal.purvar.com:7443/pail/scandocprocess v3.3             1989a8f85e29      7 days ago       2.92GB
whrp.biz.user-service                    1.0               6e4c81c02791      7 days ago       754MB
registry.internal.purvar.com:7443/pail/scandocprocess v3.2             ec3eaef050ba8     2 weeks ago      2.92GB
gitlab/gitlab-ce                         16.7.6-ce.0       128159fad5f3      4 weeks ago      2.89GB
whrp.biz.case-info-transfer-service       1.0               37ffda230f05      4 weeks ago      906MB
whrp.biz.gateway-service                 1.0               1cabae9eda5f      2 months ago     686MB
<none>                                   <none>            90b5f798deba      2 months ago     928MB
<none>                                   <none>            ee9e995d49c3      2 months ago     928MB
whrp.base.mysql                          5.7.30            4fb2023a3790      3 months ago     458MB
tesseract-openjdk                        4.0-1.8-2         c5466323daa0      4 months ago     499MB
whrp.biz.hebeiapp                        2.0               4dce68dae407      4 months ago     788MB
sba-service                             2.6.3             b7b8d6fc0b54      4 months ago     584MB
whrp.base.rabbitmq                       3.7.26            bdb6cbc53ce0      4 months ago     180MB
whrp.base.redis                          1.34.0            db1060e2263       11 months ago    207MB
```

### 3、修改 docker-compose.yml

只修改映射路径：

这里预处理的 /img\_process/output 映射出来路径要和 basic 中 /tmp/preimg 映射的路径保持一致；

input 路径无要求；

```
[root@localhost scandocprocess]# cat docker-compose.yml
version: "3"
services:
  scandocprocess:
    restart: always
    image: registry.internal.purvar.com:7443/pail/scandocprocess:v3.3
    build:
      context: ./
      dockerfile: Dockerfile
    logging:
      options:
        max-size: "10m"
        max-file: "10"
    ports:
      - "8111:8000"
    environment:
      BASE_DIR: "/img_process"
      LD_LIBRARY_PATH: "/img_process/depoly/app/orientation"
    volumes:
      - "/home/visionread/data/tmp/scandocprocess/input:img_process/input"
      - "/home/visionread/data/tmp/scandocprocess/output:/img_process/output"
[root@localhost scandocprocess]#
```

和basic中/tmp/preimg一致

### 4、启动服务

```
docker-compose up -d
```

### 5、添加定时任务，定时删除预处理图片

```
]# crontab -e

0 */1 * * * find /home/visionread/data/tmp/scandocprocess/output -type f -cmin +360
-exec rm -rf {} \;
0 */1 * * * find /home/visionread/data/tmp/scandocprocess/input -type f -cmin +360 -
exec rm -rf {} \;
```

## 业务配置

### basic application.yml 配置

可参考 131.16.42.142 中basic配置

## 1、handler 新增 preImg3ServiceImpl，代表使用第三版预处理

```
116 |
117 | # 预处理
118 | pre-img:
119 |   # 并发数
120 |   concurrency: 4
121 |   handler: preImg3ServiceImpl #可选值 preImgServiceImpl = 广州广电 或 preImg2ServiceImpl = 武汉方正 preImg3ServiceImpl-武汉方正第三版
122 |   service2:
123 |     url: http://131.16.42.152:8000
124 |     config:
125 |       auto_fill_color: false
126 |       rm_blank: true # 是否去除空白页
127 |       blank_thresh: 0.2 # 空白页判断阈值[0, 1], 值越大越容易被判断为空白页, 默认0.5
128 |       rm_noise: false # 是否去除噪点
129 |       noise_thresh: 0.1
130 |       noise_param_1: 101
131 |       noise_param_2: 51
```

## 2、阈值配置

```
119 | # 并发数
120 | concurrency: 4
121 | handler: preImg3ServiceImpl #可选值 preImgServiceImpl = 广州广电 或 preImg2ServiceImpl = 武汉方正 preImg3ServiceImpl-武汉方正第三版
122 | service2:
123 |   url: http://131.16.42.152:8000 # 第二版
124 |   config:
125 |     auto_fill_color: false
126 |     rm_blank: true # 是否去除空白页
127 |     blank_thresh: 0.2 # 空白页判断阈值[0, 1], 值越大越容易被判断为空白页, 默认0.5
128 |     rm_noise: false # 是否去除噪点
129 |     noise_thresh: 0.1
130 |     noise_param_1: 101
131 |     noise_param_2: 51
132 |     rm_black: true # 是否去黑边
133 |     black_thresh: 0.01 # 黑边判断阈值(0,0.5),值越大对黑边的容忍程度越大, 默认值为 0.005;
134 |     black_area_thresh: 1000 # 默认为5000, 范围小于图片面积即可; 黑边面积阈值, 大于阈值的区域不被判定为黑边
135 |     rm_binder: true # 是否去除印痕
136 |     binder_thresh: 0.09 #是否去除印痕阈值 [0,0.5],值越大被判定为印痕的区域数量越多, 但错误消除的概率会增加,默认 0.15;
137 |     rt_skew: true # 是否纠偏
138 |     skew_thresh: 0.001 # 歪斜角度阈值(0, 90), 歪斜角度大于阈值将被判断为歪斜, 默认0.5]
139 | service3: #第三版预处理
140 |   url: http://131.16.42.142:8111 # 第三版
141 |   detect: #检测
```

```
service3: #第三版预处理
  url: http://131.16.42.142:8111
  detect: #检测
    detect_blank: # 图像空白检测 { "isBlank": 1 }
      enabled: true
      blank_thresh: 0.5 # 默认值为0.5, 范围[0,1], 深色区域在图像中的面积占比; 值越大图像越容易被判定为空白。
    detect_black_edge: # 图像黑边检测 { "haveBlack": 1 }
      enabled: false
      black_thresh: 0.01 # 默认为0.03, 范围[0, 0.5], 黑边判定边界范围; 值越大越容易判定为有黑边, 但错误消除的概率会增加。
      black_area_thresh: 3000 # 默认为3000, 范围小于图片面积即可; 黑边面积阈值, 大于阈值的区域不被判定为黑边。
    detect_binder: # 图像印痕检测 { "haveBinder": 1 }
      enabled: false
      binder_thresh: 0.15 # 默认为0.15, 范围[0, 0.5], 印痕判定边界范围; 值越大被判定为印痕的区域数量越多, 但错误消除的概率会增加。
    detect_skew: # 图像歪斜检测 { "isSkew": 1 }
      enabled: false
      skew_thresh: 0.5 # 默认为0.5, 范围[0, 90), 图像偏转角度阈值; 图像偏转角度大于阈值为歪斜。
    detect_orientation: # 颠倒检测 { "direc": 0 } 0:图片为正
      1:图片为倒 2:图片顺时针旋转90度 3:图片逆时针旋转90度
```

```

        enabled: false
        detect_noise:                                # 噪声检测                { "haveNoise": 1 }
            enabled: false
            noise_thresh: 0.5                        # 默认:0.5 范围:[0,1], 值越小越容易判定为有噪声.
        detect_stamps_fingerprint:                    # 图像印章指纹检测                {
"haveStampsFingerprint": 1 }
            enabled: false
            lower_area: 0.1                          # 默认:0.1, 范围:[0,100](比upper_area小), 印章
和指纹与原图的面积比下限.
            upper_area: 5                            # 默认:0.1, 范围:[0,100](比lower_area大), 印章
和指纹与原图的面积比下限.
        detect_content: # 图像内容识别检测            { "rect": [214, 527, 2122, 1629] }
            enabled: false
            ignore_range: 0.5                        # 默认0.5, 范围[0, 1], 值越大检测的范围细,但可能
会产生缺失.
        detect_blur:                                # 图像虚化检测                { "isBlur": 1 }
            enabled: false
            blur_thresh: 0.2                         # 默认0.2, 范围[0,1], 越大越容易被判定为虚化.

        remove: #优化
            remove_black_edge:                       # 图像去黑边
                enabled: true
                black_thresh: 0.01                   # 默认为0.03, 范围[0, 0.5], 黑边判定边界范围; 值
越大越容易判定为有黑边, 但错误消除的概率会增加.
                black_area_thresh: 3000              # 默认为3000, 范围小于图片面积即可; 黑边面积阈值,
大于阈值的区域不被判定为黑边.
            remove_binder:                           # 图像去除印痕
                enabled: true
                binder_thresh: 0.15                  # 默认为0.15, 范围[0, 0.5], 印痕判定边界范围; 值
越大被判定为印痕的区域数量越多, 但错误消除的概率会增加.
            remove_skew:                             # 图像歪斜矫正
                enabled: true
                skew_thresh: 0.5                     # 默认为0.5, 范围[0, 90), 图像偏转角度阈值; 图像
偏转角度大于阈值为歪斜.
            correct_orientation:                     # 颠倒矫正
                enabled: false
            remove_noise:                             # 噪声去除
                enabled: false
                noise_thresh: 0.5                    # 默认:0.5 范围:[0,1], 值越小越容易判定为有噪声.
            remove_stamps_fingerprint:                # 图像印章指纹消除
                enabled: false
                lower_area: 0.1                      # 默认:0.1, 范围:[0,100](比upper_area小), 印章
和指纹与原图的面积比下限.
                upper_area": 5                      # 默认:0.1, 范围:[0,100](比lower_area大), 印章
和指纹与原图的面积比下限.
            remove_blur:                             # 图像去模糊
                enabled: false
                blur_thresh: 0.2                     # 默认0.2, 范围[0,1], 越大越容易被判定为虚化.
            trapezoid_correct:                       # 图像梯形矫正
                enabled: false
                points: "[[0,0],[1,0],[1,1],[0,1]]" # 梯形四个顶点坐标, 顺序为左上,右上,右下,
左下

```

```
get_A4:                                # A4填充
    enabled: false
    dpi: -1                             # 指定生成的A4图片dpi大小,若值为-1,则保留原图dpi,
若原图无dpi属性信息,则dpi默认为200.
    erase:                              # 图像污点擦出
    enabled: false
    rect: "[[0,0],[0,0]]" # 需要擦除的矩形的左上,右下两个点坐标
```

## 开启预处理

### 1、配置管理中，开启预处理功能

| 配置管理 |                 |  |                         |    |
|------|-----------------|--|-------------------------|----|
| 配置列表 |                 |  |                         |    |
| 序号   | 名称              | 值  | 描述                      | 操作 |
| 1    | convertFiles    | jpg, jpeg, png, html, htm, doc, docx, xls, xlsx, wps, rtf, pdf |                         | 编辑 |
| 2    | fileConvertMode | 7  | 文件转换器 1=预处理 2=缩略图 4=PDF | 编辑 |

### 2、配置管理中，imgOptimizerFlag 开关打开，如图，所有法院开启预处理；

\* 名称:

\* 值:

```
import com.whrp.dossier.caseinfo.entity.CaseInfo;
import cn.hutool.extra.spring.SpringUtil;
import com.whrp.dossier.caseinfo.service.impl.CaseInfoServiceImpl;

def apply(caseId){
    CaseInfoServiceImpl caseInfoService = SpringUtil.getBean(CaseInfoServiceImpl.class);
    CaseInfo caseInfo = null;
    if(caseId){
        caseInfo = caseInfoService.queryById(caseId);
    }
    // 开关默认值
    boolean flag = true;
    // 法院code白名单
    def whiteList = [];
    // 黑名单
    def blackList = [];
    if(caseInfo && caseInfo.getCourt()){
        // 只开启指定法院
        if(!whiteList.isEmpty()){
            if(whiteList.contains(caseInfo.getCourt().code)){
                flag = true;
            }
        }
        // 关闭黑名单
        if(!blackList.isEmpty()){
            if(blackList.contains(caseInfo.getCourt().code)){
                flag = false;
            }
        }
    }
}
```

指定法院开启或者不开启，将法院对应的code值加入对应的白名单/黑名单中即可，如 ['280','281']