

# 人脸识别、声音修复、声纹识别、语音转录与图像增强交接

## 人脸识别、声音修复、声纹识别、语音转录与图像增强

本交接内容涵盖部署于主机 **172.32.1.161**，包括人脸识别（CompreFace）、音频修复（voicefixer）、声纹识别（VoiceprintRecognition-Pytorch）、语音转录（RealtimeSTT，待部署）以及图像增强（BasicSR）等模块。

除图像增强项目 BasicSR 外，其余 均涉及语音输入处理，如麦克风录音或语音音频推理任务。可直接使用现有音频文件或简单麦克风输入（主机 **172.32.1.161** 部署不支持）。同时，所有语音模块均支持 **CPU** 推理运行，无需依赖 GPU，即可满足开发测试和基本应用需求。

BasicSR 项目主要用于图像与视频的超分、去噪与增强，例如用于视频画面清晰化与语音内容同步增强。

### 一、部署主机与目录结构

- 部署主机 IP：**172.32.1.161**
- 用户名：**ubuntu**
- 代码主目录：**/home/ubuntu/code**

当前部署的 5 个核心项目：

项目名称	本地目录	GitHub 地址	支持 CPU	核心功能
CompreFace	CompreFace/	<a href="#">🔗</a>	✓	人脸识别系统（API）
voicefixer	voicefixer/	<a href="#">🔗</a>	✓	音频修复与去噪
VoiceprintRecognition-Pytorch	VoiceprintRecognition-Pytorch/	<a href="#">🔗</a>	✓	声纹注册与识别
RealtimeSTT	<b>未部署</b>	<a href="#">🔗</a>	✓	实时语音识别与转录
BasicSR	BasicSR/	<a href="#">🔗</a>	✓	图像/视频增强、超分、去噪等

## 二、各项目说明与使用建议

### 1. CompreFace - 人脸识别平台

- 部署方式：

```
cd CompreFace
docker-compose up -d
```

- 接口支持：REST API，支持注册、识别、比对等功能。
- 默认端口：**http://localhost:8000**
- 使用建议：适用于构建身份验证系统、人脸门禁、用户匹配。

### 2. voicefixer - 音频修复与增强

- Docker** 使用（已提供 Dockerfile）：

```
docker build -t voicefixer .
docker run -v $(pwd)/data:/data voicefixer
```

- 功能：消除噪音、恢复破音、改善老旧音频质量。

### 3. VoiceprintRecognition-Pytorch - 声纹识别

- 特性功能：
  - 训练/推理/评估/声纹比对一体化
  - 支持命令行与 GUI 使用
- 识别示例（CPU）：

```
python infer_recognition.py --audio1 sample1.wav --audio2 sample2.wav
```

### 4. RealtimeSTT - 实时语音转录系统（待部署）

- 部署建议：

```
git clone https://github.com/KoljaB/RealtimeSTT.git
cd RealtimeSTT
docker-compose up -d
```

- 核心功能：
  - 支持浏览器麦克风接入
  - 实时文字流输出（WebSocket）
- 适用场景：会议转写、字幕生成、智能客服。

### 5. BasicSR - 图像/视频超分与增强

- 目录结构核心：
  - 模型代码：**basicsr/**
  - 推理脚本：**inference/**、**scripts/**
  - 配置文件：**options/**
  - 预训练模型目录：**experiments/**、**results/**
- 安装依赖：

```
pip install -r requirements.txt
python setup.py develop
```

- 推理命令（示例）：

```
python inference/inference_real_esrgan.py \
--input inputs/image.png \
--output results/output.png \
--model_path experiments/pretrained/RealESRGAN_x4.pth
```

- 功能覆盖：
  - 图像超分辨率（ESRGAN）
  - 视频增强
  - 去噪与去压缩伪影