# 语音识别-论文复现-LiveSpeechPortraits

## 电脑环境

操作系统：Windows 10

CPU处理器规格：11th Gen Intel(R) Core(TM) i5-11400H @ 2.70GHz 2.69 GHz

机带RAM：16.0 GB (15.8 GB 可用)

GPU规格：NVIDIA GeForce RTX 3050 Ti Laptop GPU

显示内存VRAM：3990MB

共享内存：8073MB

Python 3.8

CUDA 11.6

## 配置环境

# Name Version Build Channel

absl-py 2.1.0 pypi\_0 pypi

aiosignal 1.3.1 pypi\_0 pypi

albumentations 0.5.2 pypi\_0 pypi

anyio 4.2.0 pypi\_0 pypi

attrs 23.2.0 pypi\_0 pypi

audioread 3.0.1 py38haa244fe\_1 conda-forge

blas 1.0 mkl

brotli 1.0.9 h2bbff1b\_7

brotli-bin 1.0.9 h2bbff1b\_7

brotli-python 1.0.9 py38hd77b12b\_7

ca-certificates 2023.12.12 haa95532\_0

cachetools 5.3.2 pypi\_0 pypi

certifi 2023.11.17 py38haa95532\_0

cffi 1.16.0 py38h2bbff1b\_0

charset-normalizer 2.0.4 pyhd3eb1b0\_0

click 8.1.7 pypi\_0 pypi

cog 0.9.4 pypi\_0 pypi

colorama 0.4.6 py38haa95532\_0

colorlog 6.7.0 py38haa244fe\_2 conda-forge

contourpy 1.0.5 py38h59b6b97\_0

cryptography 41.0.7 py38h89fc84f\_0

cuda 11.6.1 0 nvidia

cuda-cccl 11.6.55 0 nvidia

cuda-command-line-tools 11.6.2 0 nvidia

cuda-compiler 11.6.2 0 nvidia

cuda-cudart 11.6.55 0 nvidia

cuda-cudart-dev 11.6.55 0 nvidia

cuda-cuobjdump 11.6.124 0 nvidia

cuda-cupti 11.6.124 0 nvidia

cuda-cuxxfilt 11.6.124 0 nvidia

cuda-libraries 11.6.1 0 nvidia

cuda-libraries-dev 11.6.1 0 nvidia

cuda-memcheck 11.8.86 0 nvidia

cuda-nsight-compute 12.3.2 0 nvidia

cuda-nvcc 11.6.124 0 nvidia

cuda-nvdisasm 12.3.101 0 nvidia

cuda-nvml-dev 11.6.55 0 nvidia

cuda-nvprof 12.3.101 0 nvidia

cuda-nvprune 11.6.124 0 nvidia

cuda-nvrtc 11.6.124 0 nvidia

cuda-nvrtc-dev 11.6.124 0 nvidia

cuda-nvtx 11.6.124 0 nvidia

cuda-nvvp 12.3.101 0 nvidia

cuda-runtime 11.6.1 0 nvidia

cuda-sanitizer-api 12.3.101 0 nvidia

cuda-toolkit 11.6.1 0 nvidia

cuda-tools 11.6.1 0 nvidia

cuda-visual-tools 11.6.1 0 nvidia

cycler 0.11.0 pyhd3eb1b0\_0

decorator 5.1.1 pyhd3eb1b0\_0

dominate 2.9.1 pypi\_0 pypi

exceptiongroup 1.2.0 pypi\_0 pypi

fastapi 0.98.0 pypi\_0 pypi

filelock 3.13.1 pypi\_0 pypi

fonttools 4.25.0 pyhd3eb1b0\_0

freetype 2.12.1 ha860e81\_0

frozenlist 1.4.1 pypi\_0 pypi

giflib 5.2.1 h8cc25b3\_3

google-auth 2.26.2 pypi\_0 pypi

google-auth-oauthlib 0.4.6 pypi\_0 pypi

grpcio 1.60.0 pypi\_0 pypi

h11 0.14.0 pypi\_0 pypi

h5py 3.10.0 pypi\_0 pypi

httptools 0.6.1 pypi\_0 pypi

icc\_rt 2022.1.0 h6049295\_2

idna 3.4 py38haa95532\_0

imageio 2.33.1 pypi\_0 pypi

imgaug 0.4.0 pypi\_0 pypi

importlib-metadata 7.0.1 py38haa95532\_0

importlib\_metadata 7.0.1 hd3eb1b0\_0

importlib\_resources 6.1.1 py38haa95532\_1

intel-openmp 2023.1.0 h59b6b97\_46320

jinja2 3.1.3 pypi\_0 pypi

joblib 1.2.0 py38haa95532\_0

jpeg 9e h2bbff1b\_1

jsonschema 4.21.0 pypi\_0 pypi

jsonschema-specifications 2023.12.1 pypi\_0 pypi

kiwisolver 1.4.4 py38hd77b12b\_0

lame 3.100 hcfcfb64\_1003 conda-forge

lazy\_loader 0.3 py38haa95532\_0

lerc 3.0 hd77b12b\_0

libbrotlicommon 1.0.9 h2bbff1b\_7

libbrotlidec 1.0.9 h2bbff1b\_7

libbrotlienc 1.0.9 h2bbff1b\_7

libcublas 11.9.2.110 0 nvidia

libcublas-dev 11.9.2.110 0 nvidia

libcufft 10.7.1.112 0 nvidia

libcufft-dev 10.7.1.112 0 nvidia

libcurand 10.3.4.107 0 nvidia

libcurand-dev 10.3.4.107 0 nvidia

libcusolver 11.3.4.124 0 nvidia

libcusolver-dev 11.3.4.124 0 nvidia

libcusparse 11.7.2.124 0 nvidia

libcusparse-dev 11.7.2.124 0 nvidia

libdeflate 1.17 h2bbff1b\_1

libffi 3.4.4 hd77b12b\_0

libflac 1.4.3 h63175ca\_0 conda-forge

libnpp 11.6.3.124 0 nvidia

libnpp-dev 11.6.3.124 0 nvidia

libnvjpeg 11.6.2.124 0 nvidia

libnvjpeg-dev 11.6.2.124 0 nvidia

libogg 1.3.5 h2bbff1b\_1

libopus 1.3.1 h8ffe710\_1 conda-forge

libpng 1.6.39 h8cc25b3\_0

librosa 0.7.0 pypi\_0 pypi

libsndfile 1.2.2 h81429f1\_1 conda-forge

libtiff 4.5.1 hd77b12b\_0

libuv 1.44.2 h2bbff1b\_0

libvorbis 1.3.7 he774522\_0

libwebp 1.3.2 hbc33d0d\_0

libwebp-base 1.3.2 h2bbff1b\_0

llvmlite 0.31.0 pypi\_0 pypi

lz4-c 1.9.4 h2bbff1b\_0

markdown 3.5.2 pypi\_0 pypi

markupsafe 2.1.4 pypi\_0 pypi

matplotlib-base 3.7.2 py38h4ed8f06\_0

mkl 2023.1.0 h6b88ed4\_46358

mkl-service 2.4.0 py38h2bbff1b\_1

mkl\_fft 1.3.8 py38h2bbff1b\_0

mkl\_random 1.2.4 py38h59b6b97\_0

mpg123 1.32.4 h63175ca\_0 conda-forge

mpmath 1.3.0 pypi\_0 pypi

msgpack-python 1.0.3 py38h59b6b97\_0

munkres 1.1.4 py\_0

networkx 3.1 pypi\_0 pypi

nsight-compute 2023.3.1.1 0 nvidia

numba 0.48.0 pypi\_0 pypi

numpy 1.20.3 pypi\_0 pypi

oauthlib 3.2.2 pypi\_0 pypi

opencv-python 3.4.9.33 pypi\_0 pypi

opencv-python-headless 4.9.0.80 pypi\_0 pypi

openjpeg 2.4.0 h4fc8c34\_0

openssl 3.0.12 h2bbff1b\_0

packaging 23.1 py38haa95532\_0

pandas 1.3.4 pypi\_0 pypi

pillow 10.0.1 py38h045eedc\_0

pip 23.3.1 py38haa95532\_0

pkgutil-resolve-name 1.3.10 pypi\_0 pypi

platformdirs 3.10.0 py38haa95532\_0

pooch 1.7.0 py38haa95532\_0

protobuf 3.19.0 pypi\_0 pypi

pyasn1 0.5.1 pypi\_0 pypi

pyasn1-modules 0.3.0 pypi\_0 pypi

pycparser 2.21 pyhd3eb1b0\_0

pydantic 1.10.14 pypi\_0 pypi

pyopenssl 23.2.0 py38haa95532\_0

pyparsing 3.0.9 py38haa95532\_0

pysocks 1.7.1 py38haa95532\_0

pysoundfile 0.12.1 pyhd8ed1ab\_0 conda-forge

python 3.8.18 h1aa4202\_0

python-dateutil 2.8.2 pyhd3eb1b0\_0

python-dotenv 1.0.1 pypi\_0 pypi

python-speech-features 0.6 pypi\_0 pypi

python\_abi 3.8 2\_cp38 conda-forge

pytorch 1.13.1 py3.8\_cuda11.6\_cudnn8\_0 pytorch

pytorch-cuda 11.6 h867d48c\_1 pytorch

pytorch-mutex 1.0 cuda pytorch

pytz 2023.3.post1 pypi\_0 pypi

pywavelets 1.4.1 pypi\_0 pypi

pyyaml 6.0.1 pypi\_0 pypi

ray 2.6.3 pypi\_0 pypi

referencing 0.32.1 pypi\_0 pypi

requests 2.31.0 py38haa95532\_0

requests-oauthlib 1.3.1 pypi\_0 pypi

resampy 0.3.1 pypi\_0 pypi

rpds-py 0.17.1 pypi\_0 pypi

rsa 4.9 pypi\_0 pypi

scikit-image 0.16.2 pypi\_0 pypi

scikit-learn 1.3.0 py38h4ed8f06\_1

scipy 1.10.1 py38hdcfc7df\_1

setuptools 68.2.2 py38haa95532\_0

shapely 2.0.2 pypi\_0 pypi

six 1.16.0 pyhd3eb1b0\_1

sniffio 1.3.0 pypi\_0 pypi

soxr 0.1.3 hcfcfb64\_3 conda-forge

soxr-python 0.3.7 py38he7056a7\_0 conda-forge

sqlite 3.41.2 h2bbff1b\_0

starlette 0.27.0 pypi\_0 pypi

structlog 24.1.0 pypi\_0 pypi

sympy 1.12 pypi\_0 pypi

tbb 2021.8.0 h59b6b97\_0

tensorboard 2.7.0 pypi\_0 pypi

tensorboard-data-server 0.6.1 pypi\_0 pypi

tensorboard-plugin-wit 1.8.1 pypi\_0 pypi

texttable 1.7.0 pypi\_0 pypi

threadpoolctl 2.2.0 pyh0d69192\_0

tk 8.6.12 h2bbff1b\_0

torchaudio 0.13.1 pypi\_0 pypi

torchvision 0.14.1 pypi\_0 pypi

tqdm 4.66.1 pypi\_0 pypi

typing\_extensions 4.9.0 py38haa95532\_1

tzdata 2023.4 pypi\_0 pypi

ucrt 10.0.20348.0 haa95532\_0

urllib3 1.26.18 py38haa95532\_0

uvicorn 0.27.0.post1 pypi\_0 pypi

vc 14.2 h21ff451\_1

vc14\_runtime 14.38.33130 h82b7239\_18 conda-forge

vs2015\_runtime 14.38.33130 hcb4865c\_18 conda-forge

watchfiles 0.21.0 pypi\_0 pypi

websockets 12.0 pypi\_0 pypi

werkzeug 3.0.1 pypi\_0 pypi

wheel 0.41.2 py38haa95532\_0

win\_inet\_pton 1.1.0 py38haa95532\_0

xz 5.4.5 h8cc25b3\_0

zipp 3.17.0 py38haa95532\_0

zlib 1.2.13 h8cc25b3\_0

zstd 1.5.5 hd43e919\_0

## 环境说明

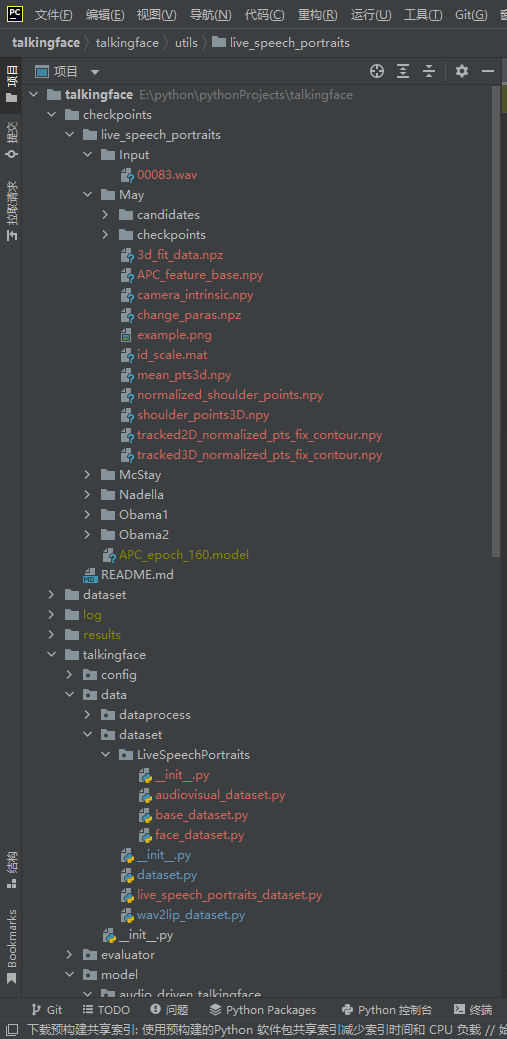
librosa==0.7.0。

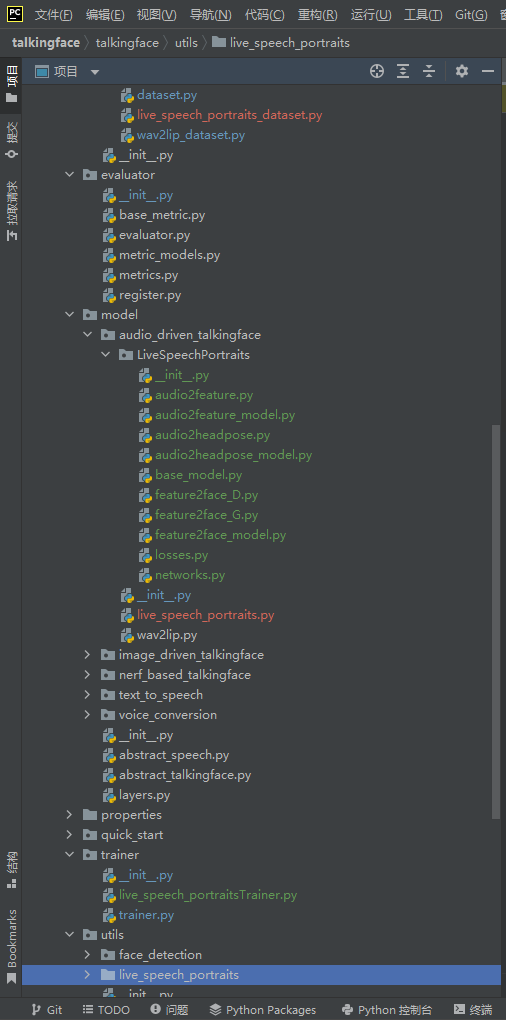
因为这是论文作者要求的，0.10.1的librosa将其中的mel()函数里的位置参数改为了关键字参数，如果不更改会报如下错误：

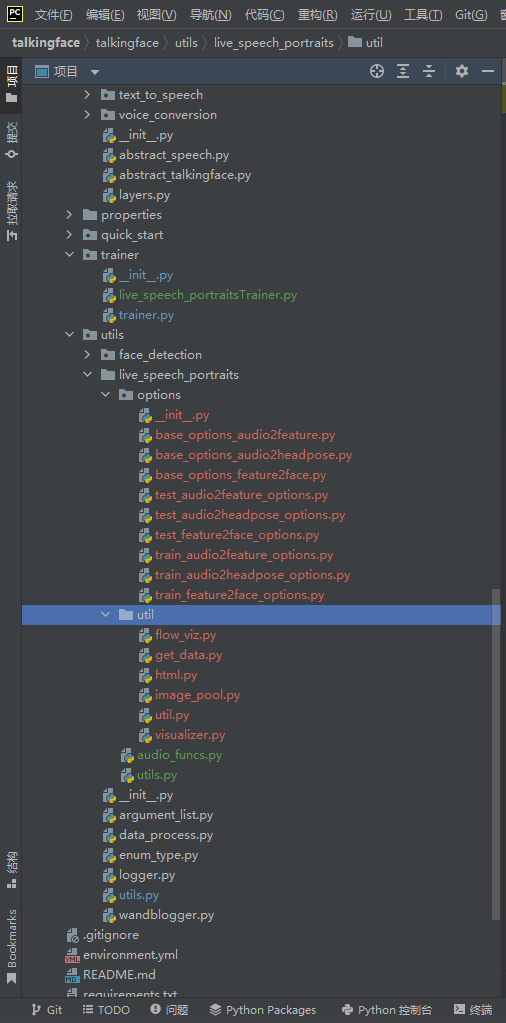
TypeError: mel() takes 0 positional arguments but 5 were given.

## 文件结构

文件结构如图所示：







## 验证

### 说明

论文作者并没有提供训练代码和对应的数据集，据作者所言此论文为他在实习期间完成，其训练代码及其相关内容均属于其公司的商业机密，不能对外分享，但是提供了生成的模型以及部分代码以供验证。

### 验证过程

运行run\_talkingface.py文件，参数为--model=live\_speech\_portraits --dataset=live\_speech\_portraits --evaluate\_model\_file=notNone

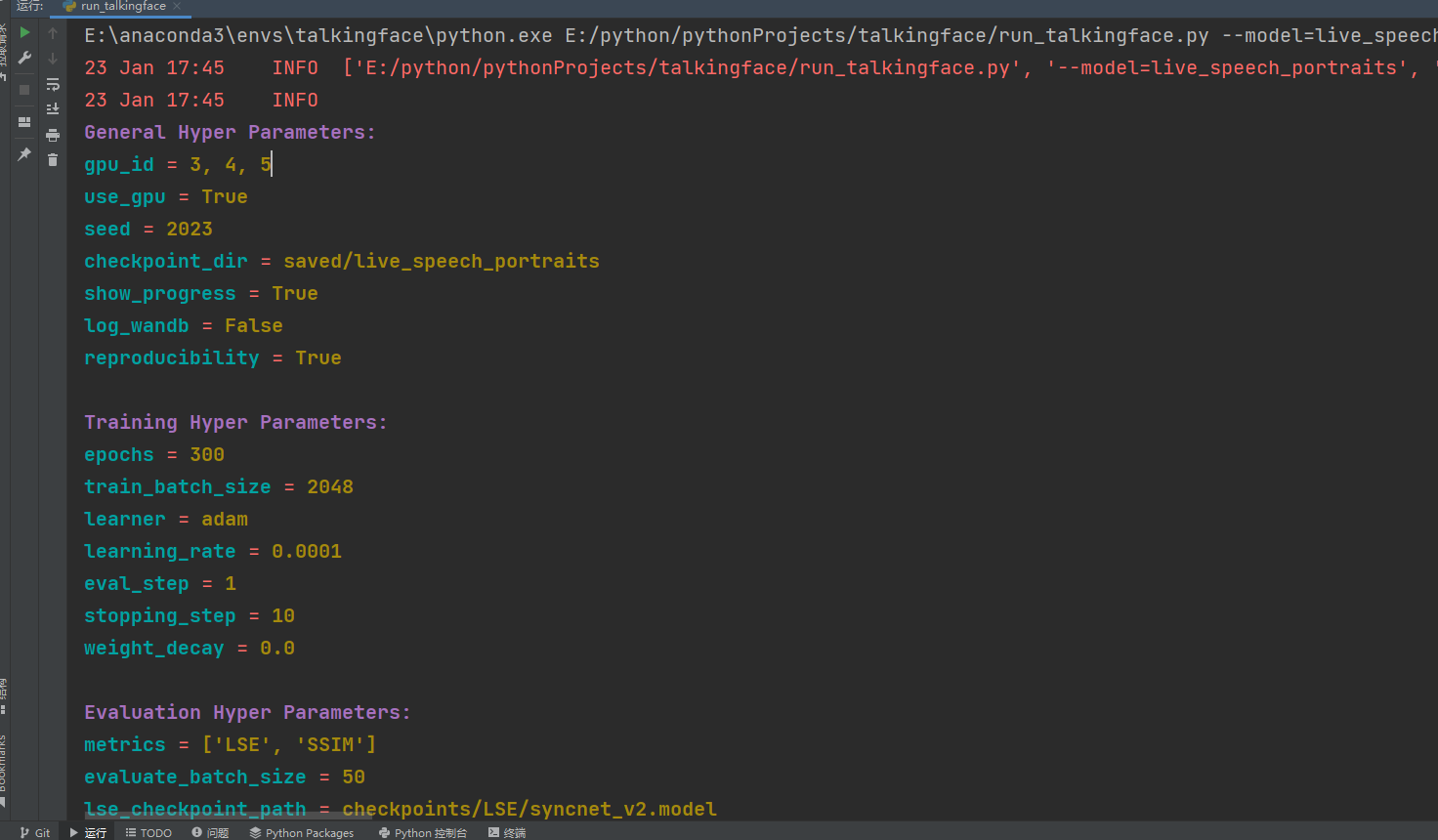
或者直接在run\_talkingface.py所在目录下通过命令行输入如下内容：

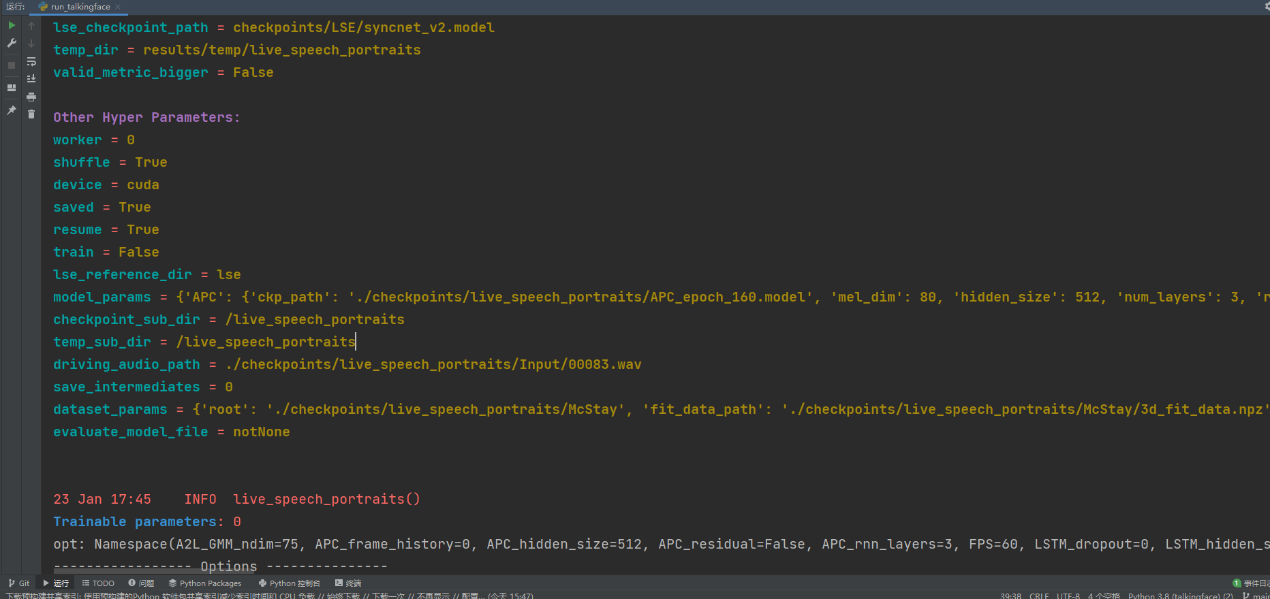
python run\_talkingface.py --model=live\_speech\_portraits --dataset=live\_speech\_portraits --evaluate\_model\_file=notNone

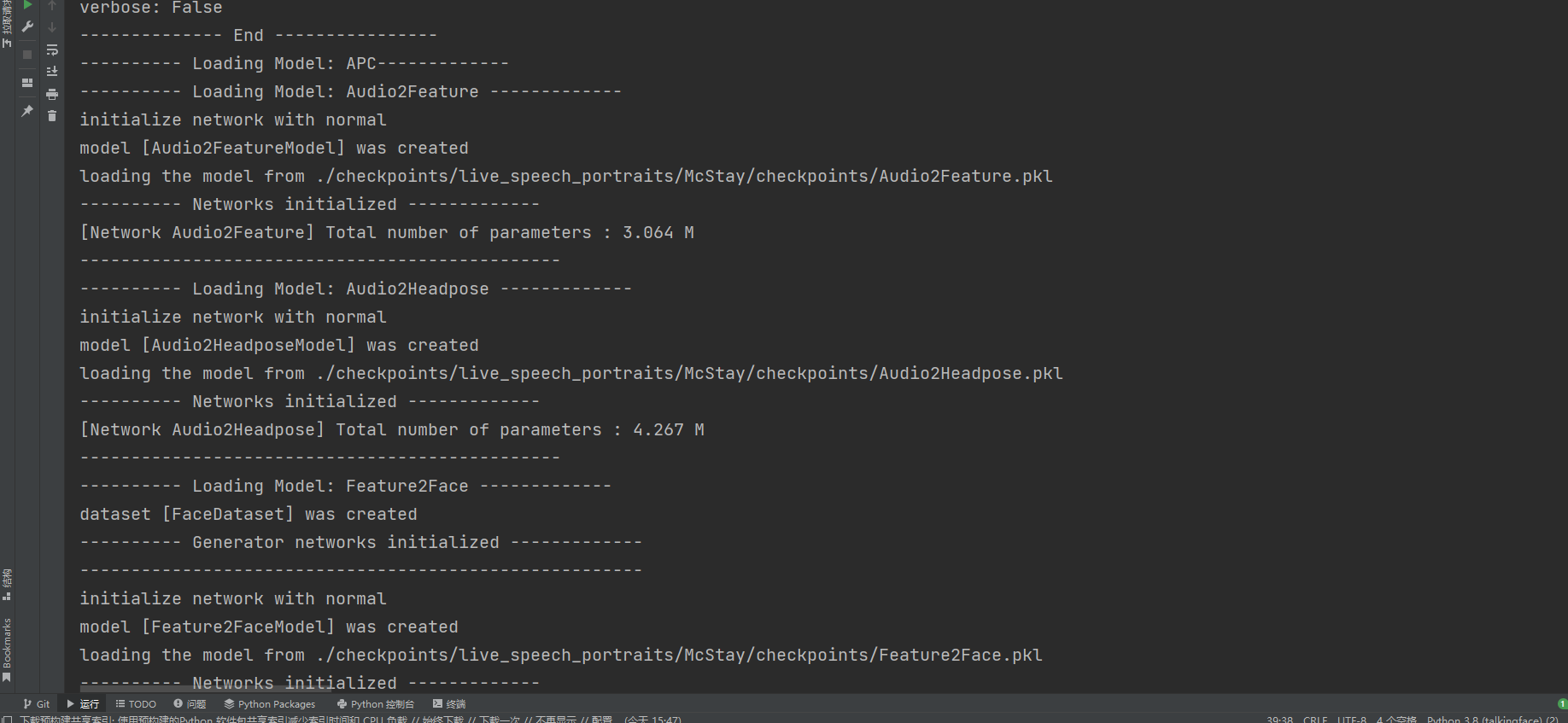
这里的evaluate\_model\_file无实际意义，但因为无训练过程，所以此处是必须的。

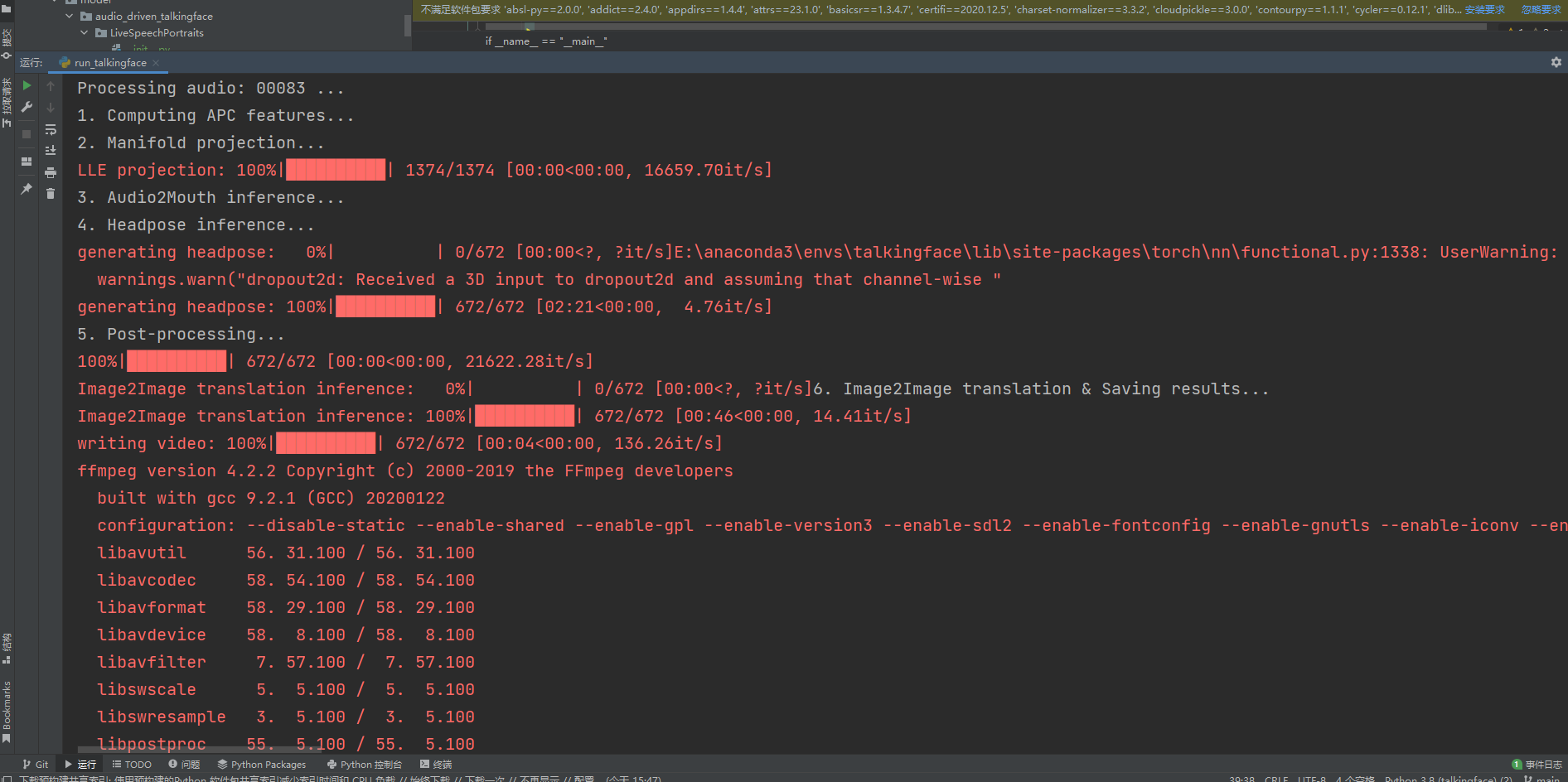
验证所需的参数均在./talkingface/properties/ live\_speech\_portraits.yaml中设定。

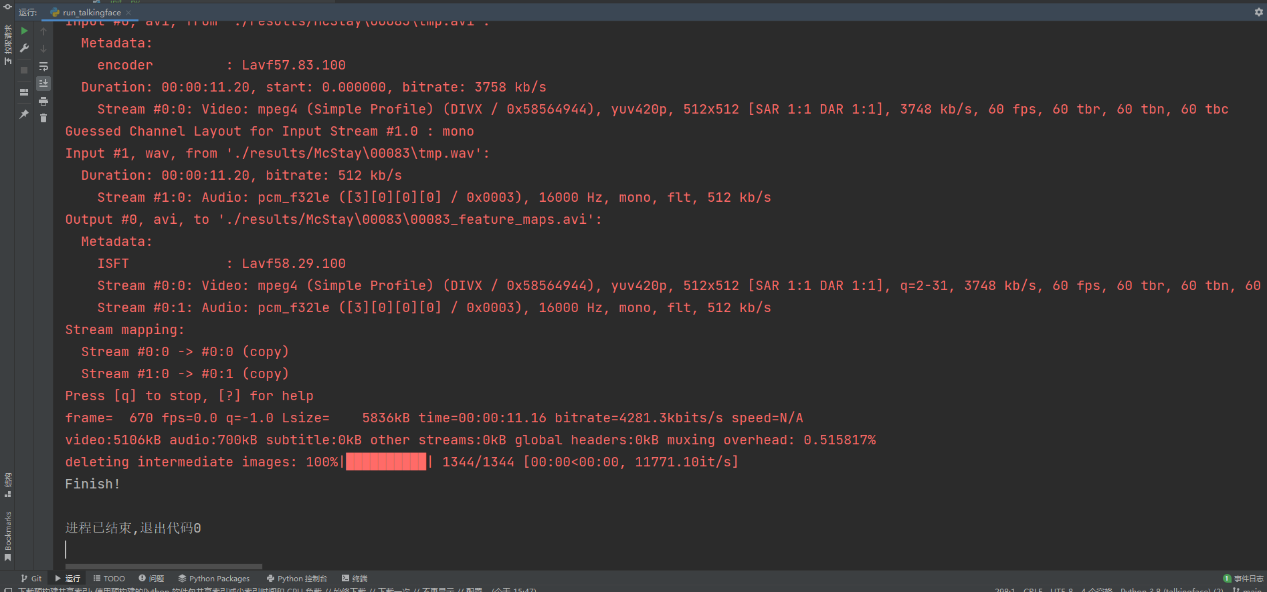
验证过程如下所示：





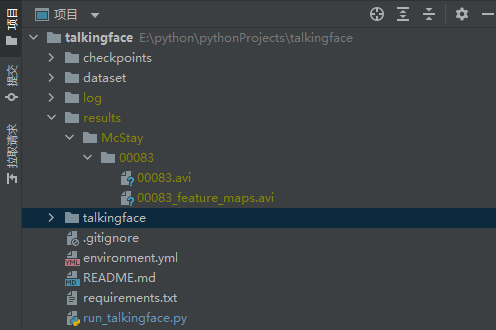






## 结果

结果在/results下，如图所示



## 演示结果

如图所示，详情请看演示视频。

