

· **Select Input (选择输入)**
点这里选择你想要处理的文件

· **Select Output (选择输出)**
点这里选择处理后文件的输出
路径

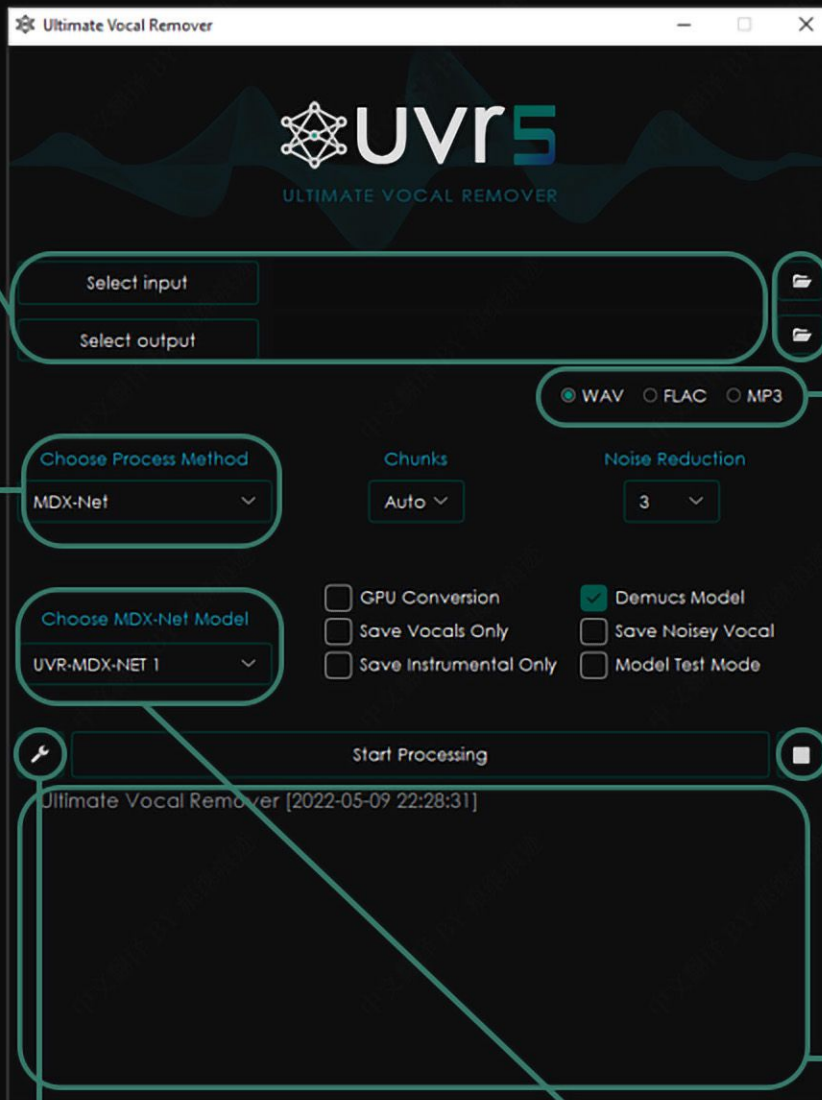
在这里，你可以选择不同的AI算法
模型来处理音频。有四个选项：

· **VR Architecture**
这些模型使用幅度谱图进行信源分离

· **MDX-Net**
这些模型使用混合频谱/波形进行源分离

· **Demucs v3**
这些模型使用混合频谱/波形进行源分离

· **Ensemble Mode (合奏模式)**
在这里，你可以从多个算法模型中
获得最佳结果。



· **打开输入文件夹**
打开输入文件所在的文件夹

· **打开输出文件夹**
打开输出文件所在的文件夹

· **保存的格式**
选择输出文件的格式，有 WAV、FLAC
或者 MP3

· **停止按钮**
停止任何正在运行的进程
会有一个弹出窗口来询问是否取消

· **进度控制台**
在这里你能看到程序处理时的一些
信息

· **设置按钮**
点这里打开设置菜单，可以访问其他
高级设置和帮助指南

每个处理算法都有自己的一组选项和模型
在这里，你可以选择与所选算法相关的模型



所提供的每个模型都经过微调，参数略有不同

以下几个模型是开箱即用的：

· **1_HP-UVR.pth**

这是一个针对器乐（伴奏）的模型

· **2_HP-UVR.pth**

这是 1_HP-UVR.pth 模型的一个微调版

· **3_HP-Vocal-UVR.pth**

这个模型强化了人声的提取。人声干会很干净，但伴奏的声音可能会很混浊

· **4_HP-Vocal-UVR.pth**

这个模型也强化了人声的提取，比前一个模型更激进

· **5_HP-Karaoke-UVR.pth**

这个模型在去除主要人声的同时保留了背景的人声

以下几个模型需要额外下载：

· **6_HP-Karaoke-UVR.pth**

这个模型在保留背景人声的同时去掉了主要人。

· **7_HP2-UVR.pth**

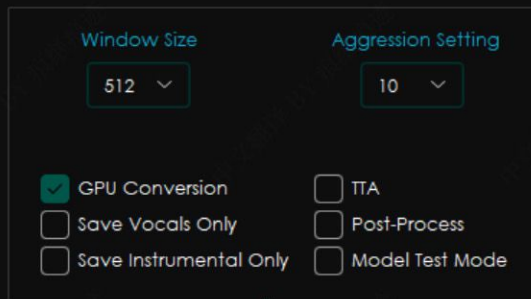
这是一个使用更多数据和新参数训练的强大的器乐模型

· **8_HP2-UVR.pth**

这是一个强大的器乐模型

· **9_HP2-UVR.pth**

这是 8_HP2-UVR.pth 模型的一个微调版



下拉菜单

· **Windows Size (窗口大小)**

Windows Size 越小，效果就越好。然而，较小的 Windows Size 意味着越长的转换时间和越大的资源占用

以下是可选择的 Windows Size 的值：

1024 - 低转换质量，最短的转换时间，低资源使用率

512 - 平均转换质量，平均转换时间，正常资源使用率

320 - 较好的转换质量，较长的转换时间，较高的资源使用率

· **Aggression Setting (力度设置)**

这个选项允许你设置去除声音的力度

范围是0-100

值越大，就会进行越深度的提取

器乐和人声模型的默认值是10

数值超过10会导致非人声模型的乐器声变得浑浊

复选框

· **GPU Conversion (GPU 转换)**

勾选此选项可以使用 GPU 来处理

注意：如果你没有一个兼容 Cuda 的 GPU，这个选项将无法正常运行

最好使用英伟达的 GPU

注意：CPU 的转换速度要比通过 GPU 处理的慢得多

· **Save Vocals Only (只保存人声)**

允许用户只保存人声部分

· **Save Instrumental Only (只保存器乐声)**

允许用户只保存伴奏部分

· **TTA**

此选项执行Test-Time-Augmentation（测试时间增强），以提高分离质量

注意：勾选此选项将增加转换的时间

· **Post-process (后处理)**

此选项也许能在人声输出中，识别出残留的乐器声。所以此选项可以改善某些歌曲的分离效果

注意：选择此选项可能会对转换过程产生不利影响，这取决于音频的情况

· **Model Test Mode (模型测试模式)**

此选项使用户更容易测试不同模型和模型组合的结果，因为它避免了用户为了在，通过多个模型处理同一曲目时，手动改变文件名和创建新文件夹的麻烦

伴奏和人声输出的文件名将包含所选模型的名称



有五个模型可供选择

所提供的每个模型，都是在稍微不同的参数，上进行微调的。不过带编号的模型，是按照 AI Crowds 官方测试集的 SDR 得分顺序排列的。

· UVR-MDX-NET Main

这是最强的模型，但占用的电脑资源也最多

· UVR-MDX-NET 1

这个模型的 SDR 得分为 9.703

· UVR-MDX-NET 2

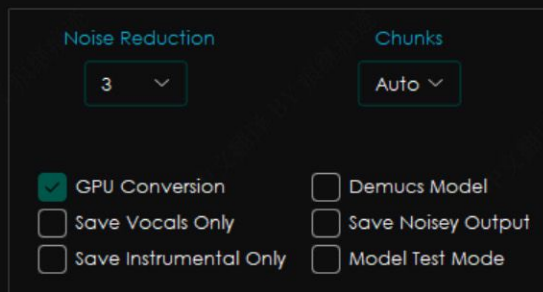
这个模型的 SDR 得分为 9.682

· UVR-MDX-NET 3

这个模型的 SDR 得分为 9.662

· UVR-MDX-NET Karaoke

这个模型删除了主要的人声，而保留了背景人声



下拉菜单

· Chunks (块)

这个选项允许用户减少（或增加）RAM 或 V-RAM 的使用

较小的块大小使用更少的 RAM 或 V-RAM，但也可能增加处理时间

较大的块大小使用更多的 RAM 或 V-RAM，但也可以减少处理时间

选择“Auto”会根据你的系统有多少 RAM 或 V-RAM 来计算一个合适的块大小

选择“Full”将把音频作为一个完整的块来处理

“Full”选项只推荐给那些电脑性能比较强的人

默认选择是“Auto”

· Noise Reduction (降噪)

这个选项将减少或消除由模型产生的任何噪音

灵敏度的设置区间为 0-10

默认设置为3

选择“None”将完全关闭降噪

复选框

· GPU Conversion (GPU 转换)

勾选此选项可以使用 GPU 来处理

注意：如果你没有一个兼容 Cuda 的 GPU，这个选项将无法正常运行

最好使用英伟达的 GPU

注意：CPU 的转换速度要比通过 GPU 处理的慢得多

· Save Vocals Only (只保存人声)

允许用户只保存人声部分

· Save Instrumental Only (只保存器乐声)

允许用户只保存伴奏部分

· Demucs Model (Demucs模型)

在 MDX-Net 模型之外，通过标准的Demucs模型处理音频，以提升分离效果

该选项会占用更多的电脑资源，并增加处理时间

· Model Test Mode (模型测试模式)

这个选项使用户更容易测试不同模型的结果，因为在通过多个模型处理同一音频时，可以省去手动改变文件名的麻烦。（详见“User Ensemble”选项卡）

伴奏和人声输出的文件名将包含所选模型的名称

· Save Noisy Output (保存噪音输出)

允许用户在不应用降噪的情况下保存额外的噪音文件

Choose Stem(s) Segment

All Stems None

GPU Conversion Shifts

Stem Only 2

Mix Without Stem Only Overlap

Split Mode 0.25

Choose Process Method

Demucs v3

Choose Demucs Model

mdx_extra

- UVR_Demucs_Model_1
- UVR_Demucs_Model_2
- UVR_Demucs_Model_Bag
- mdx_extra
- mdx_extra_q

以下模型开箱即用

· 2音轨模型

UVR_Demucs_Model_1 - 这个模型是在 UVR 数据集上训练出来的。它只能够提取2个音轨，“Vocals（人声）”或“Other（其它）”，当选择这个模型时，“Other”被保存为“Instrumental”

UVR_Demucs_Model_2 - 这个模型也是在 UVR 数据集上训练出来的。它只能够提取2个音轨，“Vocals（人声）”或“Other（其它）”，当选择这个模型时，“Other”被保存为“Instrumental”

UVR_Demucs_Model_Bag - 此项选择将模型“VR_Demucs_Model_1”和“VR_De-mucs_Model_1”打包在一起运行

· 4音轨模型

mdx_extra - 这是一个强大的4音轨模型，最初由 adefossez 和 Demucs 团队训练

mdx_extra_q - 这是由 adefossez 和 Demucs 团队最初训练的量化4音轨模型，占用的资源较少

下拉菜单

· Choose Stem(s) (选择音轨)

在这里，你可以选择使用选定的模型提取哪个音轨
音轨的选择

- 所有音轨 - 保存模型能够提取的所有音轨
- 人声 - 只提取人声音轨
- 其他 - 只提取其它音轨
- 贝斯 - 只提取贝斯音轨
- 鼓 - 只提取鼓的音轨

· Segments (分块)

这个选项允许用户减少（或增加）RAM 或 V-RAM 的使用。

- 较小的块大小使用更少的 RAM 或 V-RAM，但也可能增加处理时间
- 较大的块大小使用更多的 RAM 或 V-RAM，但也可以减少处理时间

· Shifts (偏移)

使用输入的随机移位执行多次预测并将它们平均。移位次数越多，预测所需的时间越长。除非你有 GPU，否则不推荐启用此选项

· Overlaps (重叠)

这个选项控制预测窗口之间的重叠量（对于 Demucs 模型来说，一个窗口是10秒）

复选框

· GPU Conversion (GPU 转换)

勾选此选项可以使用 GPU 来处理

- 注意：如果你没有一个兼容 Cuda 的 GPU，这个选项将无法正常运行
- 最好使用英伟达的 GPU
- 注意：CPU 的转换速度要比通过 GPU 处理的慢得多

· Stem Only (只保存音轨)

允许用户只保存所选的音轨

· Mix Without Stem Only (不含所选音轨的混合)

允许用户保存不包含所选音轨的混合音频文件

· Split Mode (分割模式)

使用 Demucs v3 模型的原始分块方法。选择它将自动禁用“分块”。

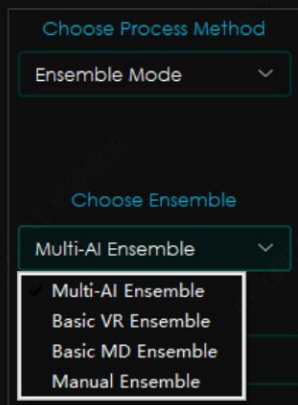
其它信息

· 添加额外的模型

你可以通过“设置”中的“Download Center”选项卡下载更多的模型，下载之后，应用程序将自动检测做下载的模型

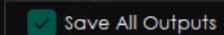
· 向下兼容

本程序也与 Demucs v1 和 v2 模型相兼容



该选项适用于以下合奏模式

- **Multi-AI Ensemble**
- **Basic VR Ensemble**
- **Basic MD Ensemble**



此选项将保存合奏模式中每个模型的所有单独的输出结果

当“Save All Outputs”（保存所有输出）未被选中时，应用程序将自动删除合奏中每个模型产生的所有单独的输出结果

以下合奏模式是开箱即用的

· **Muli-AI Ensemble**

通过以下模型处理音频文件，并使用默认参数对输出进行合奏

UVR_MDXNET1
2_HP-UVR.pth

这个合奏模式是可以自定义的，你最多可以为这个合奏选择6个模型，详情见下文

· **Basic VR Ensemble**

通过以下模型处理音频文件，并使用默认参数对输出进行合奏

1_HP-UVR.pth
2HP-UVR.pth

这个合奏模式是可以自定义的，你最多可以为这个合奏选择5个模型，详情见下文

· **Basic MD Ensemble**

通过以下模型处理音频文件，并使用默认参数对输出进行合奏

UVR MDXNET Main
UVR_MDXNET_1

这个合奏模式是可以自定义的，你最多可以为这个合奏选择5个模型，详情见下文
这个合奏模式只适用于 MDX-Net 和 Demucs 模型

· **Manual Ensemble（手动合奏）**

此选项允许用户从不同模型中选择输出方式，并进行手动合奏，详情见下文

参见“Manual Ensemble Options”（手动合奏选项）选项卡，以了解更多关于此选项的信息

自定义合奏

要自定义合奏，请点击设置按钮并选择“Ensemble Customization Options”（合奏自定义选项）按钮（或者你也可以可以点击“Choose Ensemble”（选择合奏）标签进入自定义菜单）

在你继续之前，请注意以下几点：

对于编号为6-16的 VR 模型（VR Architecture）以及 V4 模型来说，它们都是需要你取手动下载的

合奏模式将通过多个模型处理音轨，并组合生成的输出以实现更稳健的分离

所产生的文件将在输出时附加上合奏的名称

注意：每个合奏模式都有特定的选项。请阅读旁边的章节，了解更多细节

Select input: /Users/ /Desktop/伴奏提取/output/MDX_毛不易 - 借_(Ins)

Select output: C:/Users/ /Desktop/伴奏提取/output

合奏至少需要输入两个音频文件
(由不同模型提取的伴奏或人声)

Choose Process Method: Ensemble Mode

Choose Algorithm: Instrumentals (Min Spec), Vocals (Max Spec), **Instrumentals (Min Spec)**

Choose Ensemble: Manual Ensemble

下拉菜单

· 选择算法 - 你可以选择以下几种算法

Instrumentals (Min Spec) (伴奏 (最小规格))

这种算法将从你输入的文件中加载频谱图，并计算每个输入文件的最小规格值，生成的文件将通过文件中的每个模型删除所有声音数据

以下内容将被附加到输出文件中:

"_User_Ensembled_(Min Spec).wav"

Vocals (Max Spec) (人声 (最大规格))

这种算法将从你输入的文件中加载频谱图，并计算每个输入文件的最大规格值，生成的文件将包含模型中提取的所有人声数据

以下内容将被附加到输出文件中:

"_User_Ensembled_(Max Spec).wav"

示例

以“Instrumentals (Min Spec)”选项为例

1、选择由不同模型提取的伴奏

名称

- MDX_毛不易 - 借_(Instrumental).wav
- MDX_毛不易 - 借_(Vocals).wav
- VR_毛不易 - 借_(Instrumental).wav
- VR_毛不易 - 借_(Vocals).wav

2、开始处理

Select input: /Users/ /Desktop/伴奏提取/output/MDX_毛不易 - 借_(Ins)

Select output: C:/Users/ /Desktop/伴奏提取/output

Format: WAV, FLAC, MP3

Choose Process Method: Ensemble Mode

Choose Algorithm: Instrumentals (Min Spec)

Choose Ensemble: Manual Ensemble

Start Processing

3、在输出文件夹里就能找到输出的文件

名称

- MDX_毛不易 - 借_(Instrumental).wav
- MDX_毛不易 - 借_(Instrumental)_Manual_Ensemble_(Min Spec).wav
- MDX_毛不易 - 借_(Vocals).wav
- VR_毛不易 - 借_(Instrumental).wav
- VR_毛不易 - 借_(Vocals).wav