

# Whisper是什麼

亞洲大學資訊發展處 林震宇

# OpenAI Whisper

OpenAI Whisper 是一款應用人工智慧技術的「語音轉文字」(speech-to-text)工具，它能幫我們將一份錄音檔完整轉換為逐字稿，也能用於自動生成影片字幕檔。



```
PowerShell: Anaconda Powershell Prompt
(base) PS D:\whisperdesktop\data> whisper micro-machines.wav --language English --model medium
100%|#####| 1.42G/1.42G [01:14<00:00, 20.5MiB/s]
[00:00.000 -> 00:03.200] This is the Micro Machine Man presenting the most midget, miniature, motorcade of Micro Machines.
[00:03.200 -> 00:06.600] Each one has dramatic details, terrific trims, precision paint jobs, plus incredible Micro Machine Pocket play sets.
[00:06.600 -> 00:08.700] There's a police station, fire station, restaurant, service station, and more.
[00:08.700 -> 00:10.200] Perfect pocket portables to take any place.
[00:10.200 -> 00:15.200] And there are many miniature play sets to play with, and each one comes with its own special edition Micro Machine vehicle and fun, fantastic features that miraculously move.
[00:15.200 -> 00:19.200] Raise the boat lift at the airport marina, man the gun turret at the army base, clean your car at the car wash, raise the toll bridge.
[00:19.200 -> 00:21.200] And these play sets fit together to form a Micro Machine world.
[00:21.200 -> 00:25.200] Micro Machine Pocket play sets, so tremendously tiny, so perfectly precise, so dazzlingly detailed, you'll want to pocket them all.
[00:25.200 -> 00:27.600] Micro Machines and Micro Machine Pocket play sets, sold separately from Galoob.
[00:27.600 -> 00:29.600] The smaller they are, the better they are.
(base) PS D:\whisperdesktop\data>
(base) PS D:\whisperdesktop\data>
```

# Whisper安裝教學

# 1. 下載WhisperDesktop

至WhisperDesktop的  
Github上下載最新版的  
WhisperDesktop

<https://github.com/Const-me/Whisper>

master 2 Branches 17 Tags

Go to file Code

Const-me PowerShell module version 306aadd · last year 204 Commits

ComLightLib	Enabled AVX1 for C++ projects	last year
ComputeShaders	Compute shader to test a tensor for NaN values	last year
Examples	Bugfix, models source URL	last year
SampleClips	Updated performance summary	last year
Tools	The DLL might now work on processors without F16C or AVX	last year
Whisper	Version 1.12	last year
WhisperNet	NuGet package readme and metadata	last year
WhisperPS	PowerShell module version	last year
.gitignore	The DLL might now work on processors without F16C or AVX	last year
LICENSE	Initial commit	2 years ago
Readme.md	Documentation	last year
WhisperCpp.sln	The DLL might now work on processors without F16C or AVX	last year
gui-capture.png	Updated screenshots	last year
gui-load-model.png	Updated screenshots	last year
gui-transcribe.png	Updated a screenshot	last year

About

High-performance GPGPU inference of OpenAI's Whisper automatic speech recognition (ASR) model

Readme

MPL-2.0 license

Activity

7.9k stars

85 watching

679 forks

Report repository

Releases 17

Version 1.12 (Latest) on Jul 22, 2023

+ 16 releases

點擊目前最新版本

Languages

C++ 68.2% C 21.1% C# 5.5% HLSL 4.9% Other 0.3%

# 1. 下載WhisperDesktop

選擇WhisperDesktop.zip  
壓縮檔並點擊下載，下  
載完後解壓縮

Releases / 1.12.0

## Version 1.12

Latest

Compare

Const-me released this Jul 22, 2023 · 1 commit to master since this release · 1.12.0 · c5515ac

Updated models source URL in documentation and GUI  
Reliability enhancement, microphone capture less likely to transition to "Stalled" state and discard the audio

### Assets 6

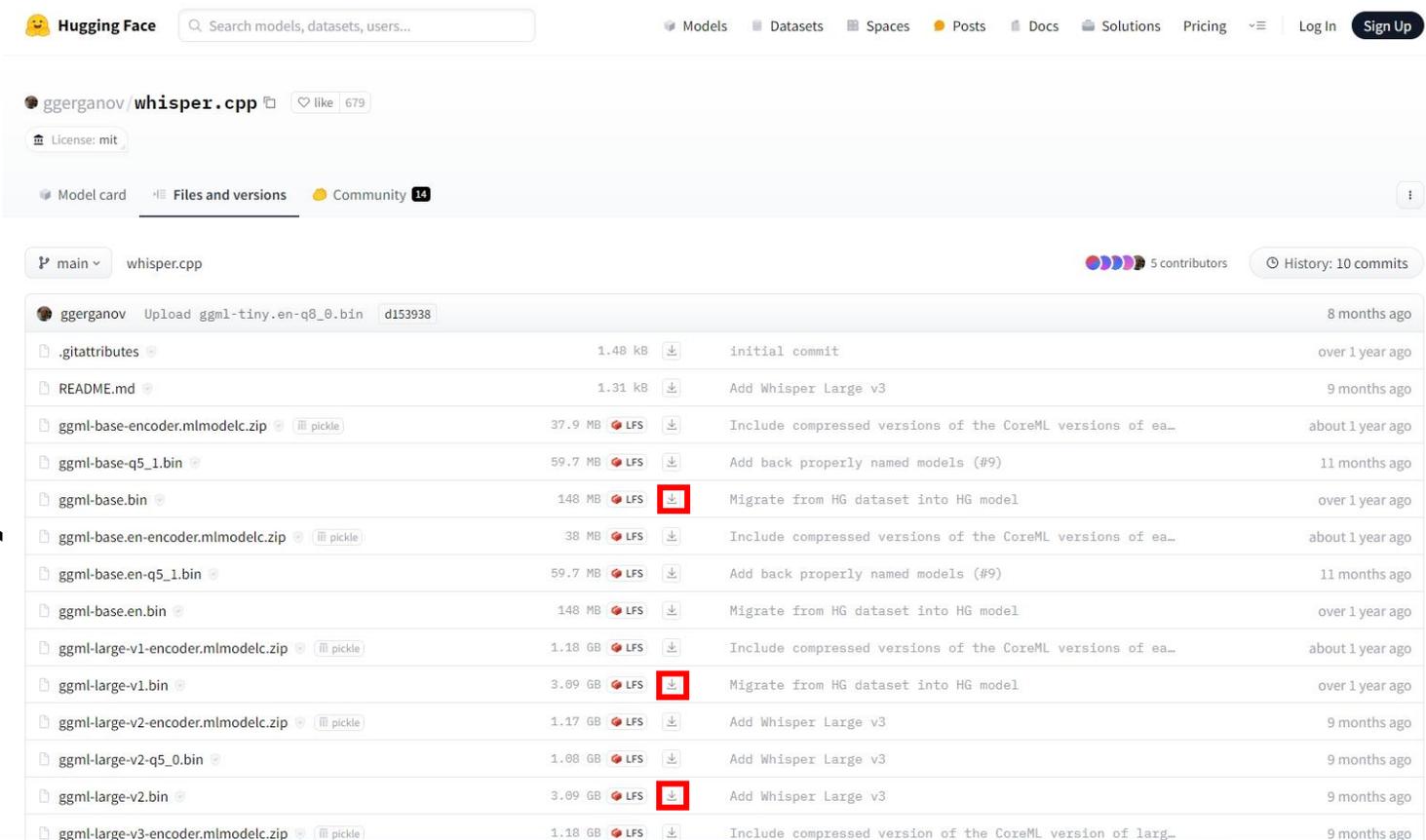
cli.zip	369 KB	Jul 22, 2023
Library.zip	286 KB	Jul 22, 2023
WhisperDesktop.zip	466 KB	Jul 22, 2023
WhisperPS.zip	406 KB	Jul 22, 2023
Source code (zip)		Jul 22, 2023
Source code (tar.gz)		Jul 22, 2023

35 🍌 1 🍌 4 🍌 32 🍌 2 🍌 1 64 people reacted

# 2. 下載語言模型

在使用whisper前最重要的要先下載語言模型，  
點擊連結進入：[連結](#)

根據模型大小依序分為  
tiny(微小)、base(基礎)、  
small(小)、medium(中)、  
large(大)、large-v1、  
large-v2、large-v3等。

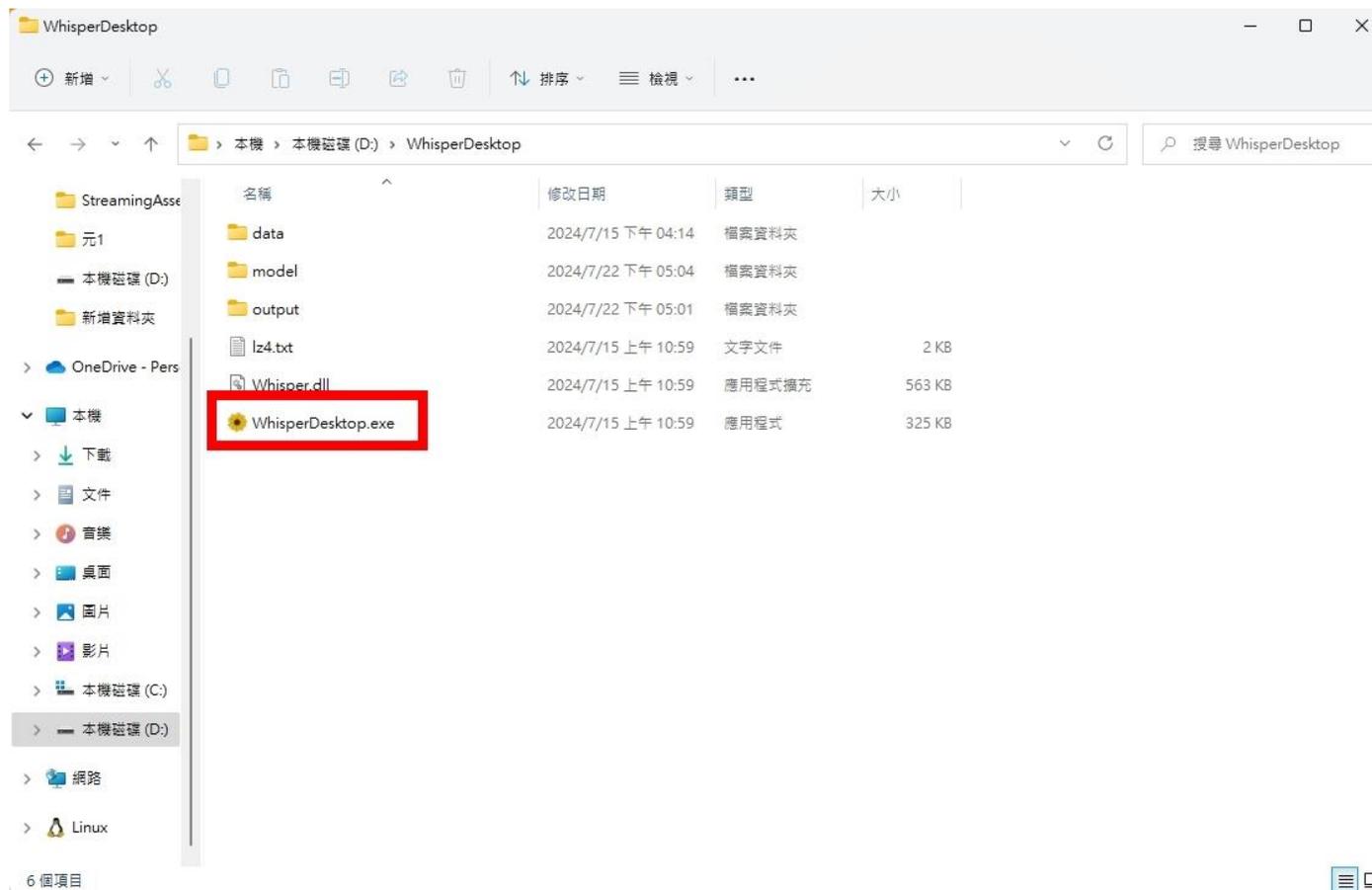


The screenshot shows the Hugging Face repository for ggerganov/whisper.cpp. The repository is licensed under MIT and has 679 likes. It contains a list of files and versions for download. The files are organized into a table with columns for filename, size, license, and commit message. The download button for each file is highlighted with a red box.

Filename	Size	License	Commit Message	Time
.gitattributes	1.48 kB		initial commit	over 1 year ago
README.md	1.31 kB		Add Whisper Large v3	9 months ago
ggml-base-encoder.mlmodelc.zip	37.9 MB	LFS	Include compressed versions of the CoreML versions of ea...	about 1 year ago
ggml-base-q5_1.bin	59.7 MB	LFS	Add back properly named models (#9)	11 months ago
ggml-base.bin	148 MB	LFS	Migrate from HG dataset into HG model	over 1 year ago
ggml-base-en-encoder.mlmodelc.zip	38 MB	LFS	Include compressed versions of the CoreML versions of ea...	about 1 year ago
ggml-base-en-q5_1.bin	59.7 MB	LFS	Add back properly named models (#9)	11 months ago
ggml-base.en.bin	148 MB	LFS	Migrate from HG dataset into HG model	over 1 year ago
ggml-large-v1-encoder.mlmodelc.zip	1.18 GB	LFS	Include compressed versions of the CoreML versions of ea...	about 1 year ago
ggml-large-v1.bin	3.09 GB	LFS	Migrate from HG dataset into HG model	over 1 year ago
ggml-large-v2-encoder.mlmodelc.zip	1.17 GB	LFS	Add Whisper Large v3	9 months ago
ggml-large-v2-q5_0.bin	1.08 GB	LFS	Add Whisper Large v3	9 months ago
ggml-large-v2.bin	3.09 GB	LFS	Add Whisper Large v3	9 months ago
ggml-large-v3-encoder.mlmodelc.zip	1.18 GB	LFS	Include compressed version of the CoreML version of larg...	9 months ago

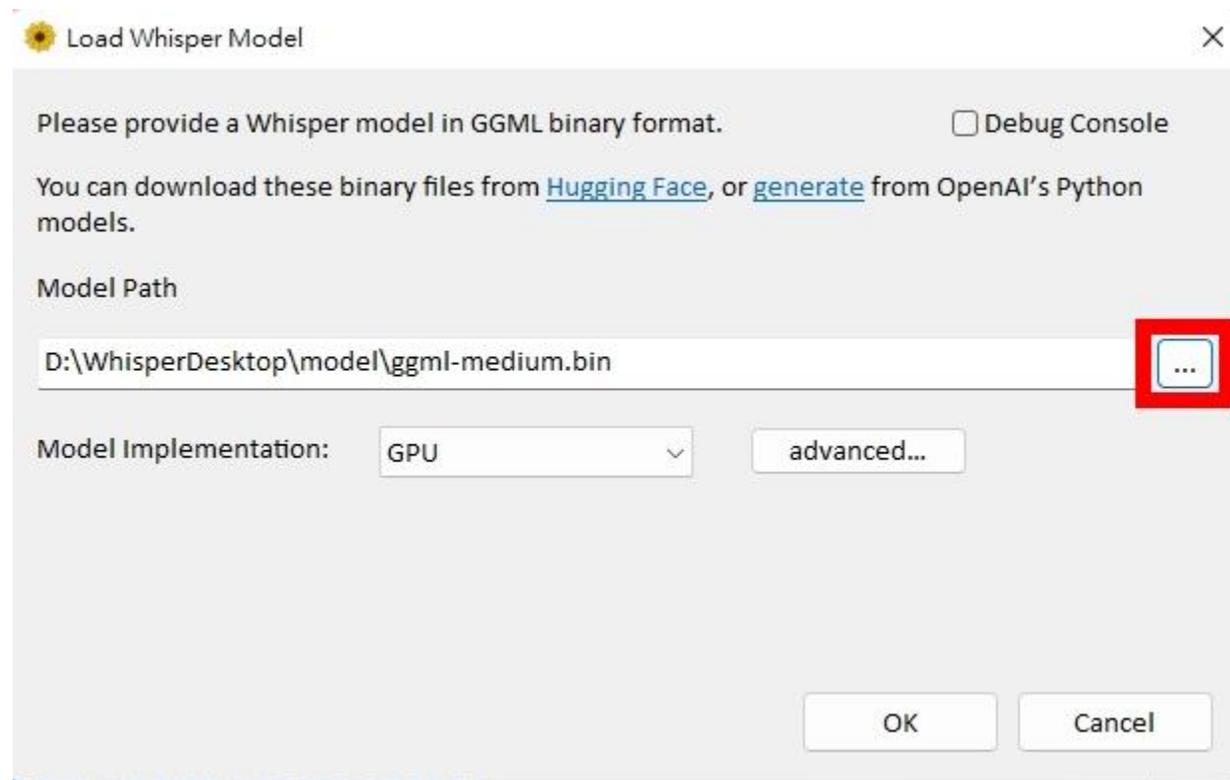
# 3. 開啟WhisperDesktop

解完壓縮後點擊  
WhisperDesktop.exe開啟  
主程式



# 4. 載入語言模型

載入剛剛下載的語言模型，並且選擇使用「GPU」來執行，接著按下「OK」鍵。



# 4.載入語言模型

「Language」根據欲辨識的影片來選擇語言，

「Transcribe File」選擇要辨識的檔案，

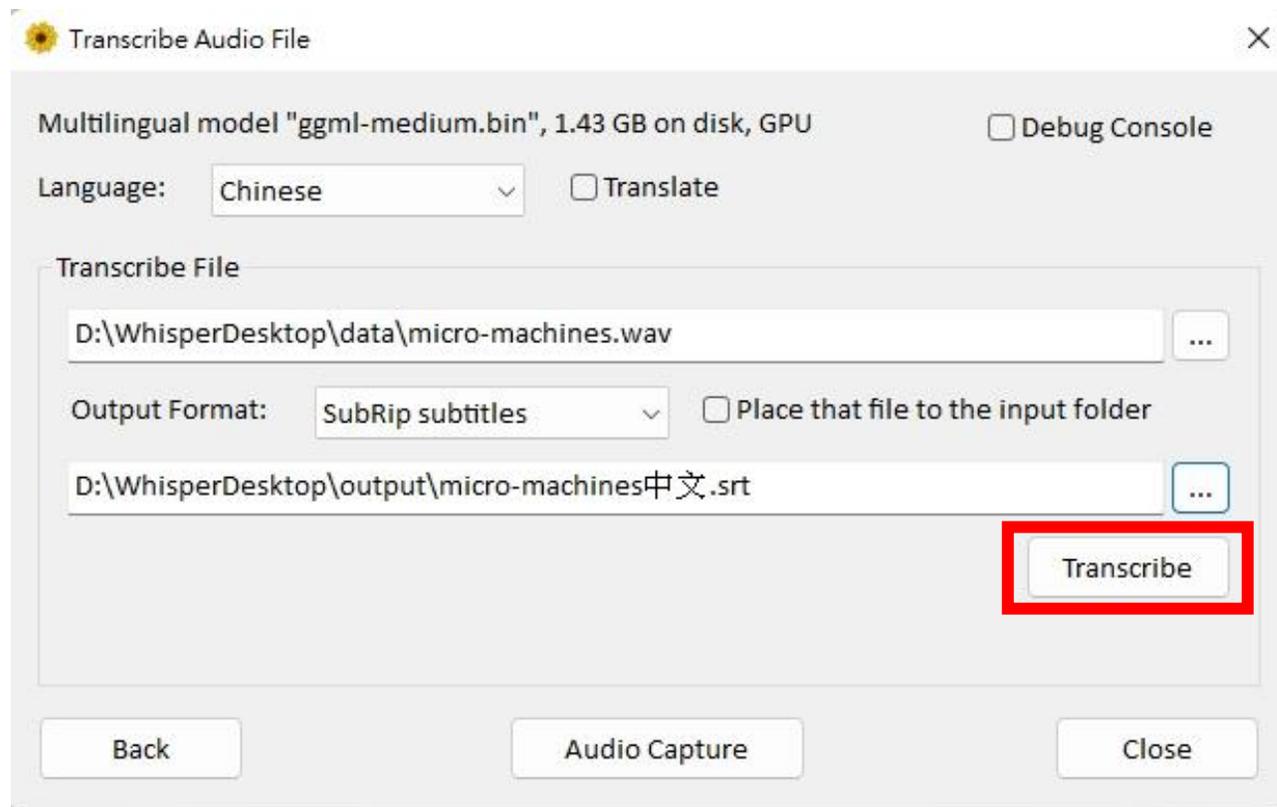
「Output Format」可選擇要輸出的檔案格式，

Text file 為純文字檔，

SubRip subtitles 為字幕

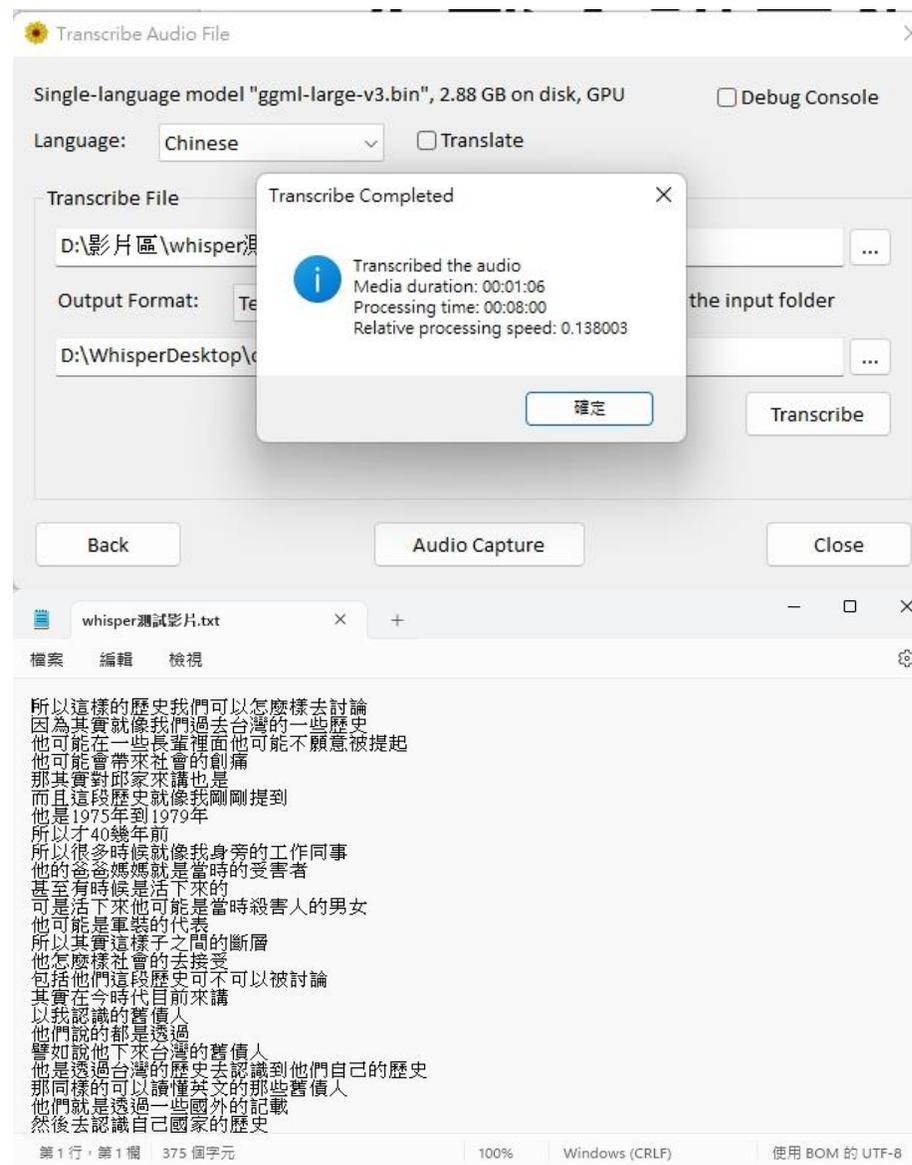
檔等，並選擇要輸出的

檔名後按下「Transcribe」



# 5.完成轉錄

根據模型的大小及影片長短，轉錄時間不同。轉錄完成後就可以打開剛剛輸出的txt檔或srt檔，欣賞轉錄的成果。



# 6. 影片教學



**END**