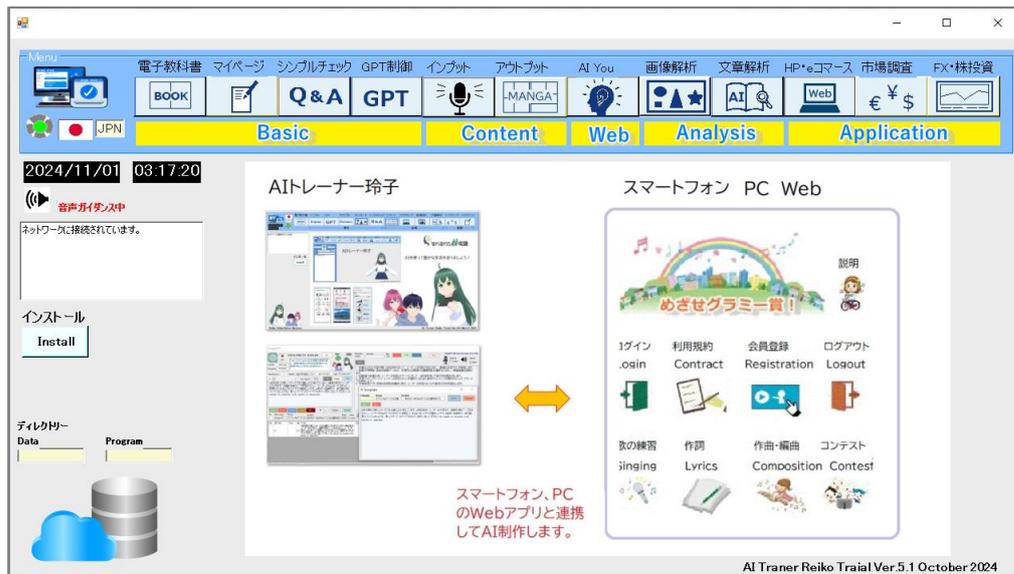


それぞれの AI 物語通信 No.20

2024年11月1日

AIトレーナー玲子の仕様改訂が行われ、これを最終形としてプログラミングしています。

1. AIトレーナー玲子改訂報告

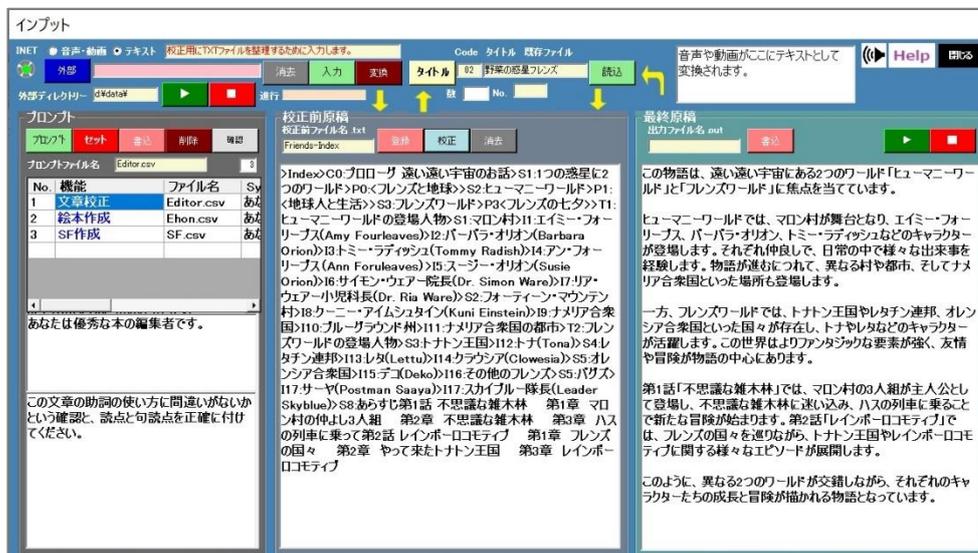


2. インプット

音声ファイル(MP3、M4A、WAV)、動画ファイル(MP4)からテキスト変換できるようになりました。Chat-GPTのWhisperを使用しているため、1時間の変換で約50円となります。1時間のテキスト変換時間は10分ほどです。その後これをさらに校正し最終テキストにします。

この右側にあるアイコンは、Web「AIアニメカレッジ」の「音声レコーダー」です。Chromeライブラリーを使っているため無料で、マイクからの直接音声変換になります。

音声レコーダー
Voice Recorder



3. アウトプット

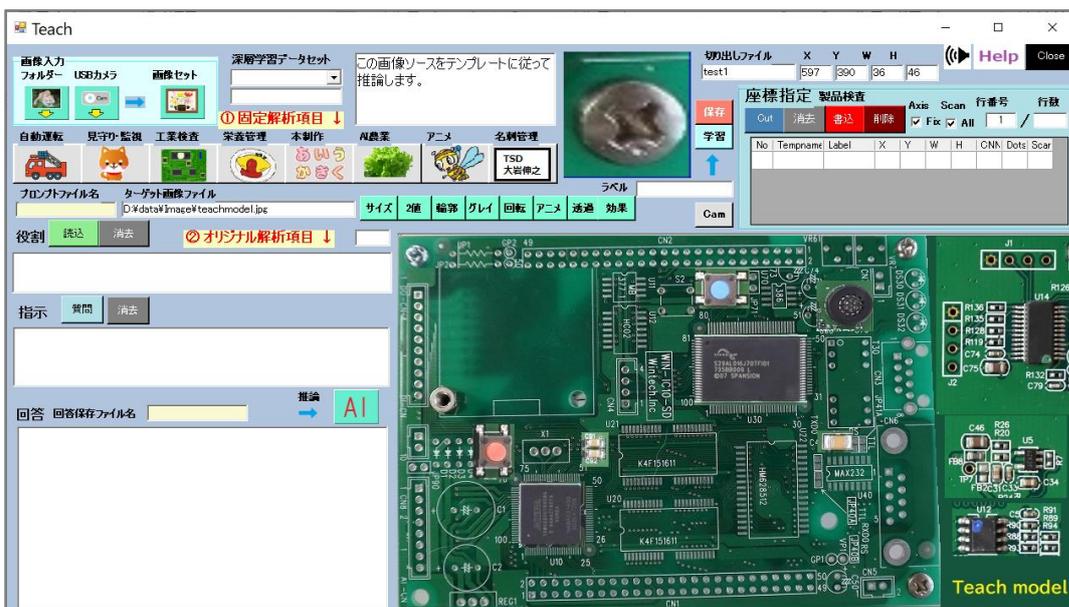
事務書類、論文、絵本、コマ割り漫画、画像、HTML、LLMなどを、PDF、PING、テキストなど、使用目的を変えていろいろな種類の媒体で出力します。



4. 画像解析

ChatGPT で画像を覚えさせてラベリングし、それと似た画像が来たらラベルを提示し、それではなければそのラベルではないと提示することはできません。現時点では、ChatGPT 自身に画像を直接アップロードしてラベリングを行う機能はありません。そのために通常の深層学習で行う機能を追加しました。

また解析だけでなく、覚えた画像を変化させることも今の Chat-GPT ではできません。ポケモンを覚えさせて、それと似たキャラクターを書くという機能はありません。この機能を今後強化して、オリジナルデータによる画像解析と、変化画像 (普通の顔→笑った顔) を作る機能を持たせます。



以前のオリジナルの CNN(C#)では理論が習えたが遅くて実用的ではなかったのですが、Python と Pytorch で、処理速度は格段に速くなりました。これで新しい画像解析はオリジナルデータ学習と gpt-4o の両刀使いです。オリジナルデータ学習では、Chat-GPT のできないラベリング式の画像解析や、希望するキャラクターを覚えさせ、いろいろな顔や形に変化させられます。

●オリジナルデータ学習

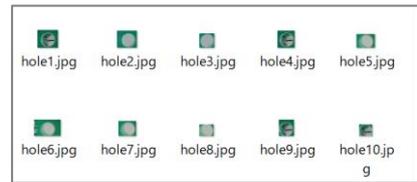
10 個ずつの 2 種類のデータセットの学習はやく 15 秒でした。画像の解像度で変わります。

<学習実行の様子>

```
D:\¥pg5¥GPT-XGA-DAO¥WindowsApp1¥bin¥Debug¥python>python teach.py
```

```
Epoch 1/10, Loss: 1.1474923014640808  
Epoch 2/10, Loss: 0.700454843044281  
Epoch 3/10, Loss: 0.6783158302307128  
Epoch 4/10, Loss: 0.6297705888748169  
Epoch 5/10, Loss: 0.5251092672348022  
Epoch 6/10, Loss: 0.45220770835876467  
Epoch 7/10, Loss: 0.38978760242462157  
Epoch 8/10, Loss: 0.23729774579405785  
Epoch 9/10, Loss: 0.14875568300485612  
Epoch 10/10, Loss: 0.14243025798350573
```

hole



screw



●推論

入力データは学習データにはない少しきり出しサイズを変えて(ビスや穴が少し小さくカット)与えています。GPU なしの推論時間は 2 秒から 8 秒くらいです。時間は速い時もあります。

<推論実行の様子>

```
D:\¥pg5¥GPT-XGA-DAO¥WindowsApp1¥bin¥Debug¥python>python recall.py
```

入力データ  ⇒ 出力 = screw

```
D:\¥pg5¥GPT-XGA-DAO¥WindowsApp1¥bin¥Debug¥python>python recall.py
```

入力データ  ⇒ 出力 = hole

5. LLM オリジナルデータ

音声入力したテキストをインプットとアウトプット機能で、オリジナルの LLM データに持っていきます。これを松崎由美子さんは健康管理の回答データの精度向上に活用します。別のグループでは FX 用の取引データとして LLM 化しますが、ビールを飲みながら、スマートフォンに録音してもらいます。この方が真実を言ってくれるからです(笑)。