

DOMAIN: <https://zenify-stt.ript.vn/>

API STT Chunkformer : <https://zenify-stt.ript.vn/api/stt/chunkformer>

API STT Zipformer : <https://zenify-stt.ript.vn/api/stt/zipformer>

1. File audio

- **Định dạng khuyến nghị:** WAV, mono, 16 kHz, 16-bit PCM
- **Kiểu gửi dữ liệu:** multipart/form-data

1.1. Lưu ý riêng cho Zipformer

Hàm `read_wave()` trong service Zipformer yêu cầu:

- Số kênh: **1** (mono)
- `sample_width` : **2 bytes** (tương ứng 16-bit PCM)

Nếu file không thỏa mãn, server sẽ raise lỗi (500).

2. Chunkformer

- **Method:** POST
- **URL:** /api/stt/chunkformer
- **Content-Type:** multipart/form-data

2.1. Input

- **Body (form-data):**
 - `file` (*bắt buộc*): file audio (khuyến nghị .wav, mono, 16kHz)
 - `chunk_size` (*optional, int, default: 64*)
Kích thước chunk (frame) cho Chunkformer.

- **left_context** (*optional, int, default: 128*)
Ngữ cảnh bên trái.
- **right_context** (*optional, int, default: 128*)
Ngữ cảnh bên phải.
- **return_timestamps** (*optional, bool, default: false*)
 - **false** : trả về **chỉ text** (transcription).
 - **true** : trả về **text + danh sách timestamps** theo từng đoạn.

2.2. Output (200 OK)

Kiểu dữ liệu trả về

```
{
  "model": "chunkformer",
  "transcription": "chuỗi văn bản sau khi nhận dạng",
  "status": "success",
  "num_chunks": null,
  "time_stamps": null
}
```

Khi `return_timestamps=true` , response sẽ có thêm trường `time_stamps` :

```
{
  "model": "chunkformer",
  "transcription": "chuỗi văn bản sau khi nhận dạng",
  "status": "success",
  "time_stamps": [
    {
      "start": 0.0,
      "end": 1.2,
      "decode": "xin chào"
    },
    {
      "start": 1.2,
      "end": 2.5,
      "decode": "mọi người"
    }
  ]
}
```

CURL mẫu

```
curl -X 'POST' \
'https://zenify-stt.ript.vn/api/v1/stt/chunkformer?chunk_size=64&left_context=128&right_context=128'
-H 'accept: application/json' \
-H 'Content-Type: multipart/form-data' \
-F 'file=@audio.wav;type=audio/wav'
```

3. Zipformer

- **Method:** POST
- **URL:** /api/stt/zipformer
- **Content-Type:** multipart/form-data

3.1. Input

- **Body (form-data):**
 - **file** (*bắt buộc*): file audio dạng WAV, **mono, 16kHz, 16-bit PCM**

3.2. Output (200 OK)

Kiểu dữ liệu trả về vẫn dùng schema `TranscriptionResponse` :

```
{
  "model": "zipformer",
  "transcription": "chuỗi văn bản sau khi nhận dạng",
  "status": "success",
  "num_chunks": null,
  "time_stamps": null
}
```

3.3. CURL mẫu

```
curl -X 'POST' \
'https://zenify-stt.ript.vn/api/v1/stt/zipformer' \
-H 'accept: application/json' \
-H 'Content-Type: multipart/form-data' \
-F 'file=@audio.webm;type=video/webm'
```