Whisper Transcriber

A modern, web-based audio transcription tool powered by OpenAI Whisper with real-time progress tracking, GPU acceleration support, and timestamped outputs.

🚀 Features

- **Drag & Drop Upload**: Intuitive web interface for audio file uploads
- **Multiple Whisper Models**: Support for tiny, base, small, medium, large, and v2/v3 variants
- **Real-time Progress**: Chunk-based progress tracking with ETA calculations
- **GPU Acceleration**: Automatic GPU detection with optimized performance estimates
- **Timestamped Transcripts**: Output includes both plain text and timestamped segments
- **Multi-job Dashboard**: Track multiple transcription jobs simultaneously
- **Audio Preprocessing**: Intelligent caching with .npy format for faster processing
- **Flexible Configuration**: Adjustable decoding parameters (temperature, beam size, etc.)
- **Production Ready**: Background processing with cancellation support

🎯 Quick Start

```
# Clone the repository
git clone https://github.com/gopichandbusam/whisper-transcriber.git
cd whisper-transcriber

# Make management script executable
chmod +x manage.sh

# Setup environment and dependencies
./manage.sh setup

# Run the application
./manage.sh run
```

Visit http://127.0.0.1:5000 in your browser and start transcribing!

📋 Requirements

- Python 3.9+
- macOS or Linux
- Sufficient disk space for model weights (100MB - 6GB depending on model)
- Optional: CUDA-compatible GPU for acceleration

🔧 Configuration

Model Selection

```
# Use a specific model size
WHISPER_MODEL_SIZE=small ./manage.sh run

# Enable chunk-based progress for better tracking
PROGRESS_MODE=chunks ./manage.sh run

# Custom port
PORT=8080 ./manage.sh run
```

Available Models

Model	Speed	Accuracy	RAM Usage	Best For
tiny	⚡⚡⚡⚡	⭐	~1 GB	Quick previews
base	⚡⚡⚡	⭐⭐	~1.3 GB	General use
small	⚡⚡	⭐⭐⭐	~2.5 GB	Balanced quality
medium	⚡	⭐⭐⭐⭐	~5 GB	High quality
large	🐌	⭐⭐⭐⭐⭐	~6+ GB	Maximum accuracy

📖 Documentation

- **Setup Guide**: See main README sections for detailed installation
- **API Reference**: Check `/health` endpoint for runtime configuration
- **Troubleshooting**: Common issues and solutions in main README

🤝 Contributing

We welcome contributions! Here's how you can help:

1. **Fork** the repository
2. **Create** a feature branch (`git checkout -b feature/amazing-feature`)
3. **Commit** your changes (`git commit -m 'Add amazing feature'`)
4. **Push** to the branch (`git push origin feature/amazing-feature`)
5. **Open** a Pull Request

Development Setup

```
./manage.sh shell   # Activate development environment
python app.py       # Run in debug mode
```

🐛 Issues & Support

- **Bug Reports**: Open an issue

- **Feature Requests**: Start a discussion
- **Questions**: Check existing issues or create a new one

📄 License

This project is licensed under the MIT License - see the LICENSE file for details.

🙏 Acknowledgments

- OpenAI Whisper for the amazing speech recognition model
- Flask for the web framework
- All contributors who help improve this project

📊 Project Stats

Stars | repo not found
Forks | repo not found
issues | repo not found
license | repo not found

Made with ❤️ by Gopichand Busam

If you find this project helpful, please consider giving it a ⭐ !