Problem Statement:

Modern classrooms often lack intelligent, real-time support tools that cater to individual student needs without interrupting the teaching flow. Students may hesitate to ask questions or struggle to understand complex visuals during lectures. To address this, the proposed solution is a multimodal AI-powered classroom assistant that enables students to interact using text, voice, or visual inputs. By integrating OCR, speech recognition, and large language models like LLaVA, the assistant can provide instant, personalized explanations for diagrams, slides, or spoken questions. This enhances student engagement, fosters independent learning, and bridges the gap between instructional delivery and real-time understanding.

Unique Idea Brief (Solution):

- We propose an AI Classroom Assistant that enables students to:
- Instantly snip any screen content (e.g., slides/diagrams)
- Get real-time AI-powered explanations
- Use voice input to ask questions
- Receive text and speech output
- The assistant uses OCR, speech recognition, and AI LLMs (via Ollama & LLaVA) to provide simplified, interactive responses that help students understand content independently.

Features Offered:

- **Snip Image:** Select any portion of the screen to extract text and diagram using OCR
- **Ask AI:** Send queries or extracted content to the LLaVA model for explanations
- **Speak:** Voice input for students to ask questions hands-free
- **Read Out:** Assistant speaks out the response using TTS
- **Clear:** Reset input/output easily for new interactions

Process Flow:

1. **Input:**

- Via Snip Tool (image)
- Via Voice (speech)
- Via Text (keyboard)

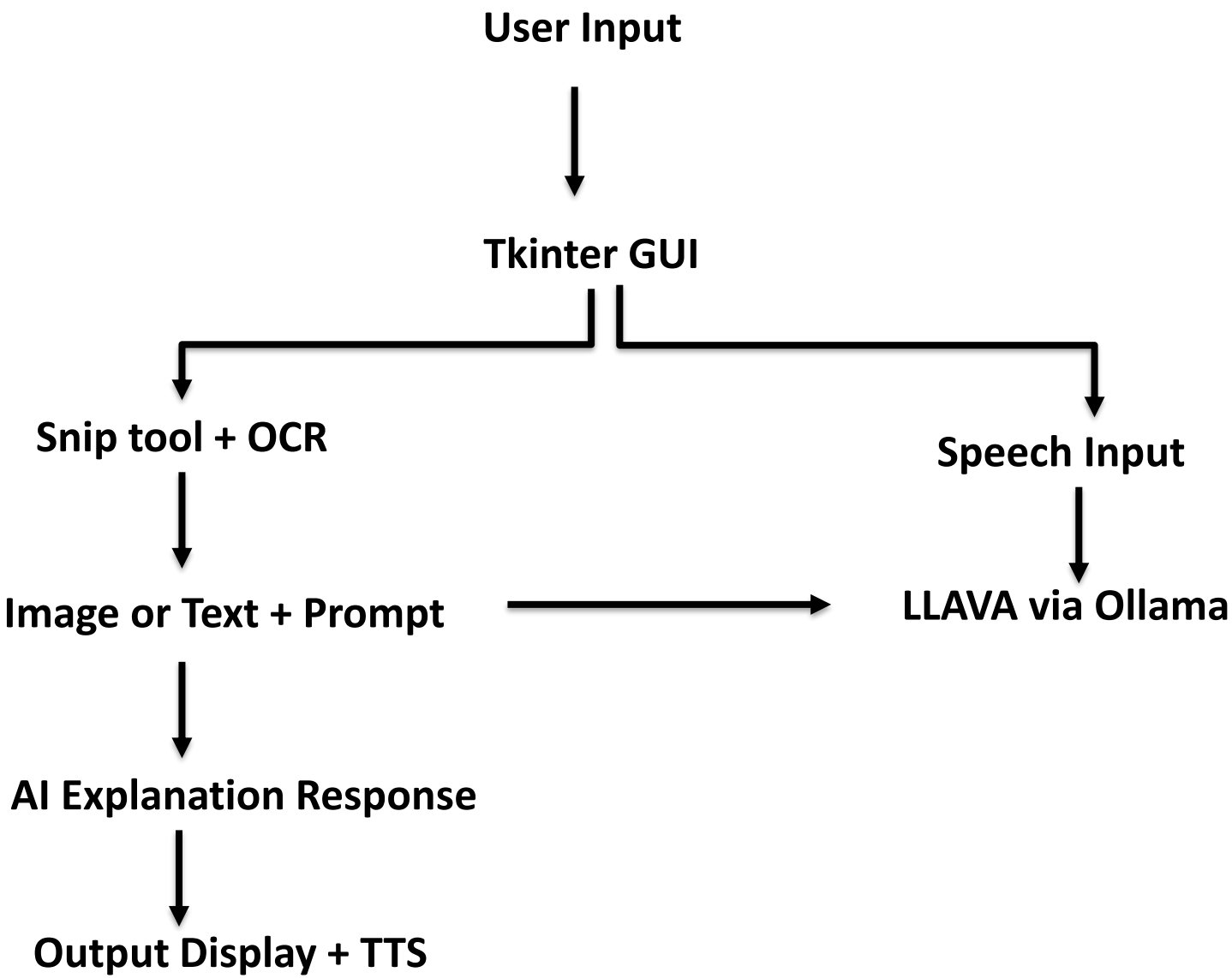
2. **Processing:**

- Image → OCR
- Speech → Text
- Text/Image + Prompt → Sent to LLaVA (via Ollama API)

3. **Output:**

- AI-generated response shown in the GUI
- Optionally read aloud using pyttsx3

Architecture Diagram:



Technologies Used:

- Python
- Tkinter – for GUI
- SpeechRecognition – for voice input
- pyttsx3 – for text-to-speech
- PIL + ImageGrab – for snipping screenshots
- pytesseract – OCR from images
- Ollama + LLaVA – for AI model interaction
- Requests – for HTTP-based LLM querying

Team Members & Contribution:

Name	Contribution
1. Singireddy Sriharsha	- Multimodal Input Processing (Text, Voice, and Image)
2. Nune Vyshali	- Real-Time AI Querying Using LLaVA via Ollama
3.Sunkoju Deekshith Chary	- Interactive GUI Built with Tkinter

Conclusion:

The AI Classroom Assistant provides a seamless, interactive way for students to understand complex content using AI. With image recognition, speech input, and AI explanations, it ensures continuous learning without disrupting the class. This tool enhances the classroom experience by bridging the gap between real-time content and immediate comprehension.