

   How do AI models process   

# multimodal inputs?

## Cover Letter

This is a slide deck intended for high-school students and undergraduate students in non-CS related majors.

The expected read time is around ten minutes.

The lesson aims to introduce early stage students to how multimodal AI can understand and process different types of information. The slideshow covers concepts such as input encoding, feature extraction, and existing multimodal architecture such as CLIP, but also emphasizes the relevance of multimodal AI to real-world problems.

Through clear visuals, the reader gains a surface-level understanding of multimodal AI without heavy technical jargon.