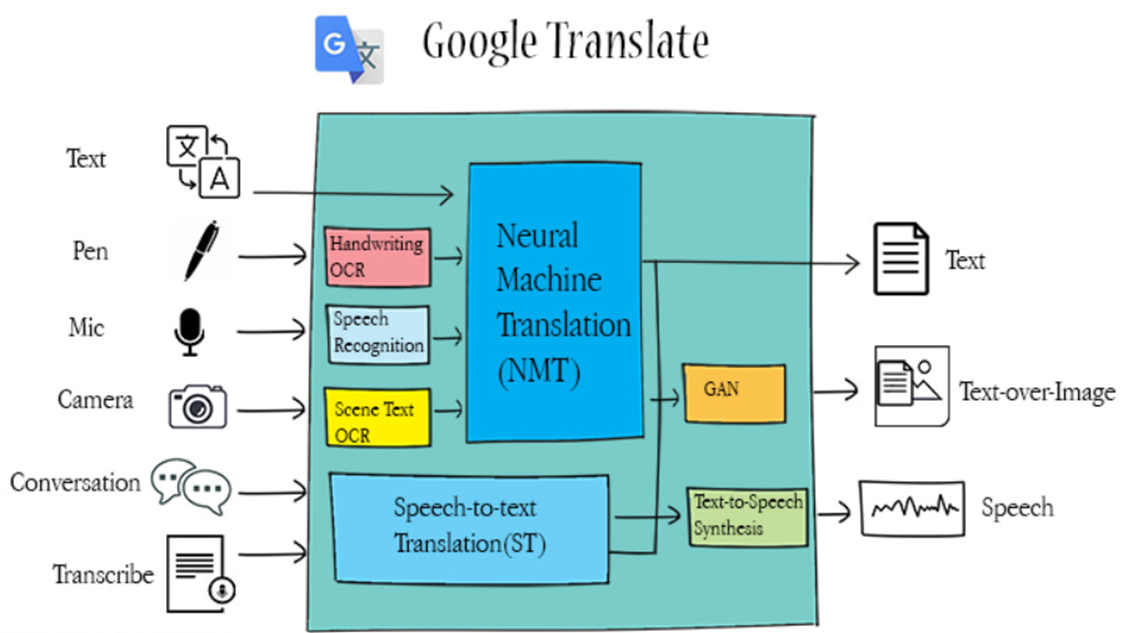
Mô hình Machine Translation của Google sẽ bao gồm:



**Input**:

* Text: Người dùng nhập văn bản trực tiếp.
* Pen: Hệ thống nhận diện chữ viết tay thông qua OCR (Optical Character Recognition).
* Mic: Nhận dạng giọng nói (Speech Recognition) cho văn bản thoại.
* Camera: OCR văn bản từ hình ảnh hoặc văn cảnh.
* Conversation: Ghi nhận các cuộc hội thoại để dịch trực tiếp.
* Transcribe: Chuyển đổi âm thanh thành văn bản.

**Xử lý chính**:

* Neural Machine Translation (NMT): Bộ máy dịch ngôn ngữ bằng trí tuệ nhân tạo sử dụng mạng nơ-ron để đảm bảo dịch chính xác và tự nhiên, được sử dụng đối với text, pen, mic và camera.
* Speech-to-Text Translation (ST): Dịch trực tiếp từ âm thanh (Conversation, Transcribe) thành văn bản.

**Ouput**:

* **Text**: Kết quả dịch dưới dạng văn bản, trực tiếp từ công nghệ NMT.
* **Text-over-Image**: Văn bản dịch chồng lên hình ảnh gốc, được hỗ trợ bởi GAN trong việc tạo hoặc chỉnh sửa văn bản trên hình ảnh để ra được output.
* **Speech**: Kết quả dịch được đọc to qua Text-to-Speech Synthesis.