

Image Captioning and Text to Image

Name: Yazeed Abdullah Al nashib

Date: 1\10\2024

Objectives

- **Automated Image Captioning:** Automatically generate a caption for a given image.
- **Text-to-Image Generation:** Create images from user-provided text descriptions or suggestions.
- **Speech Output:** Convert the generated captions to speech.
- **Multilingual Support:** Provide caption and speech output in multiple languages.

Pipeline Implementation

- **Step 1: Load Models:**
 - BLIP (Bootstrapping Language-Image Pre-training) for image captioning.
 - Stable Diffusion for text-to-image generation.
- **Step 2: Image Captioning:**
 - User uploads an image.
 - The system generates a caption using BLIP.
 - The caption is translated and converted to speech using GoogleTranslator and gTTS.
- **Step 3: Text-to-Image:**
 - User inputs a text description or selects a suggestion.
 - Stable Diffusion generates an image based on the input description.

Results

- **Image Captioning Example:**

- Input: Image of a sunset.
- Output: "A beautiful sunset over the mountains." (with speech)

- **Text-to-Image Example:**

- Input: "A robot playing chess with a human."
- Output: Generated image of a robot playing chess.

Image Captioning Results (Arabic , English)


Image Captioning with Speech

Text-to-Image

Image Captioning with Speech Output

Upload an image and get a caption in the selected language with audio output.


image




output 0


a woman with a backpack walking through a field of flowers

output 1



0:04

 1x



Flag

Select Output Language

English

Clear

Submit



Use via API  · تم الإنشاء باستخدام Gradio 


Image Captioning with Speech

Text-to-Image

Image Captioning with Speech Output

Upload an image and get a caption in the selected language with audio output.


image




output 0


امراه تحمل حقيبة ظهر وتمشي عبر حقل من الزهور

output 1



0:06

 1x





Flag

Select Output Language

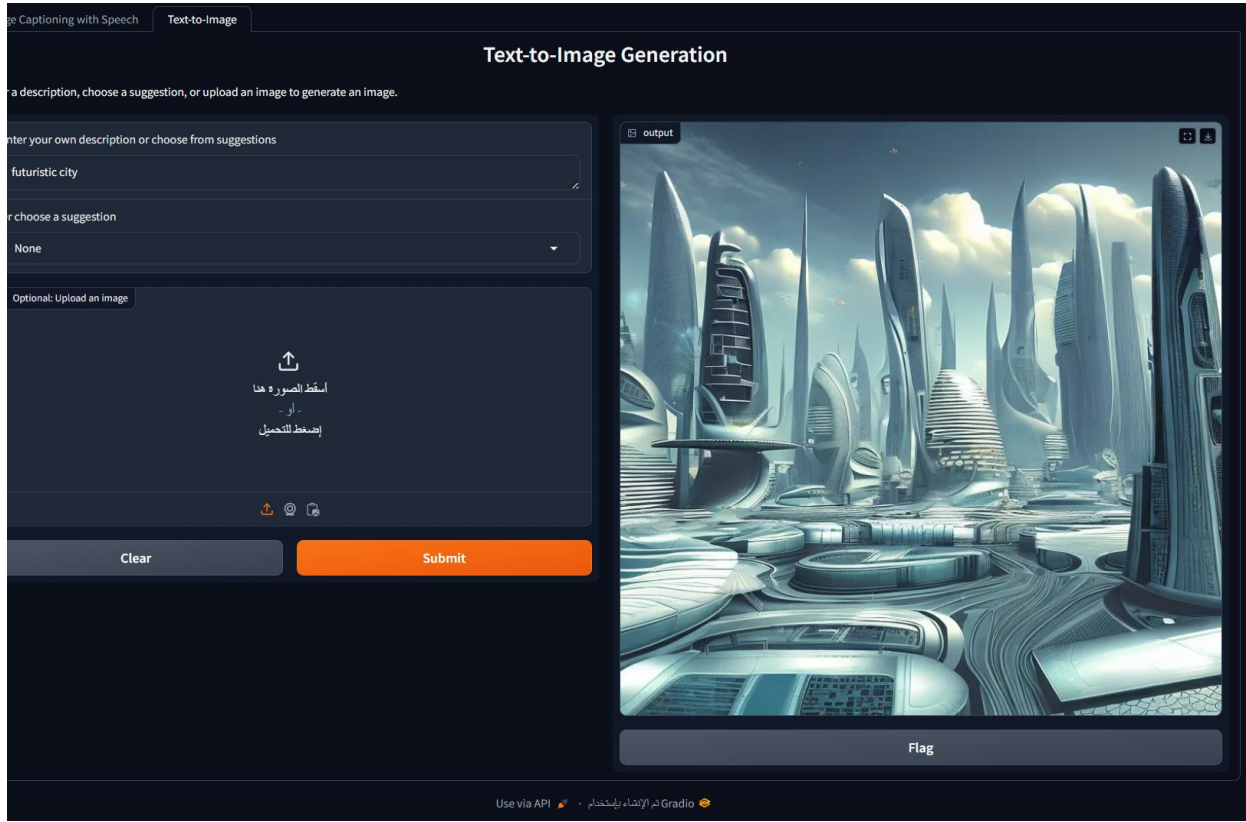
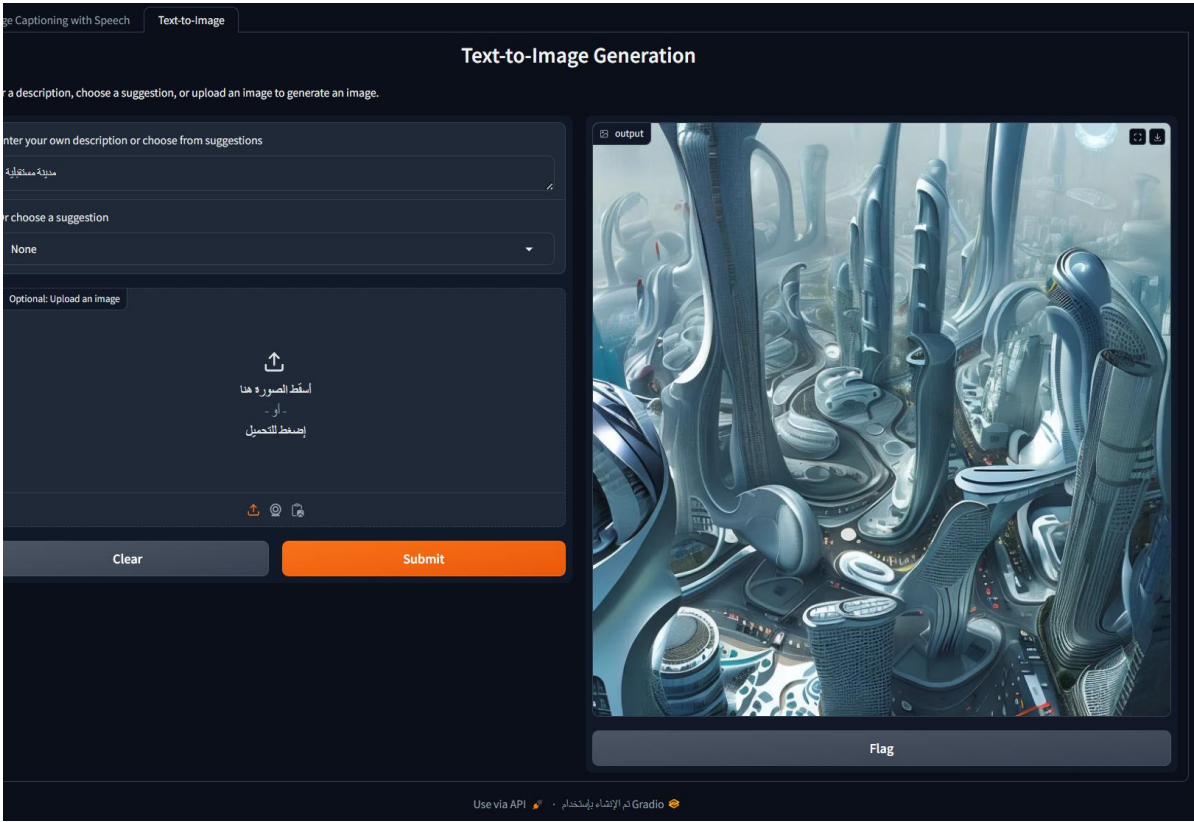
Arabic

Clear

Submit

Use via API  · تم الإنشاء باستخدام Gradio 

Text-to-Image Generation Results (Arabic , English)



Previous Work: Text Translation Project

- **Project Overview:**

This project was focused on translating text using the NLLB (No Language Left Behind) model by Facebook AI, with automatic detection of the input language.

Features:

Automatic language detection for translating text into multiple languages (English, French, Spanish, German, Arabic)

Easy-to-use web interface using Gradio.

Text-Translation Results

Text Translation with Auto Language Detection (Multiple Outputs)

Input Text

Artificial Intelligence is rapidly transforming industries and the way we live our daily lives. From healthcare to finance, AI applications are making processes more efficient and effective. As AI continues to evolve, its impact on society will grow even more profound, shaping the future in unprecedented ways.

Upload Text File (.txt)

أضف الملف هنا

أو -

إضغط للتحميل

Target Languages

English

French

Spanish

German

Arabic

Translated Text

Arabic:

ملاحظة: إن الذكاء الاصطناعي هو مجال علوم الكمبيوتر مخصص لحل المشاكل المعقدة المرتبطة بالذكاء البشري. مثال التعلم والإدماج والتعرف على الصور. تجمع المؤسسات الحديثة كميات كبيرة من البيانات من مصادر مختلفة مثل أجهزة الاستشعار الذكية والمحتوى الذي يخلقه الإنسان وأدوات المراقبة وسجلات النظام الهدف من (AI) الذكاء الاصطناعي هو إنشاء أنظمة التعلم الذاتي التي تستخرج معنى البيانات. وبدون ذلك، يمكن للذكاء الاصطناعي تطبيق هذه المعرفة لحل المشاكل الجديدة بطرق تشبه الإنسان. على مدار العقد، نتج تكملة لوجيا للذكاء الاصطناعي الأسلحة المستجيبة للتحديات البشرية. وإشياء مسورة وتصميم أسلحة واتخاذ القرارات على أساس إحصائيات البيانات في الوقت الحقيقي. يمكن لمؤسساتك دمج إمكانات الذكاء الاصطناعي في تطبيقاتك لتحسين عمليات عملك وتحسين تجربة العملاء وتسريع الابتكار.

German:

Künstliche Intelligenz verändert schnell Industrien und die Art und Weise, wie wir unser tägliches Leben leben. Von der Gesundheitsversorgung bis hin zur Finanzierung machen KI-Anwendungen die Prozesse effizienter und effizienter. Da AI weiterentwickelt wird, wird sein Einfluss auf die Gesellschaft noch tiefer werden und die Zukunft auf unvorhergesehene Weise gestalten.

Spanish:

La inteligencia artificial está transformando rápidamente las industrias y la forma en que vivimos nuestras vidas diarias. Desde la salud hasta la financiación, las aplicaciones de inteligencia artificial hacen que los procesos sean más eficientes y eficaces. A medida que la IA continúa evolucionando, su impacto en la sociedad crecerá aún más profundo, formando el futuro de maneras sin precedentes.

French:

L'intelligence artificielle transforme rapidement les industries et la façon dont nous vivons notre vie quotidienne. De la santé à la finance, les applications d'IA rendent les processus plus efficaces et plus efficaces. Au fur et à mesure que l'IA continue d'évoluer, son impact sur la société sera encore plus profond, formant l'avenir d'une manière sans précédent.

Translate Text

Translate File

Use via API

Gradio

تم الإنشاء باستخدام

Text Translation with Auto Language Detection (Multiple Outputs)

Input Text

Upload Text File (.txt)

F_T2.txt

1.4 KB

Target Languages

English

French

Spanish

German

Arabic

Translated Text

Arabic:

ملاحظة: إن الذكاء الاصطناعي هو مجال علوم الكمبيوتر مخصص لحل المشاكل المعقدة المرتبطة بالذكاء البشري. مثال التعلم والإدماج والتعرف على الصور. تجمع المؤسسات الحديثة كميات كبيرة من البيانات من مصادر مختلفة مثل أجهزة الاستشعار الذكية والمحتوى الذي يخلقه الإنسان وأدوات المراقبة وسجلات النظام الهدف من (AI) الذكاء الاصطناعي هو إنشاء أنظمة التعلم الذاتي التي تستخرج معنى البيانات. وبدون ذلك، يمكن للذكاء الاصطناعي تطبيق هذه المعرفة لحل المشاكل الجديدة بطرق تشبه الإنسان. على مدار العقد، نتج تكملة لوجيا للذكاء الاصطناعي الأسلحة المستجيبة للتحديات البشرية. وإشياء مسورة وتصميم أسلحة واتخاذ القرارات على أساس إحصائيات البيانات في الوقت الحقيقي. يمكن لمؤسساتك دمج إمكانات الذكاء الاصطناعي في تطبيقاتك لتحسين عمليات عملك وتحسين تجربة العملاء وتسريع الابتكار.

Spanish:

¿Qué es la inteligencia artificial? La inteligencia artificial (AI) es un campo de ciencias informáticas dedicado a resolver problemas cognitivos generalmente relacionados con la inteligencia humana, como el aprendizaje, la creatividad y el reconocimiento de imágenes. Las instituciones modernas recogen grandes cantidades de datos de diversas fuentes, como sensores inteligentes, contenido creado por los humanos, herramientas de vigilancia y registros de sistema. El objetivo de la inteligencia artificial es crear sistemas de autoaprendizaje que extraen significados de los datos. Luego, la inteligencia artificial puede aplicar ese conocimiento para resolver nuevos problemas de manera similar a la humana. Por ejemplo, la inteligencia artificial puede responder de manera intencional a las conversaciones humanas, crear imágenes y textos originales y tomar decisiones basadas en las entradas de datos en tiempo real. Su organización puede integrar el potencial de la inteligencia artificial en sus aplicaciones para mejorar sus procesos de negocio, mejorar la experiencia de los clientes y acelerar la innovación.

Translate Text

Translate File

Use via API

Gradio

تم الإنشاء باستخدام

Model Justification

- **Why BLIP for Captioning:**

- BLIP provides accurate and efficient image-to-text generation.

- **Why Stable Diffusion for Text-to-Image:**

- Stable Diffusion offers high-quality, detailed images based on text descriptions.

- **Why Google Translator:**

- Used for automatic language detection and translation of captions.
- Supports multiple languages to reach a broader audience.

GitHub & Hugging Face Links

- GitHub Repository:
- <https://github.com/Yaz1-e/Image-Captioning-Text2Image>
- Hugging Face Space:
- <https://huggingface.co/spaces/Yaz1-e/Image-Captioning-Text2Image>