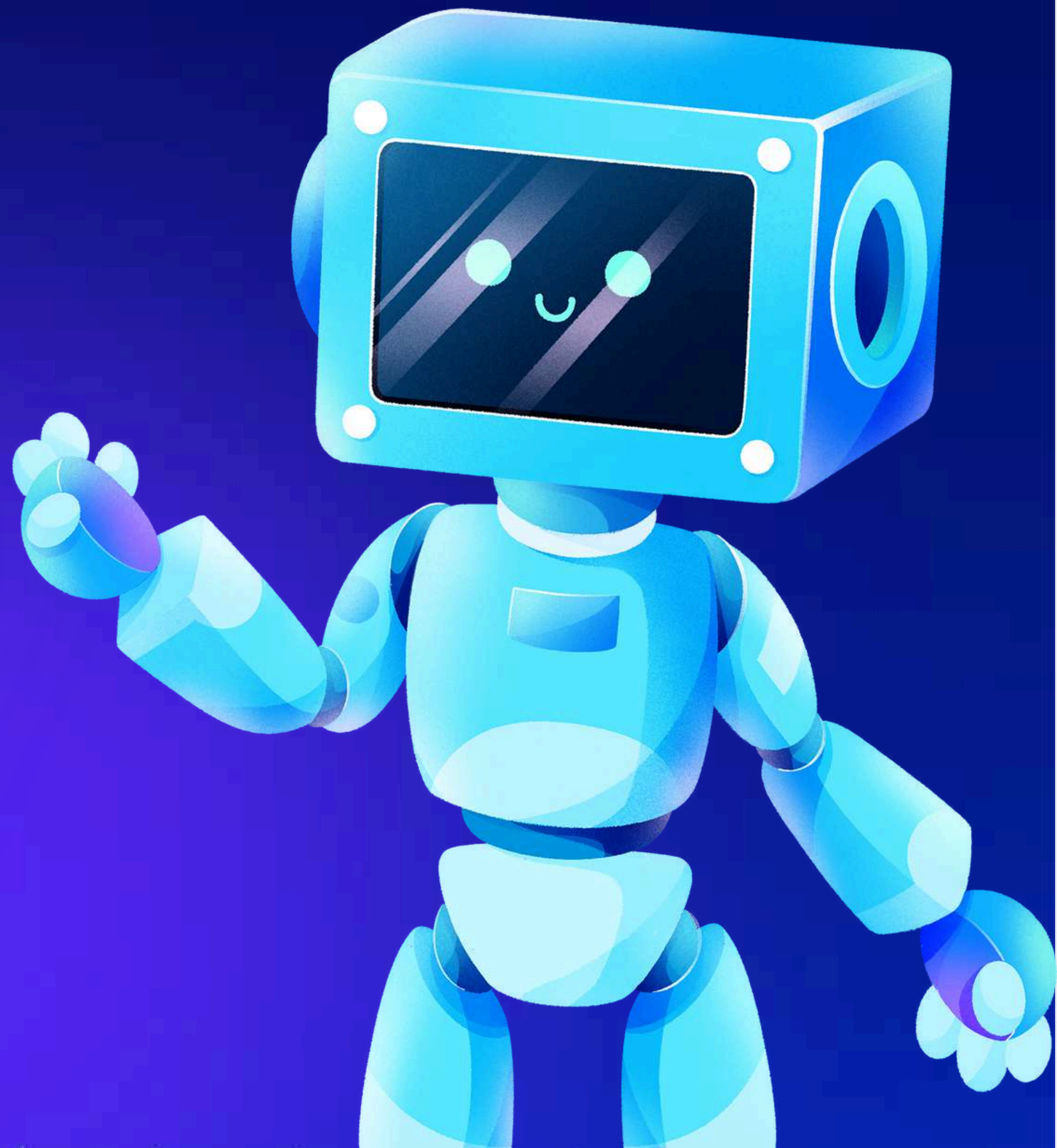


FINAL PROJECT
IMAGE GENERATION
AND CAPTIONING



OUR TEAM



Shahad Alhassani

Data Analyst

✉ Shahadbalhassani@gmail.com

[in](#) [@Shahad Alhassani](#)



Noura Abdullah

Data Analyst

✉ nourhamushi@gmail.com

[in](#) [@noura abdullah](#)

PREVIOUS WORK

This project aims to demonstrate named entity recognition (NER) for both English and Arabic languages. It utilizes Hugging Face Transformers and Gradio to extract and highlight named entities from the input text.

[PROJECT LINK](#) 



DEMO

😊 Spaces

1Noura/Named_Entity_Recognition

♡ like 0

● Running

⋮

☰

Named Entity Recognition

Select a language and enter text to extract and highlight named entities.

Input Text

Hugging Face Inc. is a company based in New York City.

Select Language

☐ English ☐ Arabic

Clear

Submit

Highlighted NER Results

☰ Examples

Input Text	Select Language
Hugging Face Inc. is a company based in New York City.	English
أحمد هو عالم في مجال الذكاء الاصطناعي	Arabic

PROJECT CONCEPTS

To generate images based on captions from uploaded images and provide translations of those captions from English to Arabic.

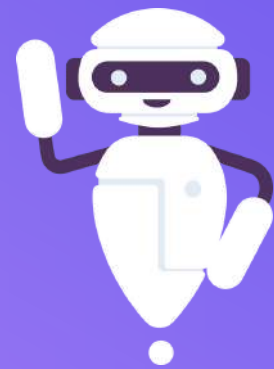


IMAGE
CAPTIONING
USING BLIP

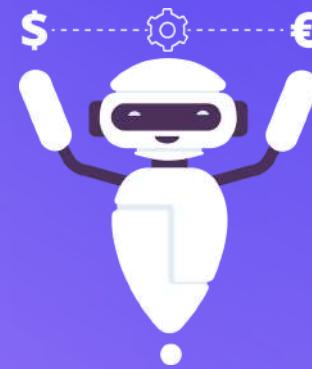






IMAGE
GENERATION
USING STABLE
DIFFUSION




TRANSLATION
USING NLLB

DEMO





 Number of Images



2

Clear

Submit


Examples



Upload Image



Number of Images

3





Generated Caption (English)

there is a small kitten that is walking in the grass

Translated Caption (Arabic)

هناك قطة صغيرة تسير في العشب

Flag

PIPELINE WORKFLOW



Upload Image

The user uploads an image.

Caption Generation

Using BLIP, a caption in English is generated based on the image content.

Translation

The caption is translated from English to Arabic using the NLLB model.

Image Generation

Stable Diffusion generates new images based on the English caption.

Display

The original and translated captions are displayed alongside the generated images.

MODEL JUSTIFICATION

01

- BLIP (Image Captioning):

1. Chosen for its state-of-the-art performance in generating descriptive captions based on images.
-

02

- Stable Diffusion (Image Generation):

1. A widely used and powerful diffusion model for generating high-quality images based on textual prompts.
-

03

- NLLB-200 (Translation):

1. A multilingual translation model, perfect for translating between English and Arabic, with efficient processing and high accuracy

CODE

```
# Load the models
caption_image = pipeline("image-to-text", model="Salesforce/blip-image-captioning-large", device=device)
sd_pipeline = StableDiffusionPipeline.from_pretrained("runwayml/stable-diffusion-v1-5").to(device)

# Load the translation model (English to Arabic)
translator = pipeline(
    task="translation",
    model="facebook/nllb-200-distilled-600M",
    torch_dtype=torch.bfloat16,
    device=device
)
```



01

PIPELINE IMPLEMENTATION

Import for libraries

Gradio

For building the user interface.

Transformers

To load and apply models (BLIP and NLLB).

Diffusers

For Stable Diffusion pipeline.

Torch

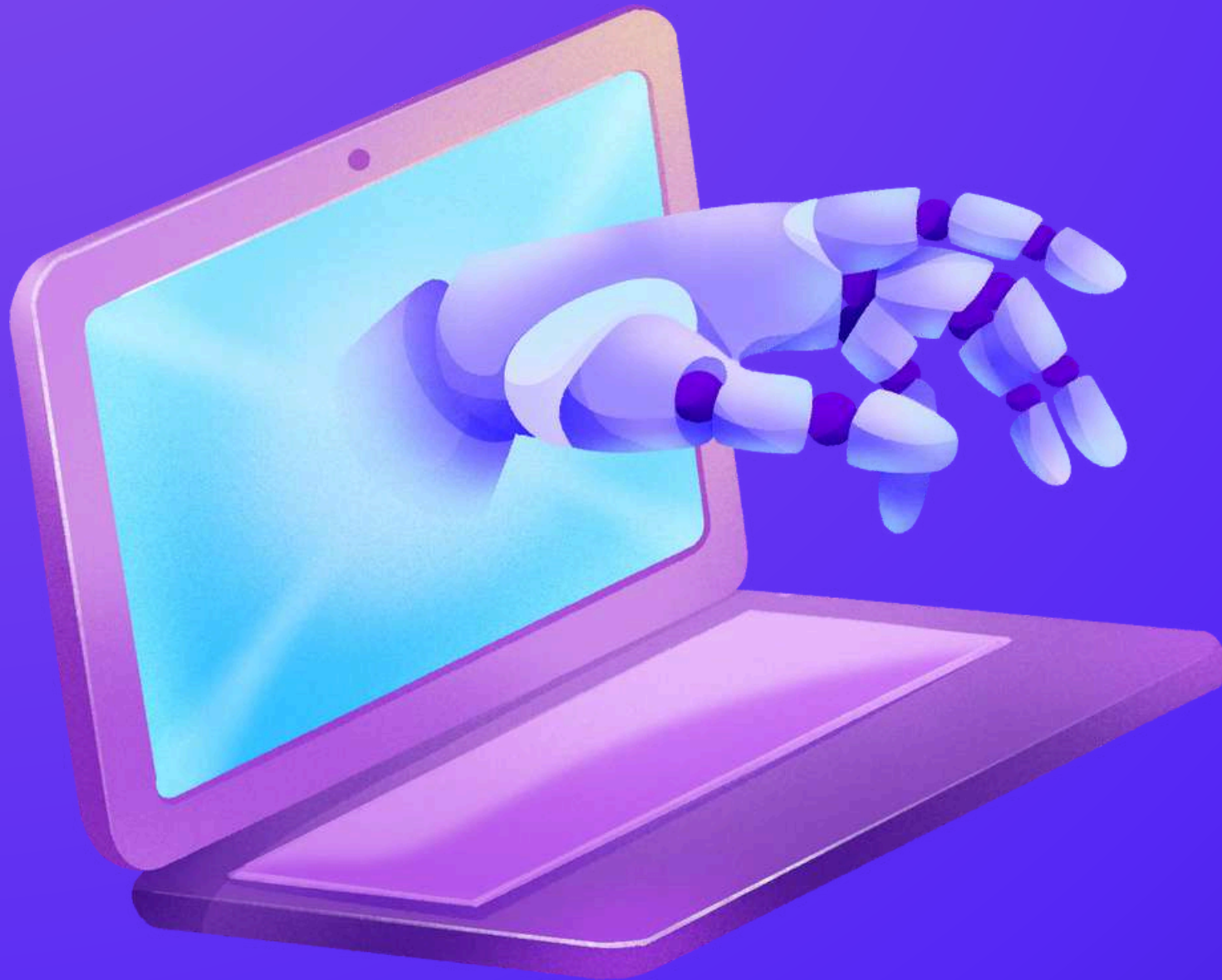
For GPU acceleration.

Wget

For downloading the images.



```
!pip install gradio  
!pip install transformers  
!pip install diffusers  
!pip install torch
```



02

PIPELINE IMPLEMENTATION

Loading Models

```
# Load the models
caption_image = pipeline("image-to-text", model="Salesforce/blip-image-captioning-large", device=device)
sd_pipeline = StableDiffusionPipeline.from_pretrained("runwayml/stable-diffusion-v1-5").to(device)

# Load the translation model (English to Arabic)
translator = pipeline(
    task="translation",
    model="facebook/nllb-200-distilled-600M",
    torch_dtype=torch.bfloat16,
    device=device
)
```

(03)

PIPELINE IMPLEMENTATION

Main Function

```
# Function to generate images based on the image's caption
def generate_image_and_translate(image, num_images=1):
    # Generate caption in English from the uploaded image
    caption_en = caption_image(image)[0]['generated_text']

    # Translate the English caption to Arabic
    caption_ar = translator(caption_en, src_lang="eng_Latn", tgt_lang="arb_Arab")[0]['translation_text']

    generated_images = []

    # Generate the specified number of images based on the English caption
    for _ in range(num_images):
        generated_image = sd_pipeline(prompt=caption_en).images[0]
        generated_images.append(generated_image)

    # Return the generated images along with both captions
    return generated_images, caption_en, caption_ar
```


04

PIPELINE IMPLEMENTATION

Interface Setup


```
# Set up the Gradio interface
interface = gr.Interface(
    fn=generate_image_and_translate, # Function to call when processing input
    inputs=[
        gr.Image(type="pil", label="Upload Image"), # Input for image upload
        gr.Slider(minimum=1, maximum=10, label="Number of Images", value=1, step=1) # Slider to select number of images
    ],
    outputs=[
        gr.Gallery(label="Generated Images"), # Output for displaying generated images
        gr.Textbox(label="Generated Caption (English)", interactive=False), # Output for English caption
        gr.Textbox(label="Translated Caption (Arabic)", interactive=False) # Output for Arabic caption
    ],
    title="Image Generation and Translation", # Title of the interface
    description="Upload an image to generate new images based on its caption and translate the caption into Arabic.",
    examples=[ # Example input
        ["sea.jpg", 3],
        ["Cat.jpeg", 4],
        ["Car.jpeg", 2]
    ]
)
```




RESULTS & EXAMPLES

Image Generation and Captioning

Upload an image to extract a caption and display it in both Arabic and English. Then, a new image will be generated based on that caption.

Upload Image






Number of Images

1

Clear

Submit

Generated Images



Generated Caption (English)

there is a small kitten that is walking in the grass

Translated Caption (Arabic)

هناك قطة صغيرة تسير في العشب


Flag




RESULTS & EXAMPLES

Image Generation and Captioning

Upload an image to extract a caption and display it in both Arabic and English. Then, a new image will be generated based on that caption.

Upload Image









Number of Images

3

Clear

Submit





Generated Caption (English)

arafeed view of a lake with a mountain in the background

Translated Caption (Arabic)

نظرة من بحيرة مع جبل في الخلفية

PROJECT ON COLAB

LINK 



• LINKS

SHAHAD ALHASSANI

[GITHUB](#) 

[HUGGING FACE SPACE](#) 

NOURA ABDULLAH

[GITHUB](#) 

[HUGGING FACE SPACE](#) 



THANK YOU!

