EX.NO.: 08
DATE: 24.07.2025

COMPUTER VISION WITH NLP FOR OCR BASED IMAGE QUESTION

AIM
To develop a Python-based system that extracts text from an image using Optical Character Recognition (OCR) and answers a natural language question based on the extracted content using a pre-trained NLP model.

ALGORITHM
1. Input: Provide an image (e.g., a scanned receipt) and a question (e.g., "What is the total amount?").
2. Text Extraction:
    a. Load the image using PIL.
    b. Use `pytesseract` to perform OCR and extract text from the image.
3. Question Answering:
    a. Load a pre-trained transformer model (`distilbert-base-uncased-distilled-squad`) using HuggingFace's `pipeline`.
    b. Feed the extracted text as context and the input question to the model.
4. Get the predicted answer based on the context.
5. Output:
    a. Display the OCR-extracted text.
    b. Display the question and the predicted answer.

CODE AND OUTPUT
```
from PIL import Image
import pytesseract
import os

def extract_text_from_image(image_path):
    image = Image.open(image_path)
    text = pytesseract.image_to_string(image)
    return text
```

```
from transformers import pipeline

def answer_question(context, question):
    qa = pipeline("question-answering",
model="distilbert-base-uncased-distilled-squad")
    result = qa(question=question, context=context)
    return result['answer']

def main(image_path, question):
    print("[1] Extracting text from image using OCR...")
    ocr_text = extract_text_from_image(image_path)
    print("\n[OCR Extracted Text]:\n", ocr_text)

    print("\n[2] Answering question using NLP...")
    answer = answer_question(ocr_text, question)
```

```
    print("\n[Question]:", question)
    print("[Answer]:", answer)

if __name__ == "__main__":
    image_path = "Receipt.jpg"
    question = "What is the total amount?"

    main(image_path, question)
```

```
[1] Extracting text from image using OCR...

[OCR Extracted Text]:
 erate,
Sensis 9?
Fate alee

movie 6 Rami
oe Teh ot

Deatte emote 4 650 OF 9.00
ata 4 Soo oF So
sdweradtael $200 OF 2a)
Nowa 8 SS OF

taal: oF 54.50

freh NoE Mat SAO 35

Enseat fra 280 OR
eset See ele

hat ese 0 2
sels a9 68
Fees a9 es 6) 8
...



[2] Answering question using NLP...
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
Device set to use cpu

[Question]: What is the total amount?
[Answer]: $200 OF 2a
```

INFERENCE

The system effectively combines OCR and NLP to automatically extract and interpret information from images. It can accurately answer context-based questions, such as identifying totals from receipts, enabling intelligent document analysis.