

## مشروع مساعد الذكي

### فكرة المشروع:

تطبيق ذكي تفاعلي يستخدم الذكاء الاصطناعي لتحليل الصور وتحديد كائن معين يطلبه المستخدم صوتياً. يعتمد على نموذج Gemini من Google لتحليل الصور، ويقدم استجابة دقيقة باللغة العربية تتضمن:

- هل الكائن موجود؟
- موقعه في الصورة
- ما حوله
- لونه أو شكله (إن أمكن)

يتم التفاعل عن طريق الصوت، ويتم الرد نصياً وشفهياً باستخدام تحويل النص إلى كلام.

### التقنيات والمكتبات المستخدمة:

المكتبة/الأداة	الاستخدام
google.generativeai	استخدام نموذج Gemini 1.5 Flash لتحليل الصور بناءً على طلب المستخدم.
PIL (Pillow)	فتح الصور، تصغيرها، وتحويلها إلى تنسيق مناسب للعرض والمعالجة.
ImageTk	لعرض الصور داخل واجهة customtkinter
customtkinter	لإنشاء واجهة مستخدم رسومية متقدمة وحديثة بدل tkinter
speech_recognition	لتحويل الكلام الصوتي إلى نص (فهم سؤال المستخدم عن الكائن المطلوب).
gTTS + pygame	تحويل النص الناتج إلى صوت طبيعي باللغة العربية وتشغيله.
pytsx3	تحويل نص إلى صوت داخل التطبيق (نسخة محلية دون اتصال بالإنترنت).
os, tempfile, time	لإدارة الملفات المؤقتة وتشغيل الملفات الصوتية وإدارتها مؤقتاً.
BytesIO	لتحويل الصور إلى تنسيق بايت من الذاكرة دون الحاجة لحفظها في ملف فعلي.

### آلية عمل المشروع:

#### 1. واجهة المستخدم

- المستخدم يفتح البرنامج ويظهر له زران:
  - تحميل صورة
  - تسجيل صوتي

## 2. تسجيل صوتي:

- عند الضغط على "تسجيل صوتي"، يبدأ البرنامج في الاستماع للمستخدم.
- يتم تحويل الصوت إلى نص باستخدام speech\_recognition

## 3. تحميل الصورة:

- بعد التعرف على طلب المستخدم (مثلاً: "ابحث عن الهاتف في الصورة")، يضغط المستخدم على "تحميل صورة".
- تُعرض الصورة داخل التطبيق.

## 4. تحليل الصورة:

- يُرسل الصوت المعالج (كنص) مع الصورة إلى نموذج Gemini عبر google.generativeai
- النموذج يبحث فقط عن الكائن المطلوب (مثلاً: الهاتف) ويُعيد وصفًا دقيقًا.

## 5. عرض النتائج:

- يتم عرض الوصف باللغة العربية في مربع النص.
- يُقرأ الوصف أيضًا صوتيًا باستخدام gTTS و pygame

## 6. زر تشغيل الصوت مجددًا:

- يتيح للمستخدم سماع الوصف مرة أخرى في أي وقت.

## شرح دور كل مكتبة:

المكتبة	دورها
google.generativeai	يستخدم لتوصيل الصورة والنص (استعلام المستخدم) إلى نموذج Gemini
Pillow (PIL)	يفتح الصور من الجهاز. يضغط الصور (thumbnail) لتناسب نافذة العرض. يحول الصور إلى بايت قبل إرسالها للنموذج.
customtkinter	لإنشاء واجهة رسومية بشكل أنيق وحديث. يتيح إضافة الأزرار، مربعات النص، الإطارات، إلخ.
speech_recognition	يستخدم Google Speech Recognition API لتحويل الصوت إلى نص عربي.
gTTS	لتحويل النص العربي الناتج من النموذج إلى ملف صوتي بصوت طبيعي.
pygame	لتشغيل ملف الصوت الناتج من gTTS
pytsx3	بديل محلي لتحويل النص إلى صوت (لا يتطلب اتصالاً بالإنترنت) يُستخدم كخيار احتياطي.

