

**جـامعـة القلمون**

**كلية الهندسة**

قسـم هندسة تقانة المعلومات

**تطبيق وصف الصور لمساعدة المكفوفين**

**(ODTV)**

يغطي هذا التقرير مشروع التخرج كأحد متطلبات نيل شهادة **البكالوريوس** في علوم هندسة تقانة المعلومات

إعداد:

عبد الهادي حيدر فاطمه ملحم روان الأحمد

**بإشراف:**

**د.**شذا شحادة

2023/2024

We certify that we have read this project and that in our opinion it is fully adequate, in scope and quality, as dissertation for (project/ final project).

|  |  |  |
| --- | --- | --- |
| **أسماء لجنة الحكم:** | التوقيع | **Exam committee :** |
| الدكتور  الدكتور  الدكتور |  | Dr  Dr  Dr |

**الاهداء**

**فاطمة**

**روان**

**عبد الهادي**

**الشكر والتقدير**

لا بد لنا ونحن نخطو خطانا الأخيرة في رحاب رحلتنا الجامعية أن نتوقف لحظة، نسترجع فيها كل اللحظات التي مررنا بها وصولاً إلى هنا ، فمنذ هذه اللحظة استحالت تلك المواقف ذكريات قد تحمل مشاعر مختلفة، لكنها حتماً لها أثر في حياتنا مستقبلاً، لذلك نتقدم بجزيل الشكر لأعضاء الهيئة التدريسية في كلية الهندسة .

وأخص بالذكر الدكتورة شذا شحادة التي تكرّمت بالإشراف على هذا المشروع، وساهمت في إثراء معرفتي وعملي وكانت رمزاً للعطاء.

# **الملخص**

مع وجود الذكاء الاصطناعي والتطور الكبير فيه والذي بدوره يقدم حلولا مساعدة في جميع المجالات, ظهرت الحاجة لإيجاد وسيلة مساعدة للمكفوفين وضعاف البصر لمساعدتهم على التنقل وحدهم في كل مكان دون الحاجة إلى الاعتماد الطرق التقليدية مثل كلب مدرب أو عصا, ومع أنتشار الهواتف المحمولة وتوفرها بأيدي الجميع تم اقتراح فكرة هذا النظام الذي يسهل على المكفوفين التحرك باستخدام هواتفهم المحمولة ,قمنا ببناء هذا النظام بعد البحث ولرغبتنا في توفير آلية أمنة وموثوقة تسهل على المكفوفين الحركة في جميع الاتجاهات بشكل امن عن طريق التقاط صور من حولهم باستخدام كاميرة الهاتف المحمول واكتشاف محتوى الصور من خلال التطبيق وتحويل محتوى الصور ألى صوت باستخدام تقنيات معينة ليتمكن المكفوف من سماع اسم الغرض الذي قد يكون في طريقه والتصرف تماما كما لو انه يبصر.

الكلمات المفتاحية - تحديد الأغراض, تعلم الألة, , تطبيق اندرويد, الرؤية الحاسوبية, الشبكات عصبية.

# 

# فهرس الأشكال

الشكل1-1 ------------------------------------------------------------------------------------3

الشكل1-2 ------------------------------------------------------------------------------------4

الشكل2-1 ------------------------------------------------------------------------------------9

الشكل 2-2 ------------------------------------------------------------------------------------12

الشكل3-1 ------------------------------------------------------------------------------------17

الشكل4-1 ------------------------------------------------------------------------------------24

الشكل4-2 ------------------------------------------------------------------------------------25

الشكل4-3 ------------------------------------------------------------------------------------25

الشكل4-4 ------------------------------------------------------------------------------------26

الشكل 4-5------------------------------------------------------------------------------------27

الشكل 4-6 ------------------------------------------------------------------------------------28

الشكل 4-7 ------------------------------------------------------------------------------------29

الشكل 4-8 ------------------------------------------------------------------------------------30

# فهرس الجداول

جدول (2- 1 )----------------------------------------------------------------------------------------------7-8

جدول (2-2)-------------------------------------------------------------------------------------------------9

جدول ( 2 – 3) --------------------------------------------------------------------------------------------- 10

جدول (2-4)-------------------------------------------------------------------------------------------------10

جدول (2-5)-------------------------------------------------------------------------------------------------11

# 

# **جدول المصطلحات والاختصارات**

**جدول الاختصارات**

|  |  |
| --- | --- |
| ML | Machine Learning **(تعلم الآلة)** |
| ODTV | Object Detection To Voice |
| TFLiteV2 | TensorFlow Lite library (version2) |
| Flutter\_tts | Flutter text-to-speech |
| YOLO(V2) | You Only Look Once Version 2 |
| DL | Deep Learning |

**المصطلحات**

|  |  |
| --- | --- |
| MachineLearning  **(تعلم الآلة)** | **التعلم الآلي هو فرع من فروع الذكاء الاصطناعي وعلوم الكمبيوتر الذي يركز على استخدام البيانات والخوارزميات لتقليد الطريقة التي يتعلم بها البشر، مما يؤدي إلى تحسين دقتها تدريجيًا.** |
| ObjectDetection | **تحديد الأغراض هي تقنية رؤية حاسوبية لتحديد مواقع الكائنات في الصور أو مقاطع الفيديو.** |
| TensorFlowLitelibrary | **هي مكتبة مفتوحة المصدر تجعل من السهل على المطورين تصميم وبناء وتدريب نماذج التعلم العميق.** |
| Mobilenet | MobileNet **عبارة عن شبكة عصبية تلافيفية تم تطويرها بواسطة** Google **والتي تم تدريبها على مجموعة بيانات** ImageNet**، والتي تستخدم بشكل رئيسي لتصنيف الصور الى فئات وتقدير الهدف.** |
| Yolo | **هو نظام متطور للكشف عن الأشياء في الوقت الحقيقي.** |
| ssd**\_**Mobilenet | **نموذج اكتشاف الكائن الذي يحسب المربع المحيط بالمخرج وفئة الغرض من الصورة المدخلة.** |
| Posenet | **هي تقنية للكشف عن الوضعيات في الوقت الفعلي والتي يمكنك من خلالها اكتشاف أوضاع البشر في الصورة أو الفيديو.** |
| Text**-**To**-** Speech | **تقنية تمكن من تحويل النص إلى أصوات كلامية مقلدة للصوت البشري.** |
| Detectron | عبارة عن قاعدة تعليمات برمجية عالية الأداء للكشف عن الكائنات، **تم انشاؤها بواسطة فيسبوك.** |
| DeepLearing  **التعلم العميق** | التعلم العميق هو أسلوب في الذكاء الاصطناعي (AI) يعلم أجهزة الكمبيوتر كيفية معالجة البيانات بطريقة مستوحاة من الدماغ البشري. يمكن لنماذج التعلم العميق التعرف على الأنماط المعقدة في الصور والنصوص والأصوات وغيرها من البيانات لإنتاج رؤى وتنبؤات دقيقة. |

**فهرس المحتويات**

المحتويات

# **الفصل الأول**

**الإطار العام للمشروع**

## 

## مقدمة

مع التقدم الكبير في التكنولوجيا والذكاء الاصطناعي، قمنا بتطوير نظام ذكي باسم ODTV (Object Detection To Voice)لمساعدة فئة خاصة من الأشخاص، وهم المكفوفين. يعتمد هذا النظام على تقنيات الذكاء الاصطناعي ويعمل على الهواتف المحمولة، مما يجعله سهل الاستخدام ويحل محل الأدوات الذكية الأخرى التي قد تكون صعبة الحصول عليها أو استخدامها بواسطة المكفوفين.

**1-2 فكرة المشروع:**

المشروع يهدف إلى تطوير نظام يساعد المستخدم المكفوف على التحرك في اتجاهات مختلفة بأستخدام كاميرا الهاتف لالتقاط صور حية. يتم تحليل هذه الصور بشكل فوري، وتفسير محتوياتها وتحديد الأشياء الموجودة ضمنها. يتم تحويل هذه المحتويات إلى نصوص قابلة للقراءة والسماع، مما يتيح للمستخدم المكفوف تجربة أفضل وزيادة استقلاليته.

**1-3 أدوات المشروع:**

بعد أن تم تحديد فكرة المشروع تم البحث عن أفضل الأدوات والتقنيات واختيار ما يناسبنا منها من اجل تنفيذ فكرة التطبيق والحصول عل أفضل أداء، ومن هذه الأدوات نذكر:

Dart: هي لغة برمجة مصممة لتطوير تطبيقات الويب وتطبيقات الاندرويد وتطبيقات IOS, هذه اللغة من تطوير شركة جوجل.

Flutter: هو برنامج مفتوح المصدر لمجموعة أدوات تطوير البرمجيات, تم تطويره من قبل جوجل لبناء واجهات المستخدم لتطبيقات نظام الاندرويد.

Android Studio: منصة لكتابة التطبيقات تسهل على المطورين كتابة الشيفرة المصدرية لتطبيقات اندرويد, كما تسمح للمطور بمعاينة هيئة تطبيقه على مختلف قياسات الشاشات بشكل فوري خلال التطوير, وتسهل تطوير التطبيقات متعددة اللغات.

## 1- 4مشكلة المشروع:

تكمن المشكلة في صعوبة تعامل المكفوفين مع البيئة المحيطة عند التحرك ، لذلك هم دائما بحاجة الى وسيلة تساعدهم على المشي والتحرك دون الاصطدام بشيء او التعرض لأي نوع من المخاطر في البيئة المحيطة وهذه الوسائل قد تكون مثل عصا او كلب مدرب، بالإضافة لأنهم بحاجة للتأقلم مع وضعهم بسبب انه في اغلب الأحيان من الصعب الخضوع للعمليات الجراحية العينية وتكاليفها المرتفعة.

وبما ان العالم اليوم أصبح الجميع يمتلكون هواتف محمولة كان لابد من إيجاد طريقة نستطيع من خلالها مساعدة المكفوفين بأفضل صورة وبأقل التكاليف باستخدام الهواتف المحمولة المتوفرة بين ايديهم وتوفر تقنيات الذكاء الاصطناعي.

## 1-5أهمية المشروع:

يعتبر المشروع ذو أهمية بالغة نظرًا لأنه يقدم نظامًا مساعدًا يسهل حركة المكفوفين في جميع الاتجاهات دون الحاجة للقلق من المخاطر، ودون الحاجة إلى أدوات تقليدية مثل العصي. يتيح النظام للمكفوفين التحكم في حركاتهم بشكل أفضل وبكل سهولة، مما يزيد من استقلاليتهم ويسهم في تحسين جودة حياتهم. بالإضافة إلى ذلك، يتميز النظام بسهولة استخدامه وتوفره على الهواتف المحمولة، مما يجعله بديلاً فعّالًا للحلول التقليدية والأجهزة الغالية.

## 1-6 الدراسات السابقة:

## في عام(2015),تم ابتكار نظام ذكي باسم "SASB" (Smart Assist System for Blind people) يعتمد على سترة ذكية مزودة بمجسات ويستخدم الأوامر الصوتية لتوجيه حركة المكفوف وتحديد المسافة بينه وبين العوائق. المرجع: 1.

## في عام (2022), قام فريق من الباحثين التونسيين بابتكار نظام يستخدم نظارة مجهزة بكاميرات ويعتمد على الذكاء الصناعي لتحويل صور الفيديو إلى إشارات صوتية. المرجع: 2

## 1-7 محتويات التقرير:

تم تنظيم هذا التقرير في خمسة فصول الفصل الأول مقدمة تتضمن الهدف من المشروع وادواته ومشكلة المشروع واهمية المشروع. ويوضح الفصل الثاني الدراسة التحليلية والتصميمية من تحليل النظام والمتطلبات وطريقة جمعها وصولا إلى المخططات الخاصة بالمشروع. كما يحتوي الفصل الثالث على الدراسة النظرية وما تم استخدامه بالمشروع، وكما تضمن الفصل الرابع اللغات والتقنيات التي تم استخدامها لإنجاز هذا المشروع ومميزات كل منها، والخامس واجهات التطبيق بالإضافة للنتائج والخاتمة.

# 

# **الفصل الثاني**

**الدراسة التحليلية والتصميمية**

## 1-2 الدراسة التحليلية

**1-1-2 تمهيد**

نستعرض في هذا القسم خصائص النظام والخدمات التي يقدمها. بالإضافة إلى المتطلبات الوظيفية والمتطلبات غير الوظيفية و مخطط عمل النظام.

2-**2 جمع المتطلبات**

في هندسة النظم وهندسة البرمجيات، تحليل المتطلبات يركز على المهام التي تحدد الاحتياجات أو الشروط اللازمة لتنفيذ مشروع، وكذلك تحليل متطلبات النظام أو البرنامج وتوثيقها والتحقق منها وإدارتها. تحليل المتطلبات مهم وأساسي لنجاح أو فشل النظام أو البرنامج. يجب توثيق المتطلبات وأن تكون قابلة للتنفيذ والقياس والاختبار ومرتبطة باحتياجات المشروع وكذلك أن تكون مفصلة بالشكل الكافي لتصميم النظام.

**2-2-1 تحليل المشكلة**

**يجب على النظام توفير السرعة والدقة في إعطاء المعلومات لمساعدة المستخدم(الكفيف) على الحركة وممارسة حياته بشكل طبيعي، بالإضافة لأن يكون موثوق وسهل الاستخدام.**

**2-2-2 خصائص النظام**

مستخدمي النظام هم فئة ذوي الاحتياجات الخاصة (المكفوفين).

يجب على النظام أن يقدم الخدمات التالية:

1.السرعة في استقبال وتحليل الصور الملتقطة عن طريق كاميرة الهاتف وفي أي اتجاه.

2. السرعة والدقة في تحديد الأغراض ووصفها للمستخدم(الكفيف) باستخدام تقنيات الصوت.

3. القدرة على الوصول للتطبيق بشكل سريع وسهل من قبل المستخدم(الكفيف) عن طريق استخدامه لاحد الأنظمة المساعدة التي تتوافر في جميع الأجهزة المحمولة مثل مساعد جوجل.

**3-2-2 المتطلبات الوظيفية**

* استخدام هاتف محمول: اذ يجب توفر هاتف شخصي للمستخدم(المكفوف) من أجل تحميل التطبيق والعمل عليه.
* استخدام كاميرة الهاتف من أجل التقاط الصور: يجب على المستخدم تمكين كاميرة الهاتف عند تحميل التطبيق حتى يتمكن المكفوف من استخدام التطبيق.
* استخدام سماعة الهاتف لسماع اسماء الأغراض: يجب على المستخدم تمكين سماعة الهاتف ورفع الصوت من اجل ان يقوم التطبيق بنطق أسماء الأغراض بعد التقاطها بالكاميرة ,او من الممكن استخدام سماعات الراس.

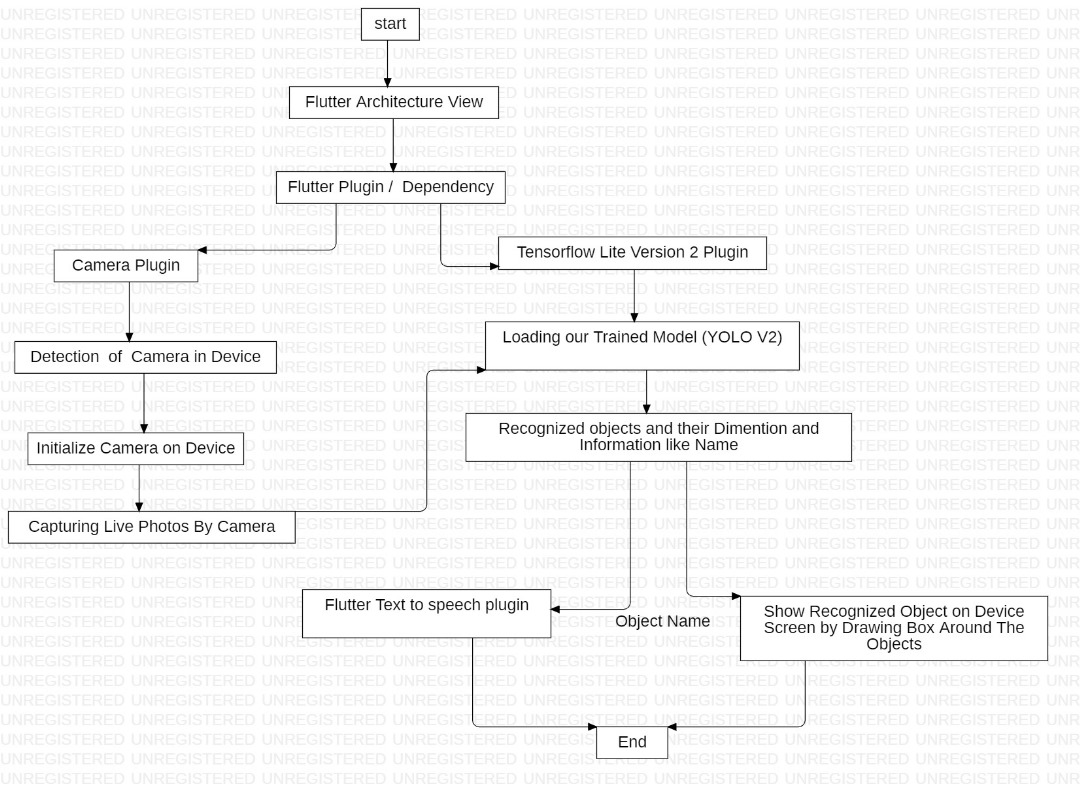
**2-2-4 المتطلبات غير الوظيفية**

* سرعة الاستجابة: اذ ان النظام بحاجة ليكون سريع في الاستجابة بعد التقاط صور من الكاميرة وسرعة في تحليلها ووصفها ونطقها باستخدام الصوت.
* سهولة الاستخدام:

النظام له واجهة واحدة فقط وهي واجهة مرتبطة مع كاميرة الهاتف للمستخدم(المكفوف) حيث ان المستخدم(المكفوف) لن يستطيع التعامل مع تعدد الواجهات لذا عند الوصول للتطبيق سوف تظهر واجهة الكاميرة مباشرة.

**الدراسة التصميمية**

تم تصميم مخطط عمل النظام والذي يصف تسلسل عمل النظام:



# **الفصل الثالث**

**الدراسة النظرية**

## 3-1 الذكاء الاصطناعي:

**1-1-3 التمهيد**

الذكاء الاصطناعي: هو سلوك وخصائص معينة تتسم بها البرامج الحاسوبية، تجعلها تحاكي القدرات الذهنية البشرية وأنماط عملها. من أهم هذه الخصائص القدرة على التعلم والاستنتاج ورد الفعل على أوضاع لم تبرمج في [الآ](https://ar.wikipedia.org/wiki/%D8%A2%D9%84%D8%A9)لة.

تأسس هذا المجال على افتراض أن مَلَكة الذكاء يمكن وصفها بدقة بدرجة تمكن الآلة من محاكاتها. كان وما زال الذكاء الاصطناعي سبباً لأفكار شديدة التفاؤل، ولقد عانى نكسات فادحة عبر التاريخ، واليوم أصبح جزءاً أساسياً من صناعة التكنولوجيا، حاملاً عبء أصعب المشاكل في علوم الحاسوب الحديثة.

إن بحوث الذكاء الاصطناعي من الأبحاث عالية التخصص والتقنيَّة، تتمحور المجالات الفرعية للذكاء الاصطناعي حول مشاكل معينة، وتطبيق أدوات خاصة وحول اختلافات نظرية قديمة في الآراء. تتضمن المشاكل الرئيسية للذكاء الاصطناعي قدراتٍ مثل التفكير المنطقي والمعرفة والتخطيط والتعلم والتواصل والإدراك والقدرة على تحريك وتغيير الأشياء. مرجع 3

**2-1-3 تفسير الصور**

في مجال تفسير الصور ( (Image Interpretation هو عملية فحص الصورة من خلال النظر في موقعها ومداها واستخدام سمات معينة للصورة، إلى جانب معرفة الأشياء وظروف الصورة، لتحديد الميزات والحكم على أهميتها.

## 3-1-3عناصر تفسير الصور

بمجرد الحصول على صورة، من المهم جدًا التأكد من دقة تفسير الصورة للغرض المقصود منها.

يستخدم تفسير الصور وتحليلها بعض المؤشرات أو المفاتيح لتحديد الميزات. بعضها مثل الشكل والحجم واللون والظل والملمس والنمط والارتباط والموقع والوقت والدقة

1. الشكل(shape): كما يوحي الاسم، فإن شكل الميزة يساعد في تحديد الميزة.
2. الحجم(size): يلعب حجم الميزة بالنسبة إلى الميزة القريبة دورًا مهمًا في تحديد الميزة بنجاح.
3. الظل(shadow): يساعد ظل الميزة في تحديد حدود الميزة. من شأن الجسم الكبير أن يلقي ظلًا كبيرًا مقارنة بالأجسام الأصغر. يستخدم الظل أيضًا لقياس ارتفاع الجسم
4. النمط(pattern): الترتيب المكاني للميزات في تنسيق معين هو نمط.
5. الارتباط (Association): العلاقة بين السمات المختلفة في المنطقة.
6. الموقع(location): الموقع هو وجود ميزة في موقع جغرافي معين.
7. الوقت(time): يمكن أن يوفر التغيير الزمني في الميزة خلال فترة زمنية الكثير من المعلومات لتفسير الصور.
8. الدقة (Resolution): تعد دقة الصورة أيضًا جانبًا رئيسيًا في تفسير الصور.

مرجع 4

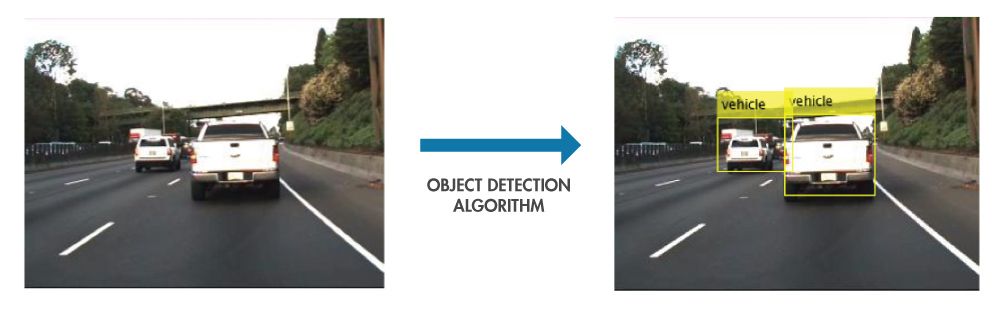
## 3-2 تحديد الأغراض :

**1-2-3تمهيد:**

ويتم عن طريق تقنية رؤية حاسوبية لتحديد مواقع الكائنات في الصور أو مقاطع الفيديو ,حيث تستفيد خوارزميات الكشف عن الكائنات عادةً من التعلم الآلي أو التعلم العميق لتحقيق نتائج ذات معنى.

**2-2-3لماذا يهم تحديد الأغراض:**

يعد اكتشاف الأشياء تقنية أساسية وراء أنظمة مساعدة السائق المتقدمة (ADAS) التي تمكن السيارات من اكتشاف ممرات القيادة أو إجراء اكتشاف المشاة لتحسين السلامة على الطريق. يعد اكتشاف الكائنات مفيدًا أيضًا في تطبيقات مثل المراقبة بالفيديو أو أنظمة استرجاع الصور ولذلك تم الاعتماد على هذه التقنية في بناء نظام مساعد للمكفوفين لأنه أيضا يمكن من اكتشاف الأغراض حول المستخدم.



**3-2-3 كيف تعمل:**

اكتشاف الأشياء غالباً يتم باستخدام التعلم العميق ويمكن استخدام مجموعة متنوعة من التقنيات لإجراء اكتشاف الكائنات. تتعلم الأساليب الشائعة القائمة على التعلم العميق والتي تستخدم الشبكات العصبية التكرارية(التحليلية) (CNNs)، مثل R-CNN وYOLO v2، وتقوم باكتشاف الكائنات داخل الصور تلقائيًا.

## 3-3 تحويل النص إلى صوت:

**1-3-3 تحويل النص الى صوت (Text-to-Speech):**

تقنية تحويل النص إلى كلام (TTS) هي تكنولوجيا مساعدة لقراءة النصوص الرقمية، تمكن من تحويل النص إلى أصوات كلامية مقلدة للصوت البشري يطلق عليها أحيانًا تقنية "القراءة بصوت عالٍ", يمكن من خلالها اخذ الكلمات الموجودة على جهاز كمبيوتر أو أي جهاز رقمي آخر وتحويلها إلى صوت.

**2-3-3 كيف تعمل تقنية تحويل النص الى صوت:**

تعمل تقنية تحويل النص إلى صوت (TTS) مع كل الأجهزة الرقمية الشخصية تقريبًا، بما في ذلك أجهزة الكمبيوتر والهواتف الذكية والأجهزة اللوحية. يمكن قراءة جميع أنواع الملفات النصية بصوت عالٍ، بما في ذلك مستندات Word وPages. حتى صفحات الويب على الإنترنت يمكن قراءتها بصوت عالٍ.

يتم إنشاء الصوت في تحويل النص إلى صوت (TTS) بواسطة الكمبيوتر، وعادةً ما يمكن تسريع سرعة القراءة أو إبطائها. تختلف جودة الصوت، لكن بعض الأصوات تبدو بشرية. حتى أن هناك أصواتًا يتم إنشاؤها بواسطة الكمبيوتر تبدو وكأنها أطفال يتحدثون.

نحن بحاجة لاستخدام هذه التقنية في بناء نظامنا حيث تقوم العديد من أدوات تحويل النص إلى صوت (TTS) بتمييز الكلمات أثناء قراءتها بصوت عالٍ. يتيح ذلك للمستخدمين من سماع مسمى الغرض المكتشف بشكل واضح.

**3-3-3 Flutter\_tts:**

أحد الإضافات الخاصة بـ Flutter لتحويل النص الى كلام وهي مدعومة على أنظمة iOS، وAndroid، والويب، وmacOS.

## 3-4 بيئة العمل flutter :

**1-4-3 تمهيد:**

هي حزمة تطوير برمجية من شركة غوغل مخصصة لتطوير تطبيقات الهواتف الذكية العاملة بنظام أندرويد و ال (IOS ) و فوشيا (نظام تشغيل جديد من غوغل) ومؤخراً يتم تطويرها لتصبح قادرة على تطوير تطبيقات الويب حيث تركز على تجربة المستخدم والمطور في نفس الوقت وهذا لآنها توفر إطار عمل شامل بلغة DART مخصص لرسم الواجهات بجودة عالية وتطبيقات بأداء ممتاز وفي نفس الوقت تعطي للمطور الأدوات التي تجعله يبني تطبيقات كاملة في اقل وقت.

**2-4-3 ميزات flutter:**

• تعتمد على Widgets وهي تعطي حرية كبيرة في التحكم وسلاسة في التحكم في عناصر البرنامج وتنظيم تشغيل التطبيق على جميع الواجهات التي يعمل عليها

•على عكس اللغات الأصيلة هي لا تدعم برمجة تطبيق يعمل على نظام التشغيل android اوعلى نظام تشغيلIOS فقط ولكن تدعم الاثنين معا ذلك يوفر وقت كبير لتعلمها بشكل شامل.

•تتيح التحكم بشكل كامل بجميع العناصر الداخلية في التطبيق , كما أن الأداء الخاص بالتطبيق يكون مستقر وقوي.

•تكلفة منخفضة مقارنة بباقي لغات البرمجة, وبالتأكيد التكلفة تعتبر عنصر أساسي من عناصر الحكم على برنامج أو لغة برمجية معينة او بيئة.

•تعتمد في تصميم الواجهات على material design التي توفر تحكم سلس جداً لجميع عناصر واجهة المستخدم ويحل مشكلة التوافق مع جميع أنواع الهواتف وأحجام الشاشة دون أن يسبب مشاكل في الواجهة أو في الأيقونات او في الشكل العام للتطبيق وهذه تعتبر من اهم مميزاتها.

## 3-5 لغة DART :

**1-5-3 تعريف :**

هي لغة من لغات البرمجة متعددة الاستخدامات مفتوحة المصدر تدعم البرمجة الموجهة بالأغراض OOP وقد طورتها شركة Google . والهدف من هذه اللغة هو إيجاد وسيلة يمكن عن طريقها كتابة كود واحد يعمل على منصات مختلفة مثل الويب ونظام الأندرويد ونظام ios دون أن يكون هناك حاجة إلى إعادة الكتابة أو إجراء تغيير في الكود.

**2-5-3 الأمور التي تميزت بها لغة البرمجة دارت Dart:**

* يمكن تعلم هذه اللغة دون الحاجة إلى تعلم لغة برمجية أخرى.
* تدعم البرمجة الكائنية .
* يمكن من خلال هذه اللغة تطوير تطبيقات الويب.
* تستطيع برمجة تطبيقات الأندرويد ونظام ios بنفس الكود.
* تمتلك لغة دارت أقوى وأذكى محررات البرمجة مع دعم كبير الكتابة وتصحيح الأخطاء.

## 3-6 نماذج التدريب التي جرت عليهم الدراسة :

**Mobilenet 1-6-3:**

**هو نموذج مفتوح المصدر بواسطة Google ومصمم لتدريب المصنفات. ويستخدم تلافيفات عميقة لتقليل عدد المعلمات بشكل كبير مقارنة بالشبكات الأخرى، مما يؤدي إلى شبكة عصبية عميقة خفيفة الوزن وهو من النماذج جيدة التنظيم لتطبيقات الرؤية المحمولة والمضمنة. تعتمد هذه الفئة من النماذج على بنية مبسطة تستخدم تلافيفات قابلة للفصل من العمق لبناء شبكات عصبية عميقة خفيفة الوزن.**

**SSD\_Mobilenet 2-6-3:**

**نموذج mobilenet Single-Shot multibox Detection (SSD)** تُستخدم **الشبكات العصبية التلافيفية لتطوير نموذج يتكون من طبقات متعددة لتصنيف كائنات معينة في إحدى الفئات المحددة.**

**يتم اكتشاف هذه الكائنات باستخدام خرائط ميزات عالية الدقة أصبحت ممكنة بفضل التطورات الحديثة في التعلم العميق باستخدام معالجة الصور.**

**Mobilenet SSD هو نموذج للكشف عن الكائنات يحسب المربع المحيط بالمخرج وفئة الكائن من الصورة المدخلة. يستخدم نموذج الكشف عن الكائنات باستخدام كاشف الطلقة الواحدة (SSD) Mobilenet كعمود فقري ويمكنه تحقيق اكتشاف سريع للكائنات مُحسّن للأجهزة المحمولة.**

**YOLO(v2,v5) 3-6-3:**

**YOLO هو اختصار ل "you only look once" أي تم تعريف مصطلح YOLO على أنه "أنت تنظر مرة واحدة فقط". تقوم خوارزمية YOLO بالكشف عن الكائنات في الوقت الفعلي باستخدام شبكة عصبية تلافيفية (CNN).**

**PoseNet Model 4-6-3:**

**PoseNet هي تقنية للكشف عن الوضعيات في الوقت الفعلي والتي يمكنك من خلالها اكتشاف أوضاع البشر في الصورة أو الفيديو.**

**إذ أنه يعمل في كلتا الحالتين كوضع فردي (اكتشاف وضعية بشرية واحدة) واكتشاف متعدد الأوضاع (اكتشاف وضعية متعددة للبشر).**

**تكتشف أداة PoseNet نقاط الجسم الرئيسية في الشخصيات البشرية باستخدام نموذج PoseNet.**

**يعد كاشف الشخص الواحد أسرع وأبسط ولكنه يتطلب وجود شخص واحد فقط على الشاشة، في حين يمكن للكاشف متعدد الأشخاص اكتشاف العديد من الأشخاص، ولكنه أبطأ قليلاً من خوارزمية الشخص الواحد.**

**بكلمات بسيطة، Posenet هو نموذج TensorFlow للتعلم العميق يسمح لك بتقدير وضعية الإنسان عن طريق اكتشاف أجزاء الجسم مثل المرفقين والوركين والمعصمين والركبتين والكاحلين وتشكيل هيكل عظمي لوضعيتك من خلال ضم هذه النقاط**

**تم تدريب PoseNet على بنية MobileNet.**

**اين يستخدم نموذج PoseNet:**

1. **يستخدم في فلاتر سناب شات حيث ترى اللسان والجوانب واللمحات والوجوه السخيفة.**
2. **تطبيقات اللياقة البدنية مثل cult التي تستخدم للكشف عن تمريناتك.**
3. **تستخدم Instagram Reels المشهورة جدًا اكتشاف الوضعية لتزويدك بميزات مختلفة لتطبيقها على وجهك والمناطق المحيطة بك.**
4. **ألعاب افتراضية لتحليل لقطات اللاعبين.**

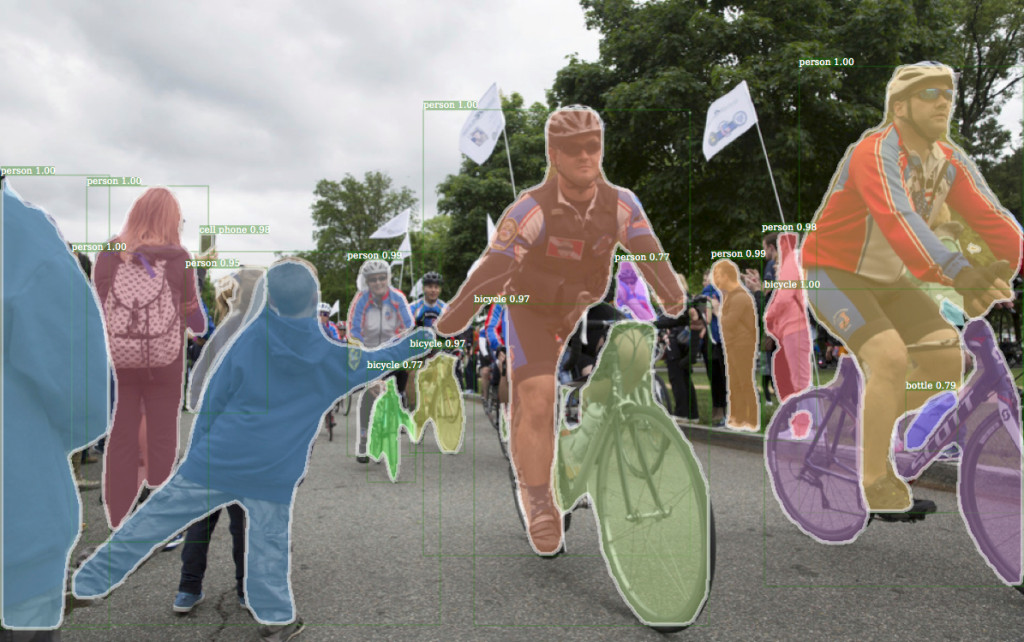
**يمنحنا Posenet إجمالي 17 نقطة أساسية يمكننا استخدامها، بدءًا من أعيننا وحتى آذاننا وحتى الركبتين والكاحلين.**

**Detectron 5-6-3:**

عبارة عن قاعدة تعليمات برمجية عالية الأداء للكشف عن الكائنات، وتغطي المربع المحيط ومخرجات تجزئة مثيل الكائن.

تم إنشاء Detectron بواسطة Facebook AI Research (FAIR) لدعم التنفيذ والتقييم السريع لأبحاث رؤية الكمبيوتر الجديدة. ويتضمن تطبيقات لخوارزميات الكشف عن الكائنات التالية:

* Mask R-CNN
* RetinaNet
* Faster R-CNN
* RPN
* Fast R-CNN
* R-FCN
* يمكن استخدام Detectron خارج الصندوق لاكتشاف الكائنات بشكل عام أو تعديله للتدريب وتشغيل الاستدلال على مجموعات البيانات الخاصة بك. إنه مكتوب بلغة Python وسيتم تشغيله بواسطة إطار التعلم العميق PyTorch 1.0. (مرجع 2)



# 

# **الفصل الرابع**

**التطبيق العملي والنتائج**

## 4-1 التطبيق العملي:

**4-1-1 التقنيات والأدوات ولغات البرمجة المستخدمة :**

* **أطر العمل واللغات المستخدمة**

جرى العمل على تطبيق الموبايل الخاص بالمشروع باستخدام لغة DART ، وهي لغة من تطوير google , تحديداً حزمة التطوير البرمجية Flutter

* **بيئة التطوير**

Android Studio 2022

* **تصميم الواجهات**

لا يوجد سوى شاشة إظهار خرج الكاميرا وقد تم الاتصال والتعامل مها عن طريق الحزمةcamera

* **تخزين البيانات والتعامل معها**

تم استخدام عدة نماذج مُدربة مسبقاً, على مجموعة بيانات نصية وصورية ضخمة.

**4-2-1 إنشاء الواجهات الخاصة بالتطبيق:**

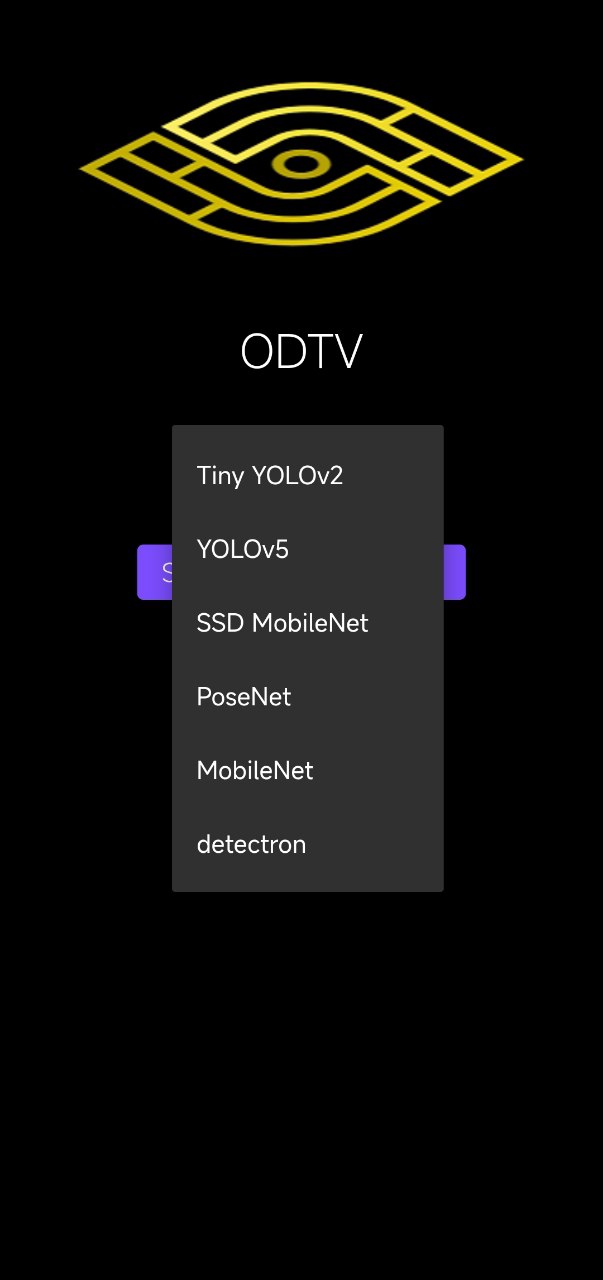
وفقاً للمتطلبات الوظيفية للتطبيق وكونه موجه للمكفوفين ,فإنه ليس بحاجة لتصميم واجهات ،وتحقيقاً للمتطلبات غير الوظيفية من السرعة والدقة ، تم العمل على عدة مراحل:

**4-2-2** **المرحلة الأولى** **:** **واجهات تجريب وتحديد النموذج المناسب**

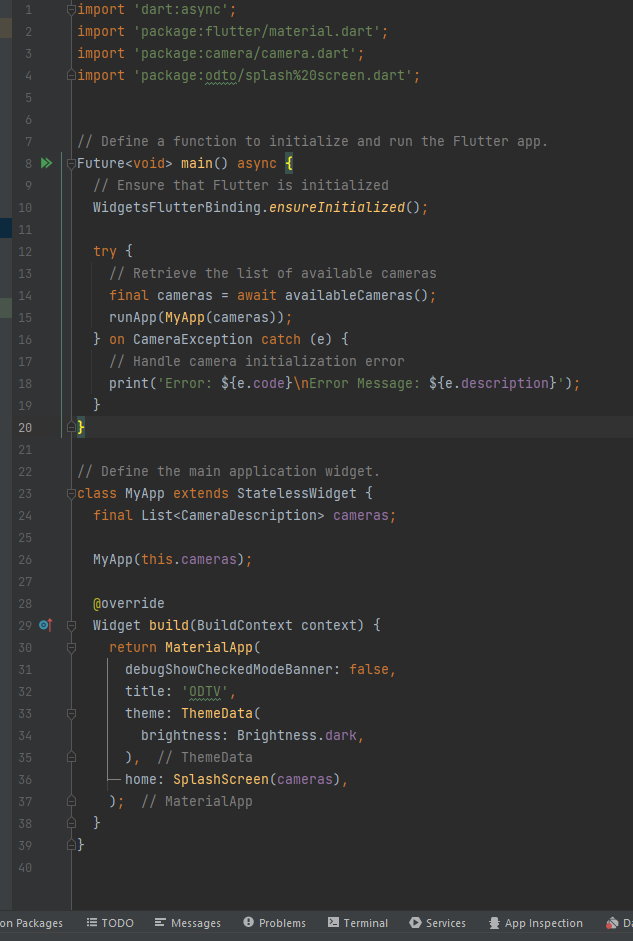
باستخدام حزمة تطوير البرمجيات Flutter (Tensorflowlite) والتي تمكننا من التعامل مع النماذج المدربة بإطار العمل TensorFlow ,وقد تم العمل على مرحلتين :

* **مرحلة تجريب وتحقق من النتائج:**

وقد تم فيها استخدام واستدعاء العديد من النماذج ,كما وقد تم وضع بعض الواجهات لتسهيل عملية التجريب والتحقق ,يظهر الشكل التالي الواجهة التجريبية الأولى وقائمة النماذج :

 الشكل

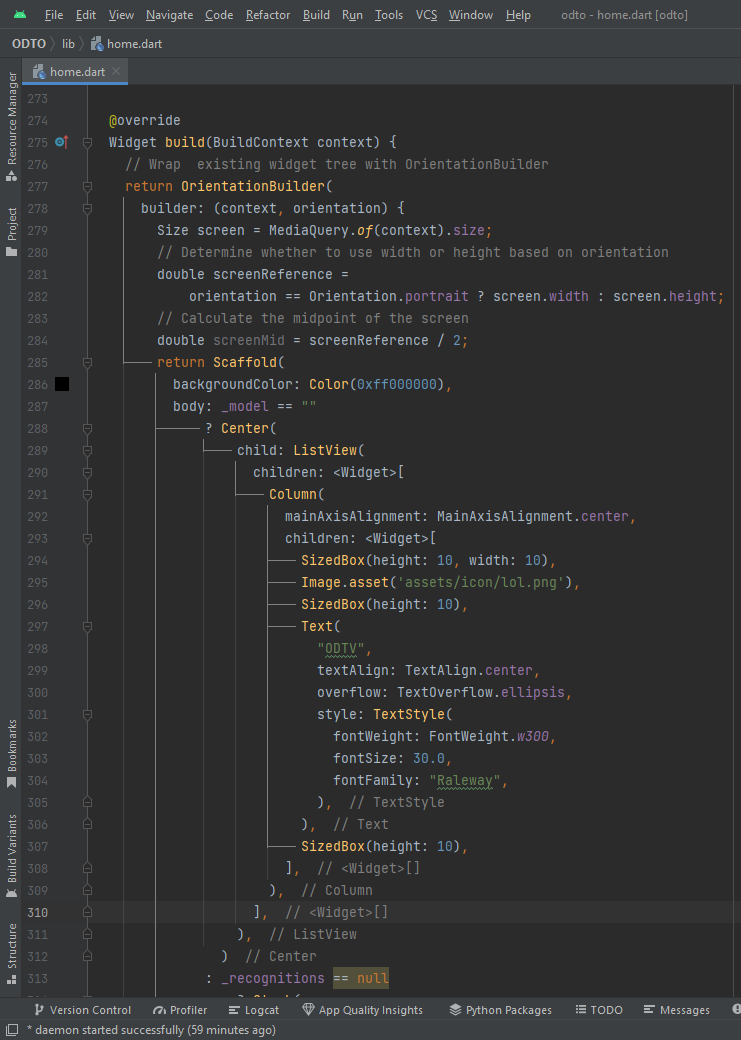
يجدر ذكر أن حزمة TFLITE لتقوم بقراءة النموذج واستدعاءه تحتاج لمصفوفة من البيانات حتى يتم معالجتها من قبل النموذج نقوم بتعريفها وتمريرها من خلال الصف الرئيسي لصفحة Home كما موضح في الجزء البرمجي :

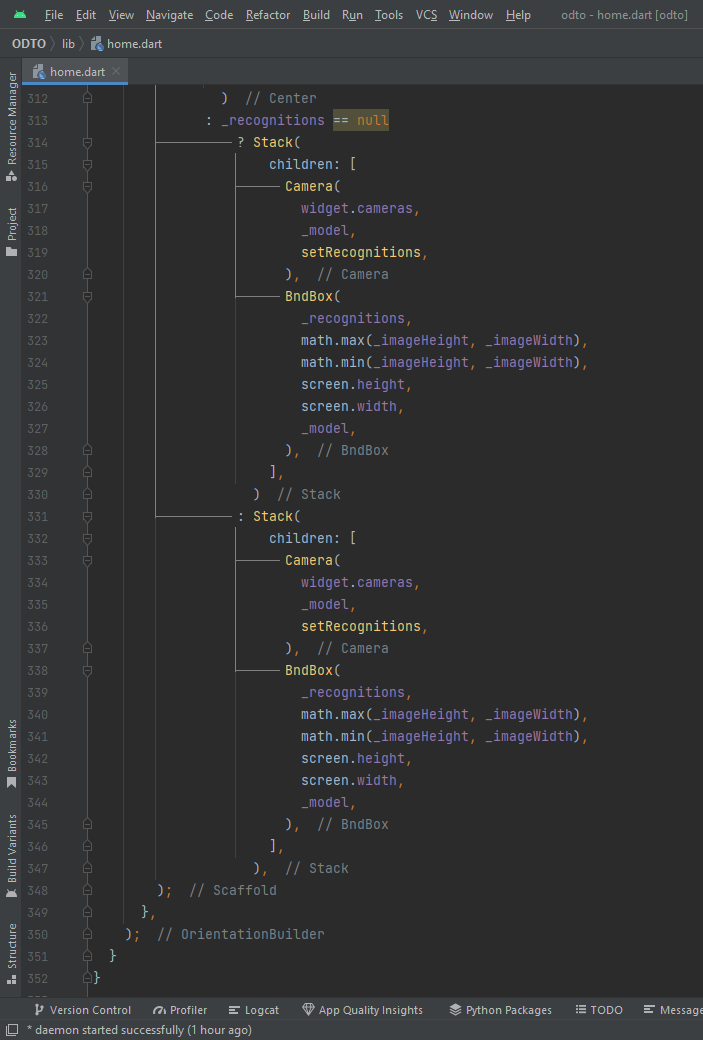


الشكل

(ملاحظة: في الجزء البرمجي الموضح سابقاً تم التمرير لsplash screen نيابة عن الhome لأن الجزء الموضح تم أخذ صورة عنه في النتيجة النهائية للتطبيق الحالي).

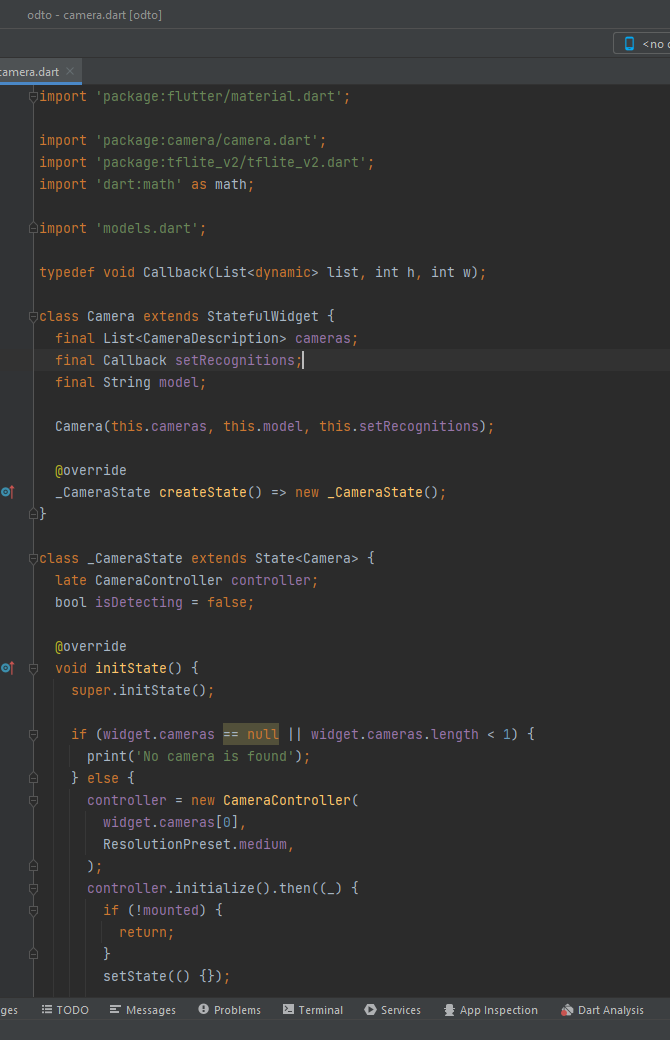
في الواجهة HOME الموضحة بالشكل () والتي تم بناءها برمجياً وفق:





الشكل

يظهر الجزء البرمجي أن الواجهة تحتوي جزئية خاصة بالكاميرا إذ يتم اختبار وجود نموذج ,إن وجد حينها نستطيع استدعاء الواجهة الخاصة بالكاميرا في هذه الواجهة نستخدم الحزمة TFLite\_v2 ويتم عن طريقها معالجة البيانات حسب النموذج كما هو موضح :



الشكل

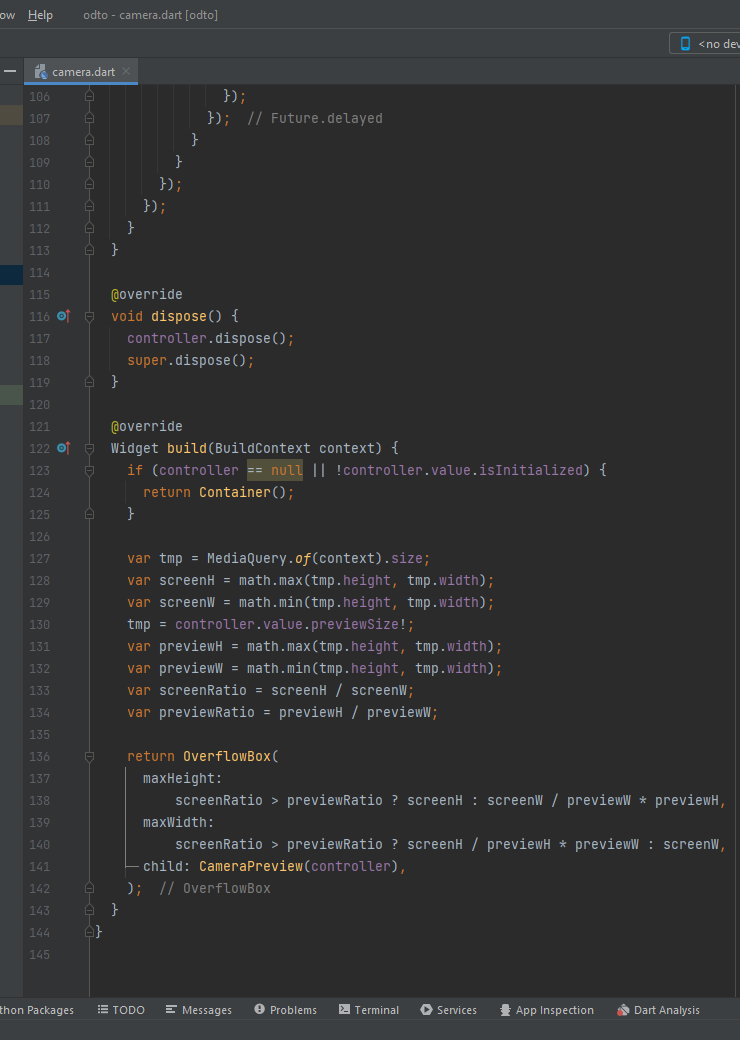


الشكل

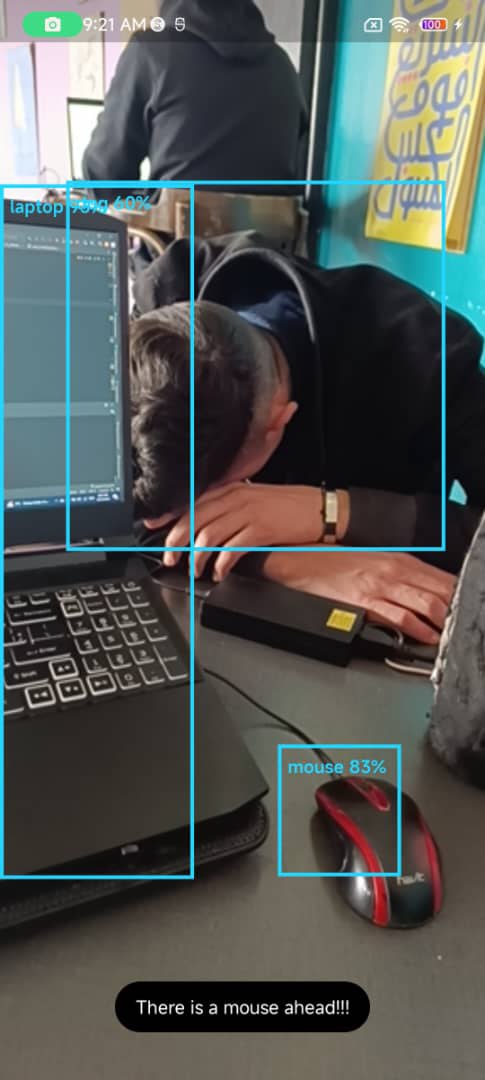


الشكل

كما هو موضح في الأكواد السابقة يتم معالجة المعطيات واستدعاء الإجرائية المناسبة للتحقق حسب النموذج المرسل ومن ثم بناء الواجهة الخاصة بcamera وفق الآتي:



الشكل

بالعودة للصف home نلاحظ أيضاً وجود جزئية مسمات بBND Box وهي تختص بالرسم الظاهر في الكاميرا من مربعات ترسم حول الأشياء المكتشفة وقد قم بناءها في هذه المرحلة لتظهر بلون ازرق مضافة له نسبة التعرف واسم الغرض :

الشكل

يجدر الحديث في هذه المرحلة عن اعتماد النموذج المبني عن طريق YOLOV2 كونه الأكثر كفاءة كتزامن ما بين اكتشاف الأغراض والقراءة الصوتية ،ونتيجة لهذه المرحلة ستقسم الحزم والنماذج إلى حزم ونماذج قد تم استخدامها تجريبياً وجزم ونماذج معتمدة .

*الحزم والنماذج المستخدمة المعتمدة:*

1. *TensorFlow Lite version2:*

**TensorFlow Lite والتي يشار إليها غالبًا باسم TFLite، هي حزمة مفتوحة المصدر تم تطويرها بواسطة Google لنشر نماذج التعلم الآلي على الأجهزة المتطورة، كما أنها شائعة لحالات الاستخدام التي تتراوح بين تصنيف الصور واكتشاف الكائنات والتعرف على الكلام ومهام اللغة الطبيعية .**

**وهو إصدار خفيف الوزن مصمم خصيصًا لمنصة الهاتف المحمول والأجهزة المدمجة.**

**توفر TensorFlow Lite مجموعة من الأدوات التي تتيح التعلم الآلي على الجهاز من خلال السماح للمطورين بتشغيل نماذجهم المدربة على الأجهزة المحمولة والمدمجة وأجهزة إنترنت الأشياء وأجهزة الكمبيوتر. وهو يدعم منصات مثل Linux وAndroid وiOS وMCU المضمنة.**

**يوفر حل التعلم الآلي للجوال مع زمن وصول منخفض وحجم ثنائي صغير.**

**عند استخدام TFLite لرؤية الكمبيوتر، يمكنك إجراء الاستدلال على الجهاز لأشياء مثل تصنيف الصور أو اكتشاف الكائنات.**

1. *Flutter TTS:*

**تستخدم هذه الحزمة Flutter\_tts للتفاعل مع الوظائف الأصلية. تستخدم TextToSpeech لنظام Android، وAVSpeechSynthesizer لمنصة IOS.**

*3- YOLO(v2):*

**تقوم خوارزمية YOLO بالكشف عن الكائنات في الوقت الفعلي باستخدام شبكة عصبية تلافيفية (CNN).**

**كما يوحي الاسم، تحتاج خوارزمية الكشف عن الكائنات فقط إلى انتشار أمامي واحد عبر الشبكة العصبية. وهذا يعني أن التنبؤ في الإطار بأكمله يتم من خلال تشغيل الخوارزمية مرة واحدة. يتم استخدام CNN للتنبؤ بفئات الاحتمالية المختلفة والحقول الحدودية في وقت واحد.**

*من أهم مزايا نموذج* ***YOLO****:*

1. **السمات الرئيسية لـ YOLO هي السرعة والدقة العالية والقدرة على التعلم.**
2. **من حيث السرعة – قامت هذه الخوارزمية بتحسين سرعة اكتشاف الكائنات في الوقت الفعلي.**
3. **من حيث الدقة العالية – تعطي هذه التقنية نتائج دقيقة مع الحد الأدنى من الأخطاء في الخلفية.**
4. **القدرة على التعلم - تتيح القدرة التعليمية الممتازة لهذه الخوارزمية تعلم تمثيلات الكائنات وتطبيقها في اكتشاف الكائنات.**

*كيف تعمل* ***YOLO****:*

**يمكن شرح تشغيل خوارزمية YOLO باستخدام ثلاث تقنيات على النحو التالي:**

1. **الكتل المتبقية(Residual blocks):**

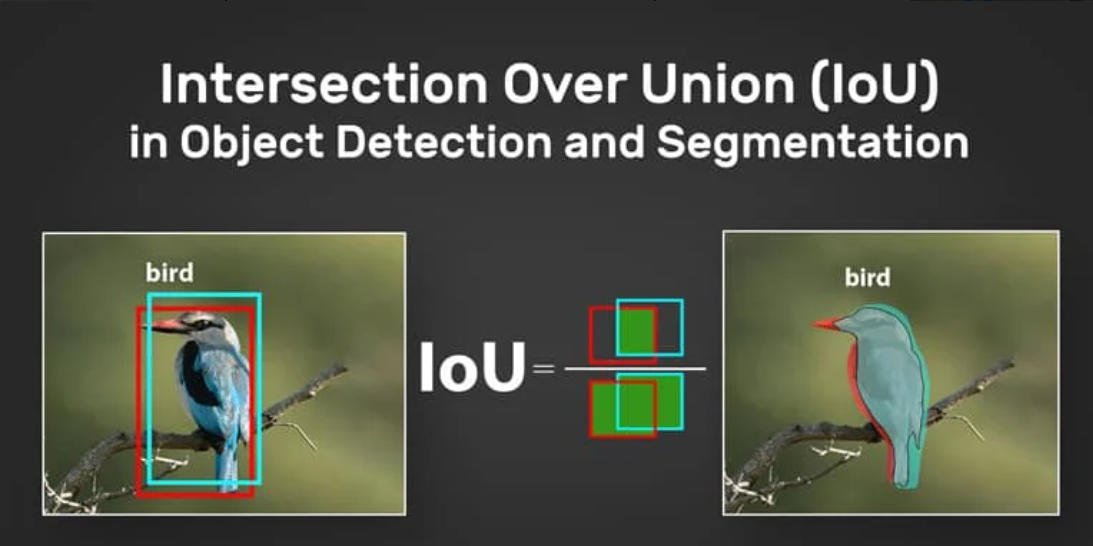
**يتم تقسيم الصورة إلى شبكات مختلفة. كل شبكة لها البعد S x S. تكتشف كل خلية شبكة الكائنات التي تظهر فيها.**

1. **انحدار المربع المحيط(Bounding box regression):**

**المربع المحيط عبارة عن مخطط تفصيلي يبرز كائنًا في صورة مع بعض السمات مثل العرض (border width)، والارتفاع (border height)، والفئة (مثل شخص، سيارة، إشارة مرور، وما إلى ذلك)، والتي يتم تمثيلها بالحرف C. مركز المربع المحيط (bx، will). يستخدم YOLO انحدار المربع المحيطي الفردي للتنبؤ بارتفاع الكائنات وعرضها ومركزها وفئتها.**

1. **التقاطع فوق الاتحاد (IOU):**

**يعد التقاطع عبر الاتحاد (IOU) Intersection over Union أداة لاكتشاف الكائنات تشرح كيفية تداخل المربعات.**

****

**الشكل**

*الحزم والنماذج المستخدمة تجريبياً:*

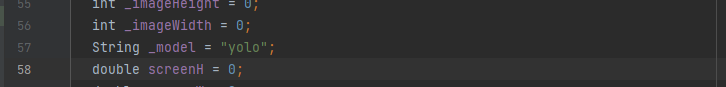
1. **Flutter\_Pytorch:  
    وقد تم استخدام هذه الحزمة للتعامل مع YOLOV5 وقد تم الاستغناء عنها كونها أظهرت نسبة مشاكل أعلى في قراءة النموذج .**
2. **Flutter d2go:**

**وتتميز هذه الحزمة بإمكانية التعامل والدقة في قراءة نموذج detectron2 وهي تتيح التعامل مع الصورة او مع رتل صوري (فيديو)، لم يتم اعتماد هذه الحزمة لأن النموذج الذي تقوم بقراءته يسبب مشاكل عدة في التوافقية مع الصوت وسرعة التحقق من النتائج.**

1. **النماذج المتبقية قد تم ذكرها والتحدث عنها في الدراسة النظرية وقد تم التخلي عنها في النظام كونها ذات مشاكل عالية إما في التوافقية مع الصوت او في دقة وصحة النتائج.  
   علاوة على ما ذكر سابقاً يجدر الذكر بأن النموذج المعتمد حالياً YOLO بإصداراتها هيي نماذج اكتشاف أغراض ذات دقة عالية وحجم متوسط نسبياً للتطبيقات المستعملة لها .**

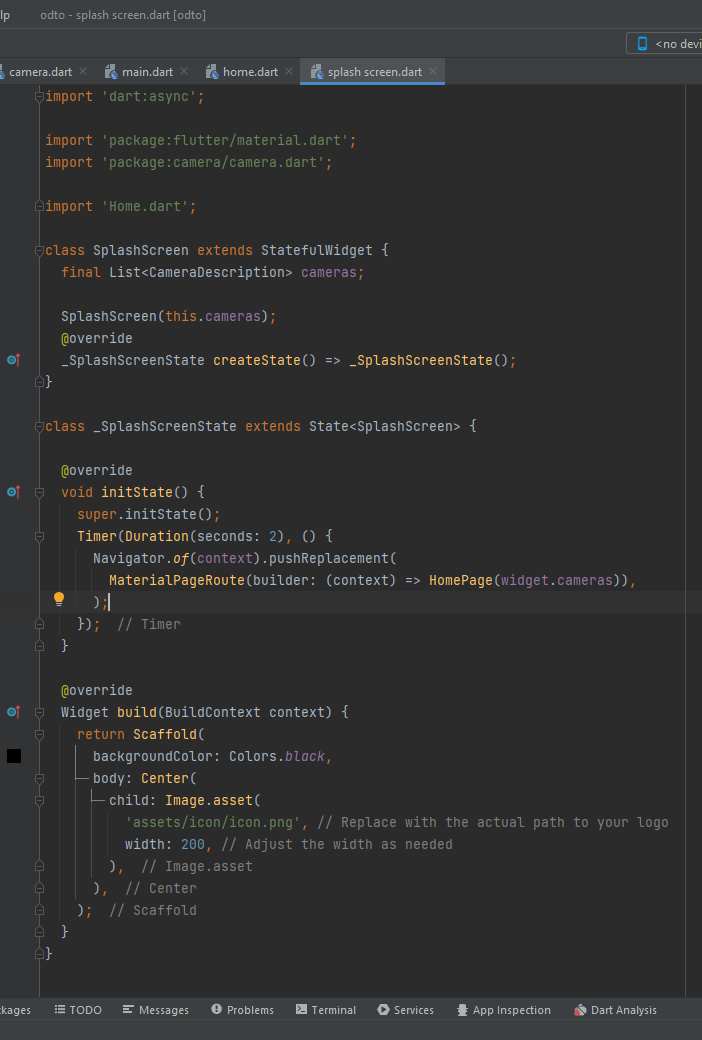
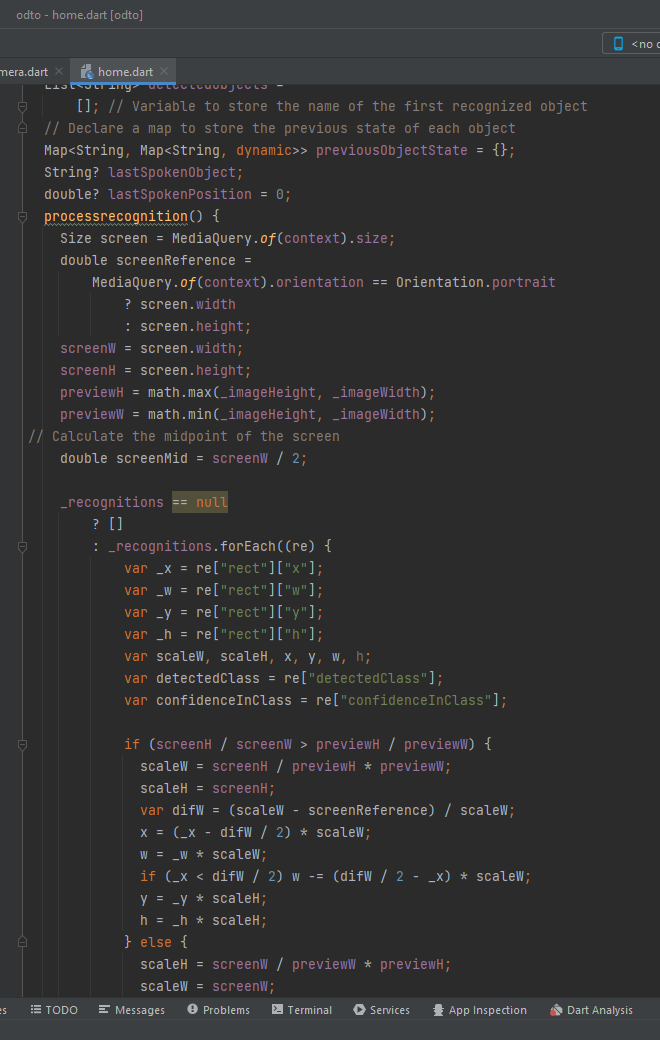
* **مرحلة تحسين القراءة الصوتية وضبط المخرجات:**

في هذه المرحلة قمنا بعدة خطوات لضبط النسخة :

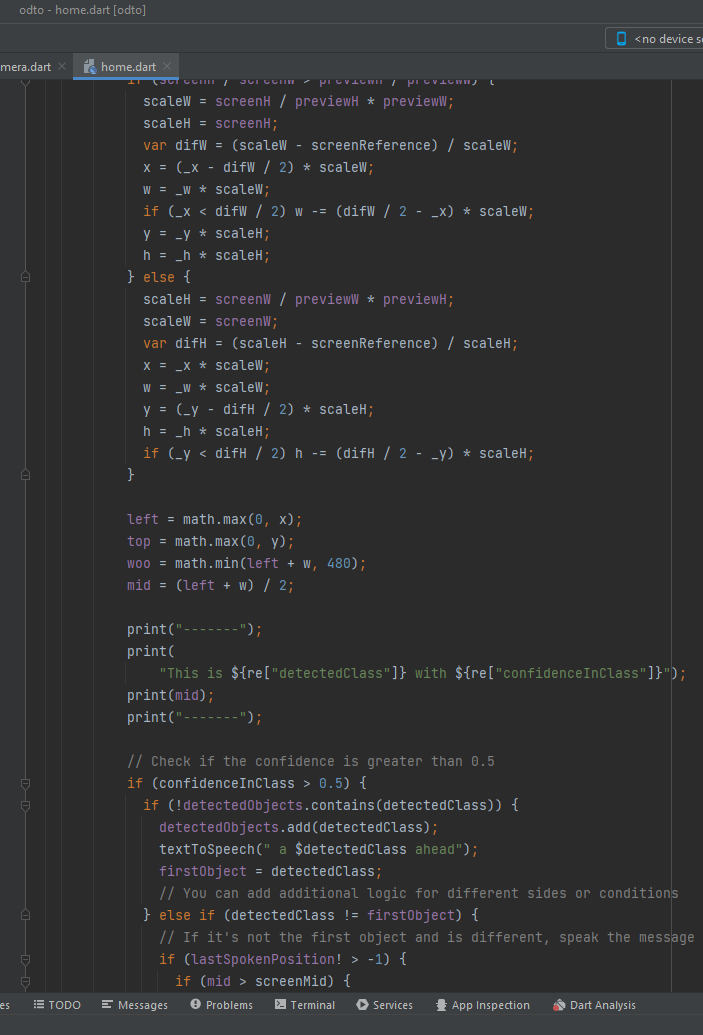
1. تم تثبيت النموذج YOLOV2:  
    تفيد هذه الخطوة بأننا سابقاً كنا نقوم بعملية تبديل ما بين واجهة تحديد النموذج والكاميرا، الآن سيقوم بفتح الكاميرا مباشرة، وكذلك تفيد الآلية هذه بعدم حد إمكانياتنا كباحثين ببناء تطبيق مغاير تماماً لأجل تثبيت نتائج كل بحث أو تطوير قد يطرأ على نماذج التطبيق.

الشكل

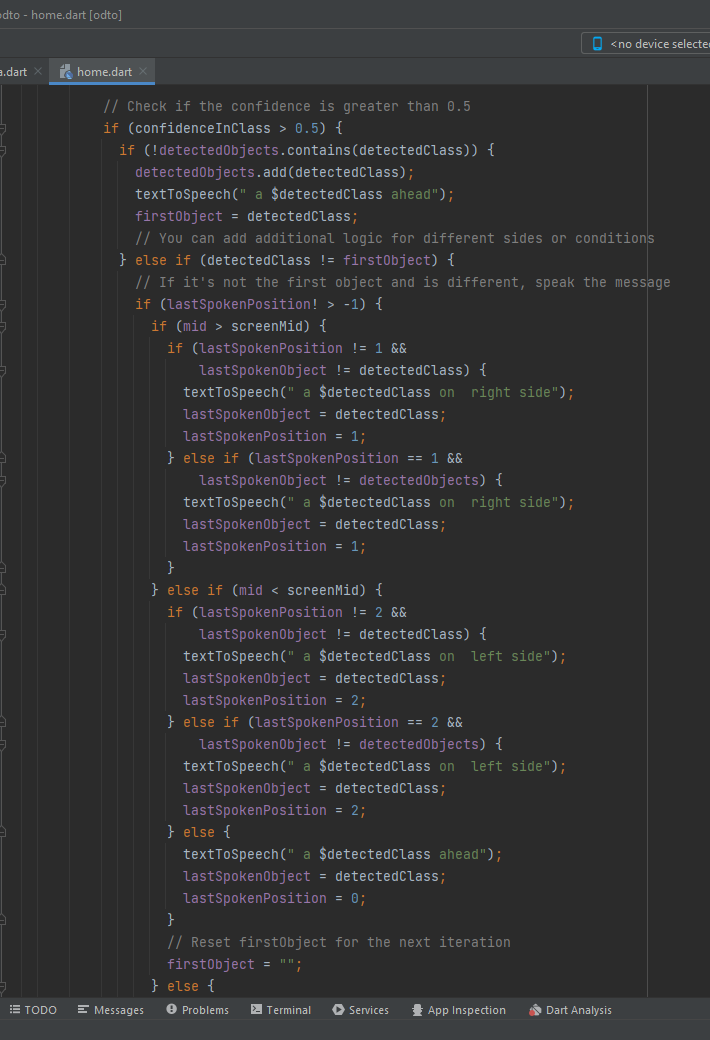
ويجدر أن نذكر أن النموذج الذي تم استخدامه هو نموذج مدرب مسبقاً وذلك اختصاراً لوقت عملية جمع البيانات والتدريب ،بالإضافة إلى أن النموذج المدرب يحوي بعض الشمولية في البيانات المدرب عليها .

1. تم إنشاء واجهة ترحيبية بسيطة بمؤقت زمني:  
    قد تم وضع الشاشة الترحيبية لإنشاء تأخير زمني بسيط نضمن عن طريقه، أنه تم تهيئة جميع مجريات التطبيق، لضمان عدم حدوث خطأ.وقد تم بناء الواجهة وفق الآتي:  
   الشكل
2. بعد انتهاء المؤقت يتم الانتقال مباشرة للكاميرا وبدأ عملية اكتشاف الأغراض:  
    كما يظهر المؤقت في الشكل السابق ، الواجهة تظهر لمدة ثانيتين وبعدها يتم استدعاء الواجهة Home وتمرير وصف الكاميرا لها
3. تم إلغاء كتابة اسم الغرض ونسبة التعرف عليه:  
    وذلك لعدم وجود أهمية لهم لدى المستخدم العادي .
4. تم بناء آلية بسيطة لمعالجة قراءة الأغراض وتحديد مواقعها:  
    يتم في هذه العملية فحص عدة شروط على مصفوفة البيانات العائدة من النموذج نستعرض الشروط وفق الآتي:

الشكل

****

**الشكل**



الشكل



الشكل

# **الفصل الخامس**

**مناقشة وخاتمة**

## 

## -1 الخلاصة:

جرى عرض مجمل مراحل تطوير المشروع في هذه الوثيقة بدءاً من طرح المشكلة الموجودة والحل وصولاً لتنفيذ ذلك الحل من خلال العثور على النموذج الأكثر توافقاً مع عملية التحويل الصوتي.

وهنا لابد من الإشارة إلى أهمية وعظمة الخبرة التي اكتسبناها من المشروع على الصعيد الهندسي عبر المرور بمراحل تصميم المنتج البرمجي كافة اعتماداً على الأسس العلمية لبناء التطبيقات البرمجية وعلى الصعيد التقني من خلال تعلم تقنيات جديدة ساعدت على تقديم مشروع يواكب التطورات التقنية والبرمجية .فكان المشروع خير تطبيق لما تعلمناه في مواد هندسة البرمجيات 1 وهندسة البرمجيات2 وبرمجة التطبيقات والشبكات العصبونية

مروراً بجميع مراحله آخذين بعين الاعتبار قابلية المشروع للتوسع والصيانة .

## 5-2 الآفاق المستقبلية:

القدرة على تحقيق قاعدة معرفة واكتشاف أغراض بشكل أكبر وأكثر دقة، تطوير التطبيق ليصبع مساعد تفاعلي بشكل أكبر للمكفوفين ،تحقيق عمليات تتبع للأغراض أكثر فاعلية ودقة .

# المراجع

1- <https://scholar.ppu.edu/handle/123456789/6816> جامعة بوليتكنك فلسطين.

2-https://www.aljazeera.net/tech/2022/8/2/ أخبار الجزيرة -تكنولوجيا

3-https://www.oracle.com/ae-ar/artificial-intelligence/what-is-ai/?source=rh-rail ما هو الذكاء الاصطناعي

4-https://en.wikipedia.org/wiki/ image\_interpretation تفسير الصور.

2-https://ai.meta.com/tools/detectron/

3-https://www.analyticsvidhya.com/blog/2022/06/yolo-algorithm-for-custom-object-detection/

5-© 2021 JETIR April 2021, Volume 8, Issue 4 www.jetir.org (ISSN-2349-5162)

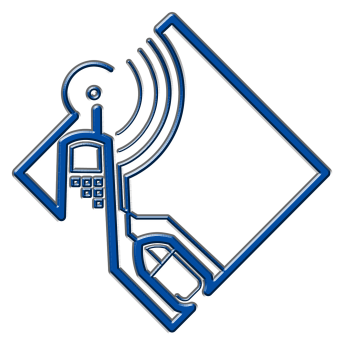
# الملاحق

قد تم إلحاق ملف نصي مع المشروع يحوي عناوين الأغراض التي يمكن للنموذج المستخدم اكتشافها.

**ABSTRACT**

In their daily lives, visually impaired and blind people face a variety of obstacles. The proposed project's aim is to create an Android-based application for people who are blind or visually impaired. The system's goal is to detect real-time objects and notify blind people about them through audio or vocal information. The proposed framework makes use of Machine Learning and Google's TensorFlow Lite object detection API. It's a useful app that makes it much easier for blind and visually disabled people to complete daily tasks with less effort.

Keywords - Object Detection, Machine Learning ,Android Application ,Image, Visually Impaired People, Computer Vision, Algorithms ,Neural Networks.



**UNIVERSITY OF KALAMOON**

**FACULTY OF ENGINEERING**

DEPARTMENT OF INFORMATION TECHNOLOGY ENGINEERING

Real Time Object Detection Android Application using Machine Learning Concepts for Visually Impaired People

A (project/( Final project) submitted in partial fulfillment of the requirement of the degree B.Sc. in the Department of Information Technology Engineering

By

AbdAlhadi Haidar

Fatimah Melhem

Rawan Alahmad

Supervised by

Dr. Shaza Shehada

Preparing year 2023