

Chapitre 9 : Android ML Kit

Introduction

ML Kit est un SDK de Google conçu pour intégrer facilement des fonctionnalités de Machine Learning dans les applications Android et iOS. Il propose des API prêtes à l'emploi pour des tâches courantes comme la reconnaissance de texte, la détection de visages, le balayage de codes-barres, l'étiquetage d'images et la détection d'objets.

Annoncé lors de Google I/O 2018, Google ML Kit simplifie l'intégration de technologies de ML les applications mobiles, rendant ces fonctionnalités accessibles même aux développeurs n'ayant pas une expertise approfondie en apprentissage automatique.

ML Kit fonctionne directement sur l'appareil et offre des performances rapides et peut être utilisé hors ligne. Cette exécution locale permet aussi de préserver la confidentialité des données, qui restent sur le mobile sans nécessiter de connexion cloud.

Le SDK ML Kit est optimisé pour les appareils mobiles, garantissant ainsi une expérience utilisateur de qualité. Avant ML Kit, l'intégration de modèles de Machine Learning nécessitait souvent un serveur dédié et des échanges constants avec le cloud, rendant le processus dépendant d'une connexion internet. Avec ML Kit, tout se passe sur l'appareil, ce qui permet d'utiliser ces fonctionnalités même en mode hors ligne. De plus, pour les appareils compatibles, l'accélération matérielle augmente les performances et réduit le temps de réponse.

I. API Google ML Kit

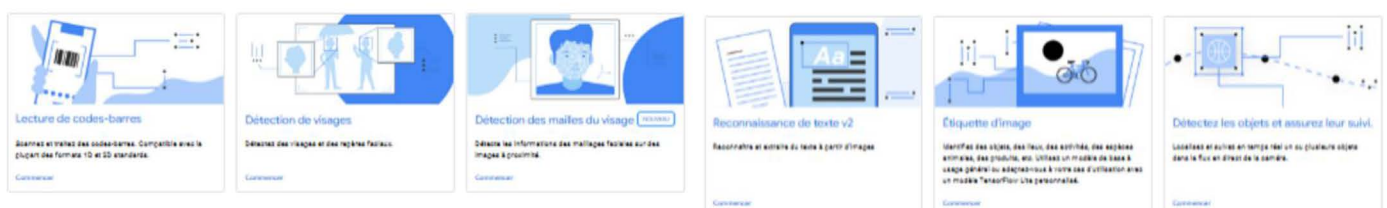
Google offre aux développeurs mobiles 15 API clés en main dans leurs applications grâce à Google ML Kit. Google continue d'améliorer et d'optimiser ces modèles pré-entraînés, tout en prévoyant d'en ajouter de nouveaux pour offrir des fonctionnalités toujours plus performantes au public. Un didacticiel interactif est également disponible pour guider les développeurs, quel que soit leur niveau de compétence en apprentissage automatique, dans la création de leur première application ML sur appareil.

<https://developers.google.com/ml-kit/guides?hl=fr>

Les API disponibles dans Google ML Kit incluent :

A. API Vision :

API d'analyse de vidéos et d'images permettant d'étiqueter des images et de détecter des codes-barres, du texte, des visages et des objets





1. Lecture de codes-barres

L'API de lecture de codes-barres de ML Kit permet de décoder localement la plupart des formats standards sans connexion réseau. Elle peut extraire des informations structurées, comme des coordonnées ou des identifiants Wi-Fi, pour permettre à l'application de réagir intelligemment. Pour une solution simple et rapide, le lecteur de code Google utilise le même modèle, se concentrant sur le code-barres centralisé sans nécessiter de permissions liées à l'appareil photo.



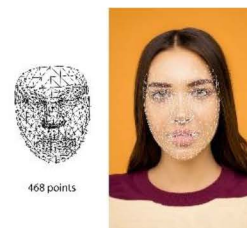
2. Détection de visages



L'API de détection de visages de ML Kit permet de repérer les visages dans une image, de localiser leurs principales caractéristiques et de tracer leurs contours sans identifier les personnes. Elle est utile pour des applications de retouche de selfies, de création d'avatars et peut fonctionner en temps réel, par exemple dans les chats vidéo ou les jeux réactifs aux expressions faciales.

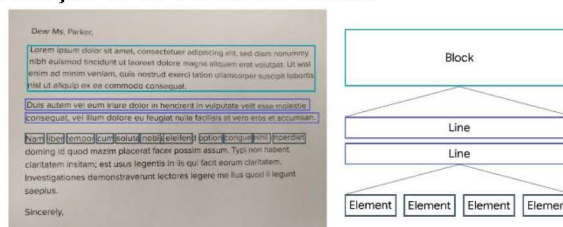
3. Détection des mailles du visage

L'API de détection de maillages faciaux de ML Kit permet de générer en temps réel un maillage 3D de 468 points pour les selfies, lorsque les visages se trouvent à environ 2 mètres de la caméra. Ce maillage détaillé, avec un cadre de délimitation et des points 3D, est idéal pour des applications de filtres de réalité augmentée, de selfies, et de chat vidéo, offrant une grande précision pour des effets visuels avancés.



4. Reconnaissance de texte

L'API de reconnaissance de texte de ML Kit v2 identifie le texte dans les jeux de caractères chinois, devanagari, japonais, coréen et latin. Elle peut automatiser des tâches de saisie de données, comme la lecture de cartes de crédit, de reçus et de cartes de visite.



5. Étiquetage d'images



Les API d'étiquetage d'images de ML Kit détectent et extraient des informations sur des entités d'une image, couvrant diverses catégories comme les objets, lieux, activités, animaux et produits. Un modèle personnalisé peut également être utilisé pour des cas spécifiques de classification d'images.

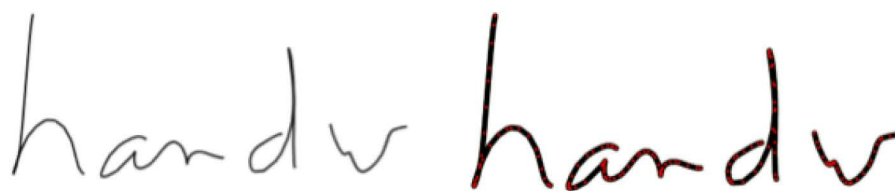
6. Détection et suivi d'objets

L'API de détection et de suivi d'objets de ML Kit permet d'identifier et de suivre des objets dans une image ou un flux de caméra en direct, avec la possibilité de les classer via un modèle intégré ou personnalisé. Opérant directement sur l'appareil, elle peut servir de frontend pour un pipeline de recherche visuelle, envoyant ensuite les objets détectés à un backend cloud pour une analyse plus approfondie.



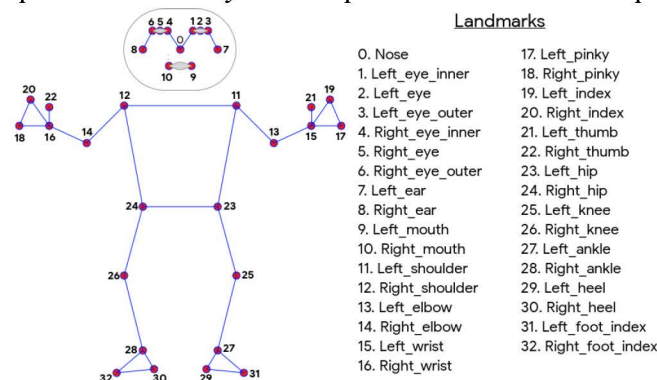
7. Reconnaissance d'encre numérique

L'API de reconnaissance d'encre numérique de ML Kit permet de reconnaître du texte manuscrit, de classer des gestes, et de reconnaître des croquis en plusieurs langues. Elle utilise la technologie de reconnaissance d'écriture manuscrite de Google (Gboard, Google Traduction). Adaptée aux caractères non disponibles sur les claviers, elle permet aussi de naviguer avec des gestes et de reconnaître des emojis dessinés, fonctionnant hors ligne sur Android et iOS.



8. Détection des postures

L'API ML Kit Pose Detection est une solution légère et flexible pour détecter les postures corporelles en temps réel à partir d'une vidéo ou d'une image. Elle identifie 33 points de repère squelettiques couvrant le corps, y compris le visage, les mains et les pieds. Cette correspondance squelettique permet de distinguer différentes postures en analysant les positions relatives des points de repère.



9. Segmentation des selfies

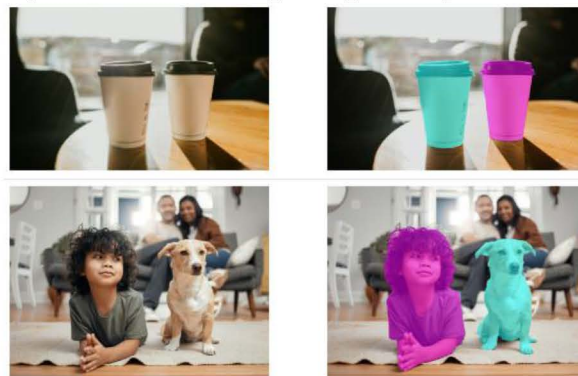


L'API de segmentation de selfies de ML Kit permet aux développeurs de séparer facilement l'arrière-plan des utilisateurs dans une scène, mettant ainsi l'accent sur l'essentiel. Elle génère un masque à partir d'une image d'entrée, où chaque pixel est assigné une valeur flottante entre $[0.0, 1.0]$, indiquant la confiance que le pixel représente une personne. Cette API fonctionne avec des images statiques et des vidéos en direct, utilisant les résultats des images précédentes pour obtenir des segments plus fluides lors des vidéos.

10. Segmentation des sujets

L'API de segmentation du sujet de ML Kit permet aux développeurs de distinguer facilement plusieurs sujets de l'arrière-plan d'une image, facilitant des applications comme la création d'autocollants, le changement d'arrière-plan ou l'ajout d'effets visuels. Elle génère un masque et un bitmap de sortie pour

le premier plan, ainsi que des masques pour chaque sujet détecté, où les pixels sont attribués une valeur flottante entre 0.0 et 1.0, indiquant la confiance que le pixel représente un sujet.



11. Scan de documents

L'API d'analyse de documents de ML Kit permet de numériser facilement des documents physiques avec un flux d'interface utilisateur de haute qualité sur Android. Elle donne aux utilisateurs un contrôle complet sur la numérisation, avec des options pour recadrer, appliquer des filtres, et corriger les ombres ou taches. Les modèles ML et ressources nécessaires sont centralisés via les services Google Play, réduisant ainsi l'impact binaire et éliminant le besoin d'autorisations directes pour l'appareil photo, assurant une numérisation entièrement sur l'appareil.



B. API de langage naturel :

Les API de traitement du langage naturel permettent d'identifier et de traduire 58 langues, et de fournir des suggestions de réponse.



1. Identification de la langue

L'API d'identification de la langue de ML Kit permet de détecter la langue d'une chaîne de texte, utile pour du texte saisi par l'utilisateur sans information linguistique associée. Elle prend en charge plus d'une centaine de langues, y compris les textes en caractères natifs et romanisés pour des langues comme l'arabe, le russe et le chinois.

Simple language identification	
"My hovercraft is full of eels."	en (English)
"Dao shan xue hai"	zh-Latn (Latinized Chinese)
"ph'nglui mglw'nafh wgah'nagl fhtagn"	und (undetermined)

Confidence distribution	
"an amicable coup d'etat"	en (0.52) fr (0.44) ca (0.03)

2. Traduction de texte

L'API de traduction de ML Kit permet de traduire du texte en temps réel dans plus de 50 langues directement sur l'appareil, sans connexion internet. Elle utilise les mêmes modèles que le mode hors ligne de Google Traduction, avec une gestion dynamique des packs de langue pour optimiser le stockage.



3. Réponse suggérée

L'API Réponse suggérée de ML Kit propose des réponses automatiques pertinentes aux messages, aidant les utilisateurs à répondre rapidement, notamment sur des appareils à saisie limitée. Elle prend en compte le contexte complet d'une conversation pour générer des suggestions plus précises, en traitant les messages localement sans nécessiter de connexion internet.

Code temporel	User-ID	Utilisateur local ?	Message
Jeu. 21 févr., 13:13:39 PST 2019		true	tu es en route ?
Jeu. 21 févr., 13:15:03 PST 2019	FRIEND0	false	Je suis en retard, désolé.

Réponses suggérées

Suggestion n° 1	Suggestion n° 2	Suggestion n° 3
Pas de problème.	😊	Aucun problème.

4. Extraction d'entités

L'API d'extraction d'entités de ML Kit enrichit l'expérience utilisateur en reconnaissant des entités spécifiques dans le texte et en permettant d'ajouter des actions contextuelles. Elle va au-delà des simples opérations de couper/copier/coller, en facilitant une interaction intelligente avec le texte dans les applications.

Texte d'entrée	Entités détectées
Rendez-vous au 1600 Amphitheatre Parkway, Mountain View, CA 94043, États-Unis Organisons un rendez-vous pour en discuter.	Type d'entité 1: Adresse Texte de l'entité 1: "1600 Ampitheatre Parkway, Mountain View, CA 94043, États-Unis"
Vous pouvez contacter l'équipe de test demain à l'adresse info@google.com pour déterminer le meilleur calendrier.	Type d'entité 1: Date-heure Texte de l'entité 1 : "24 juin 2020"
	Type d'entité 2: Adresse e-mail Texte de l'entité 2: info@google.com
Votre commande auprès de Google a été expédiée. Pour suivre l'avancement de votre livraison, veuillez utiliser ce numéro de suivi: 9612804152073070474837.	Type d'entité: Numéro de suivi Texte de l'entité: "9612804152073070474837"
Appelez le restaurant au 555-555-1234 pour payer votre dîner. Mon numéro de carte est le 4111-1111-1111-1111.	Type d'entité 1: numéro de téléphone Texte de l'entité 1: "555-555-1234"
	Type d'entité 2: Carte de paiement Texte de l'entité 2: "4111 1111 1111 1111"

II. Exemple de reconnaissance des codes-barres

```
//PrixProduit
public class MainActivity extends AppCompatActivity {
    private Button btnSuivant;
    private Button btnAnalyser;
    private ImageView imgProduit;
    private TextView tvPrix;
    private int[] tImage = new int[] {
        R.drawable.eau_m1,R.drawable.eau_m2,R.drawable.eau_safia };
    private int indiceCourant;

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        ...
        initGraphique();
    }

    private void initGraphique() {
        ...
        indiceCourant=0;
        ajouterEcouteurs();
    }

    private void ajouterEcouteurs() {
        btnAnalyser.setOnClickListener(view -> analyser());

        btnSuivant.setOnClickListener(view -> suivant());
    }

    private void suivant() {
        indiceCourant = (indiceCourant + 1) % tImage.length;
    }
}
```

```
imgProduit.setImageResource(tImage[indiceCourant]);
analyser();
}

private void analyser() {
    Bitmap bitmap = ((BitmapDrawable) imgProduit.getDrawable()).getBitmap();
    InputImage image = InputImage.fromBitmap(bitmap, 0);
    BarcodeScanner scanner = BarcodeScanning.getClient();
    Task<List<Barcode>> result = scanner.process(image)
        .addOnSuccessListener(new OnSuccessListener<List<Barcode>>() {
            @Override
            public void onSuccess(List<Barcode> barcodes) {
                afficherPrix(barcodes);
            }
        })
        .addOnFailureListener(new OnFailureListener() {
            @Override
            public void onFailure(@NonNull Exception e) {
                Toast.makeText(MainActivity.this, "Failure:" + e.getMessage(),
                    Toast.LENGTH_SHORT).show();
            }
        });
}

private void afficherPrix(List<Barcode> barcodes) {
    if(barcodes.size() > 0) {
        if(barcodes.size() == 1) {
            Barcode bar = barcodes.get(0);
            tvPrix.setText(getPrix(bar.getRawValue()));
        } else {
            Toast.makeText(this, "Plusieurs code à barre dans l'image !",
                Toast.LENGTH_SHORT).show();
        } else {
            Toast.makeText(this, "Aucun code à barre dans l'image !", Toast.LENGTH_SHORT).show();
        }
    }

    private String getPrix(String rawValue) {
        String prix = "Le prix ";
        if(rawValue.equals("6191577600015"))
            prix += "de Eau Safia est 0.900 DT";
        else if(rawValue.equals("123456789012"))
            prix += "de Eau M1 est 6.700 DT";
        else if(rawValue.equals("3401312345624"))
            prix += "de Eau M2 est 4.300 DT";
        else
            prix += "Inconnu";
        return prix;
    }
}
```