

Mini-Projets NoSQL

Reconnaissance Automatique de la Langue des Signes Marocaine

Approche par Vision Informatique
et Stockage MongoDB



MAHDAD ABDELHAKIM
ACHRAF BIRHRISSEN

Encadré par Mr
IDRISS MOUMEN

- 1. Introduction
- 2. Problématique
- 3. Objectifs
- 4. Méthodologie
- 5. Résultats et discussion
- 6. Conclusion



Introduction

La reconnaissance automatique de la langue des signes est un domaine de recherche en pleine expansion, avec un potentiel considérable pour améliorer la communication et l'accessibilité pour les personnes sourdes et malentendantes.

Ce projet vise à développer un système capable de traduire les signes MSL en texte, comblant ainsi le fossé de communication entre les signeurs et les non-signeurs. Notre travail s'appuie sur l'ensemble de données ArASL, une ressource précieuse pour la recherche dans ce domaine, et propose une approche basée sur l'apprentissage profond pour atteindre une haute précision.



Problématiques

- Le fossé de communication
- Variabilité de la langue des signes
- Défis techniques de la reconnaissance visuelle.
- Disponibilité limitée des ressources
- Adaptation aux conditions réelles.



Objectifs



- 01** Développement d'un modèle de reconnaissance MSL performant.
- 02** Atteindre une haute précision de classification.
- 03** Création d'un système de détection en temps réel.
- 04** Intégration de MongoDB pour le stockage et l'analyse des données.
- 05** Analyse des erreurs de classification.

Méthodologie

Description du Dataset

- **Source de données** : ArASL Database 54K Final
- **Description du dataset** : Ce projet utilise l'ensemble de données ArASL Database 54K Final, une collection exhaustive de signes de la langue des signes arabe. L'ensemble comprend [insérer le nombre exact] images couvrant un total de 32 classes, chacune représentant une lettre de l'alphabet arabe. Chaque classe contient des exemples de différents locuteurs et dans diverses conditions, assurant ainsi une certaine robustesse et généralisation du modèle.

Méthodologie

Prétraitement des Données

- Redimensionnement des images à 64x64 pixels : Toutes les images ont été redimensionnées à une résolution uniforme de 64x64 pixels pour garantir une entrée cohérente pour le modèle CNN.
- Normalisation : Les valeurs de pixel ont été normalisées entre 0 et 1 afin d'améliorer la stabilité de l'entraînement et d'accélérer la convergence du modèle.
- Partition en ensembles d'entraînement, de validation et de test : L'ensemble de données a été divisé en trois parties, avec un ratio de 70% pour l'entraînement (pour l'apprentissage du modèle), 15% pour la validation (pour le réglage des hyperparamètres) et 15% pour le test (pour l'évaluation finale). Cette division nous permet de mesurer les performances du modèle sur des données jamais vues auparavant, évitant ainsi le surapprentissage.

Méthodologie

Modèle CNN



Méthodologie

Processus d'entraînement

- **Optimiseur (Optimizer):** Adam.
- **Fonction de perte (Loss Function):** Categorical cross-entropy.
- **Nombre d'époques (Epochs):** 15.
- **Taille du lot (Batch Size):** 32.
- **Validation et Ajustement:** L'ensemble de validation à jouer le rôle de sur-apprentissage
- **Early Stopping:** L'apprentissage s'arrête quand les performances ne s'améliorent pas pour un certain nombre d'époques.

Méthodologie

Métriques d'évaluation

- **Précision (Accuracy)** : Mesure la proportion globale de prédictions correctes.
- **Précision (Precision)** : Indique la proportion de prédictions positives qui étaient réellement correctes.
- **Rappel (Recall)** : Indique la proportion d'instances positives réelles qui ont été correctement identifiées.
- **Score F1 (F1-Score)** : La moyenne harmonique de la précision et du rappel, fournissant une mesure équilibrée des performances du modèle.
- **Matrice de confusion (Confusion Matrix)** : Visualisation matricielle des performances de classification, montrant les vrais positifs, les faux positifs, les vrais négatifs et les faux négatifs.
-

Méthodologie

Intégration de MongoDB

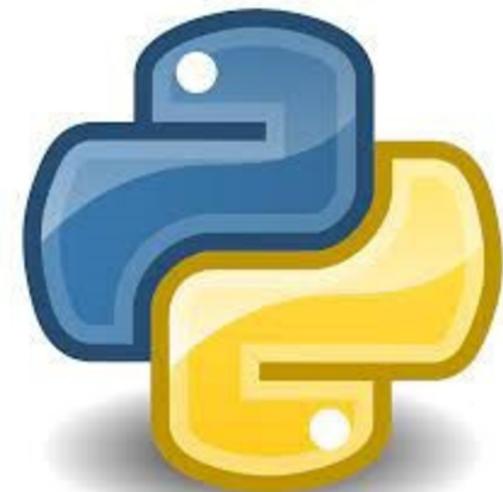
- **Stockage des données** : Utilisation de **MongoDB Atlas**, une base de données NoSQL basée sur le cloud, pour stocker les données suivantes :
 - Métadonnées de l'ensemble de données (nombre de classes, taille des images, etc.).
 - Quelques exemples d'images.
 - Architecture du modèle (au format JSON).
 - Historique d'entraînement (précision et perte à chaque époque).
 - Matrice de confusion.
 - Exemples mal classés (images et prédictions incorrectes).

Méthodologie

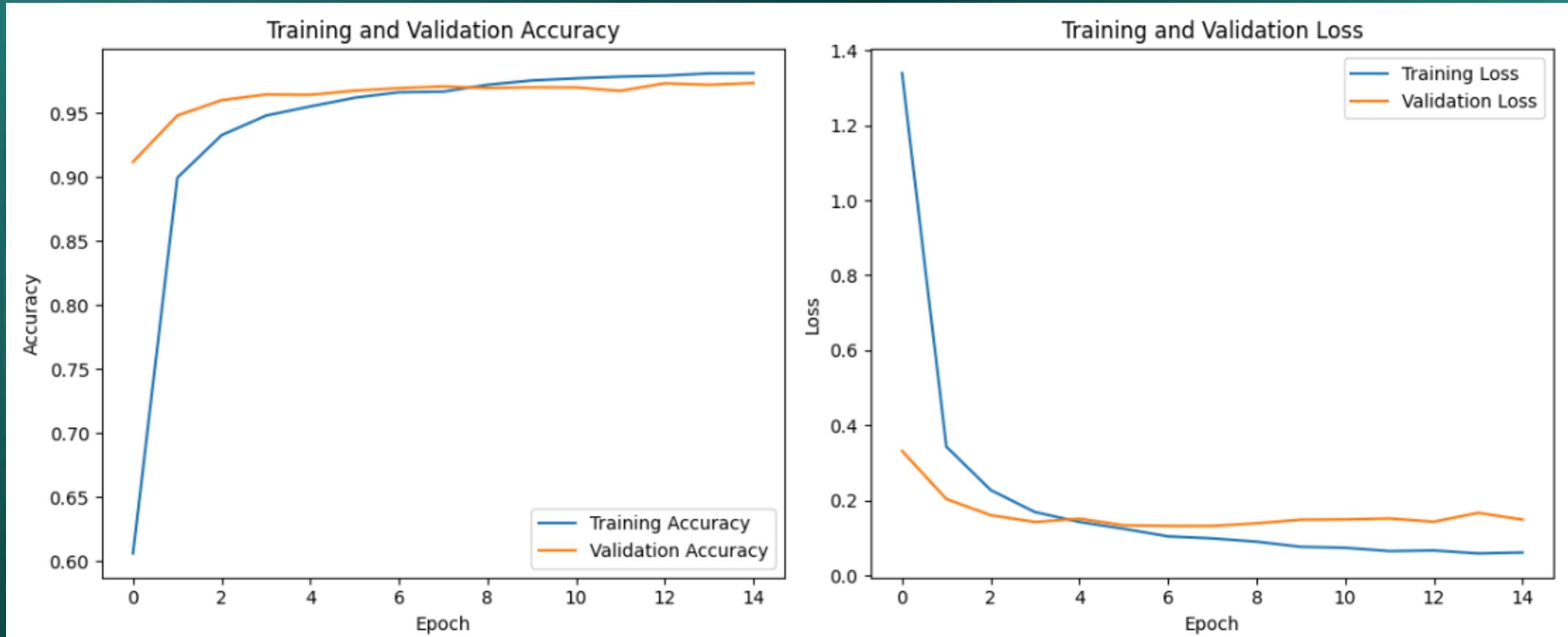
Détection en temps réel

- **Acquisition vidéo** : Utilisation d'OpenCV pour capturer le flux vidéo de la webcam.
- **Segmentation de la main** : Algorithmes de traitement d'image pour détecter et segmenter la région de la main dans chaque trame.
- **Prétraitement des images** : Redimensionnement et normalisation des images de la main segmentées.
- **Prédiction du modèle** : Application du modèle CNN entraîné à l'image de la main prétraitée.
- **Affichage des résultats** : Affichage de la prédiction et du score de confiance sur le flux vidéo.

Outils et Environnement de Réalisation

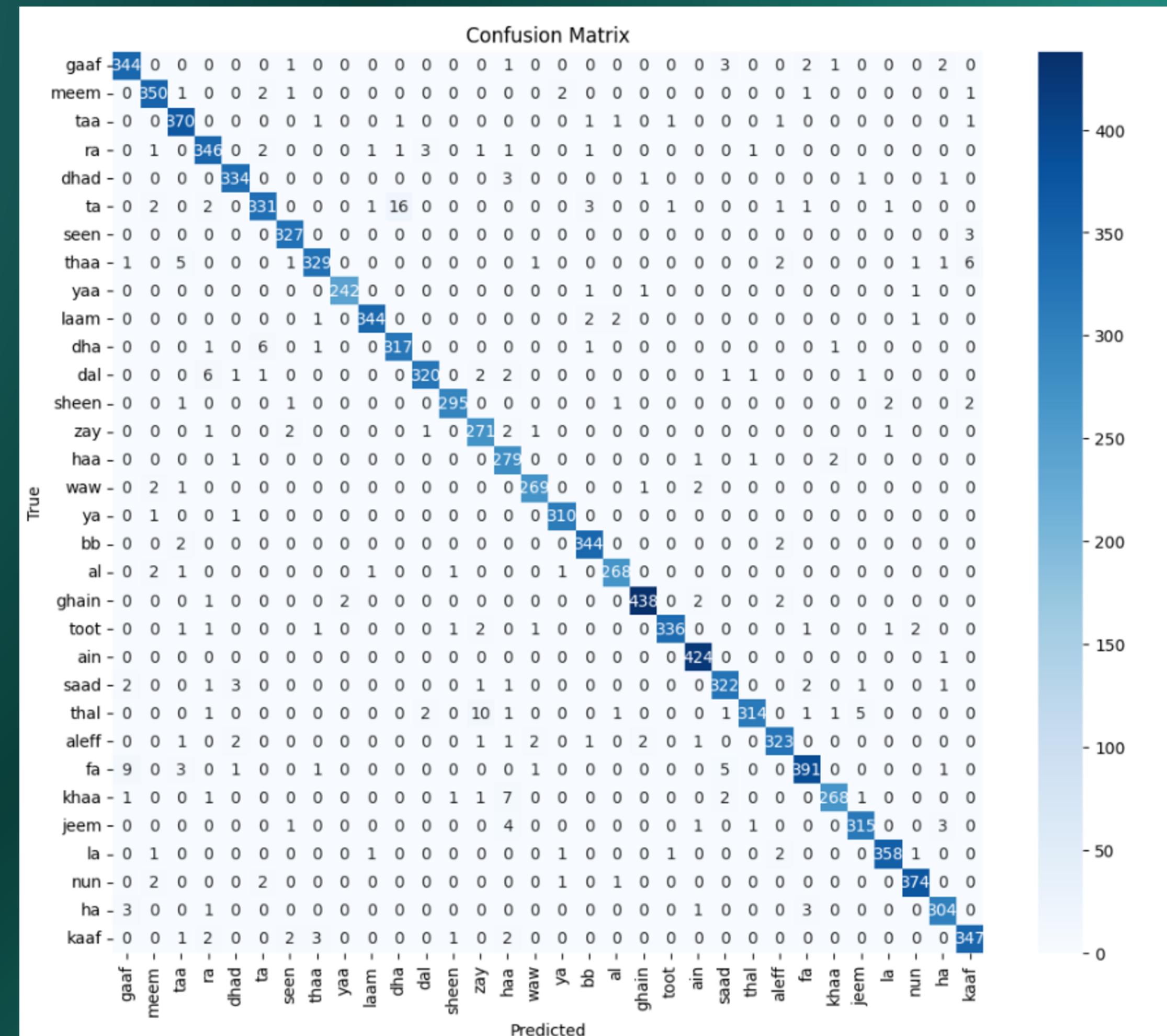


Résultats et discussion



Test accuracy: 0.9717

Résultats et discussion



Résultats et discussion

	Precision	Recall	F1-Score	Support
gaaf	0.96	0.97	0.96	354
meem	0.97	0.98	0.97	358
taa	0.96	0.98	0.97	377
ra	0.95	0.97	0.96	358
dhad	0.97	0.98	0.97	340
ta	0.96	0.92	0.94	349
seen	0.98	0.96	0.97	348
thaa	0.98	0.95	0.96	347
yaa	0.99	0.98	0.99	345
laam	0.99	0.98	0.99	350
dha	0.98	0.99	0.98	327
dal	0.98	0.96	0.97	335
sheen	0.97	0.98	0.97	302
zay	0.98	0.96	0.97	279
haa	0.98	0.98	0.98	284
waw	0.98	0.98	0.98	275
ya	0.98	0.98	0.98	312
bb	0.97	0.99	0.98	348
al	0.98	0.98	0.98	274
ghain	0.98	0.98	0.98	445
toot	0.98	0.97	0.98	447
ain	0.96	1.00	0.98	425
saad	0.98	0.96	0.97	434
thal	0.97	0.97	0.97	334
aleff	0.97	0.98	0.97	334
fa	0.96	0.98	0.97	282
khaa	0.97	0.97	0.97	282
jeem	0.98	0.97	0.97	412
la	0.96	0.98	0.97	365
nun	0.98	0.98	0.98	380
ha	0.97	0.98	0.97	338
kaaf	0.96	0.97	0.96	358

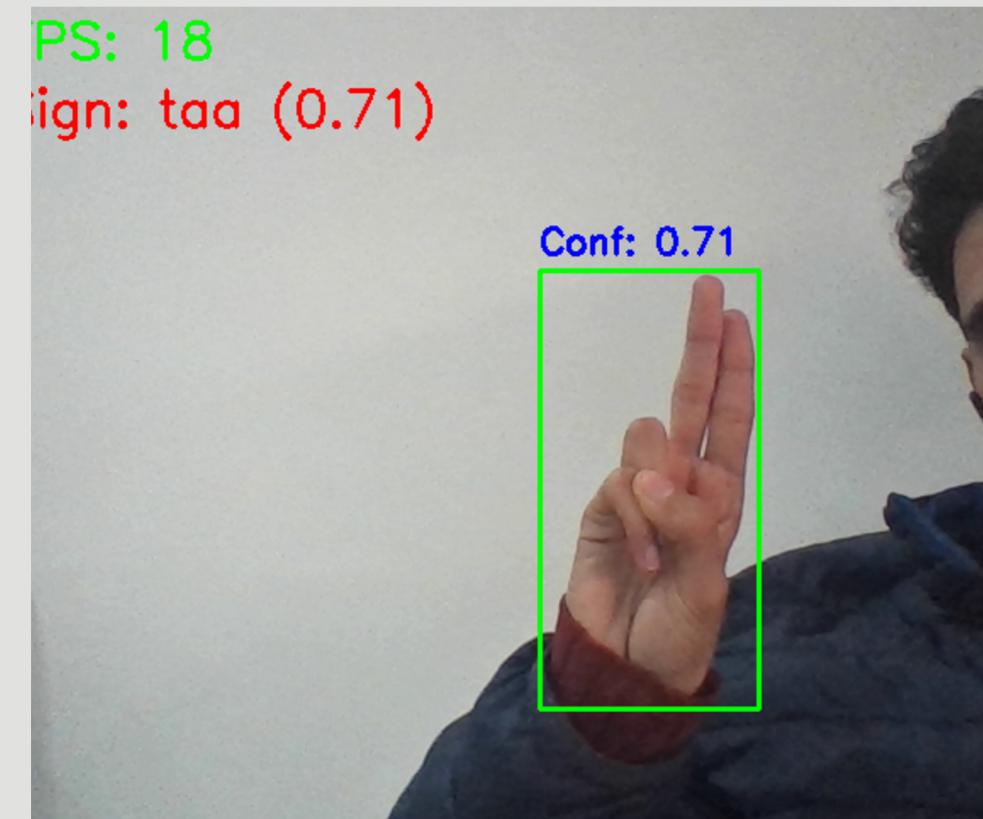
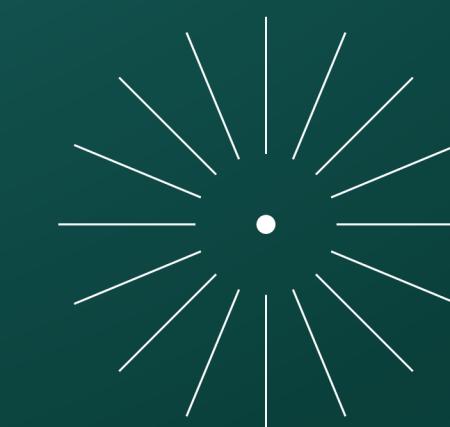
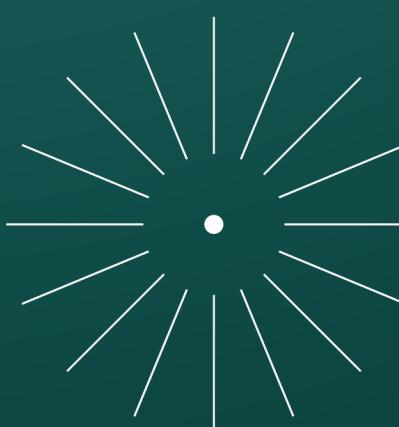
The screenshot shows the Apache Atlas web interface. The top navigation bar includes 'Atlas', 'pfe team - 2...', 'Access Manager', 'Billing', 'All Clusters', 'Get Help', and 'Achraf'. The left sidebar has sections for 'Overview', 'DATABASE', 'Clusters' (which is selected), 'SERVICES', 'Atlas Search', 'Stream Processing', 'Triggers', 'Migration', 'Data Federation', 'SECURITY', 'Quickstart', 'Backup', 'Database Access', 'Network Access', 'Advanced', and 'Goto'. The main panel displays the 'sign_language.performance_reports' database. It shows storage size (36KB), logical data size (10.14KB), total documents (2), and index sizes (36KB). There are tabs for 'Find', 'Indexes', 'Schema Anti-Patterns', 'Aggregation', and 'Search Indexes'. A search bar for 'Search Namespaces' is present. An 'INSERT DOCUMENT' button is at the top right. Below it is a 'Filter' section with a query input field containing '{ field: \'value\' }'. The results section shows 'QUERY RESULTS: 1-2 OF 2' with the following document content:

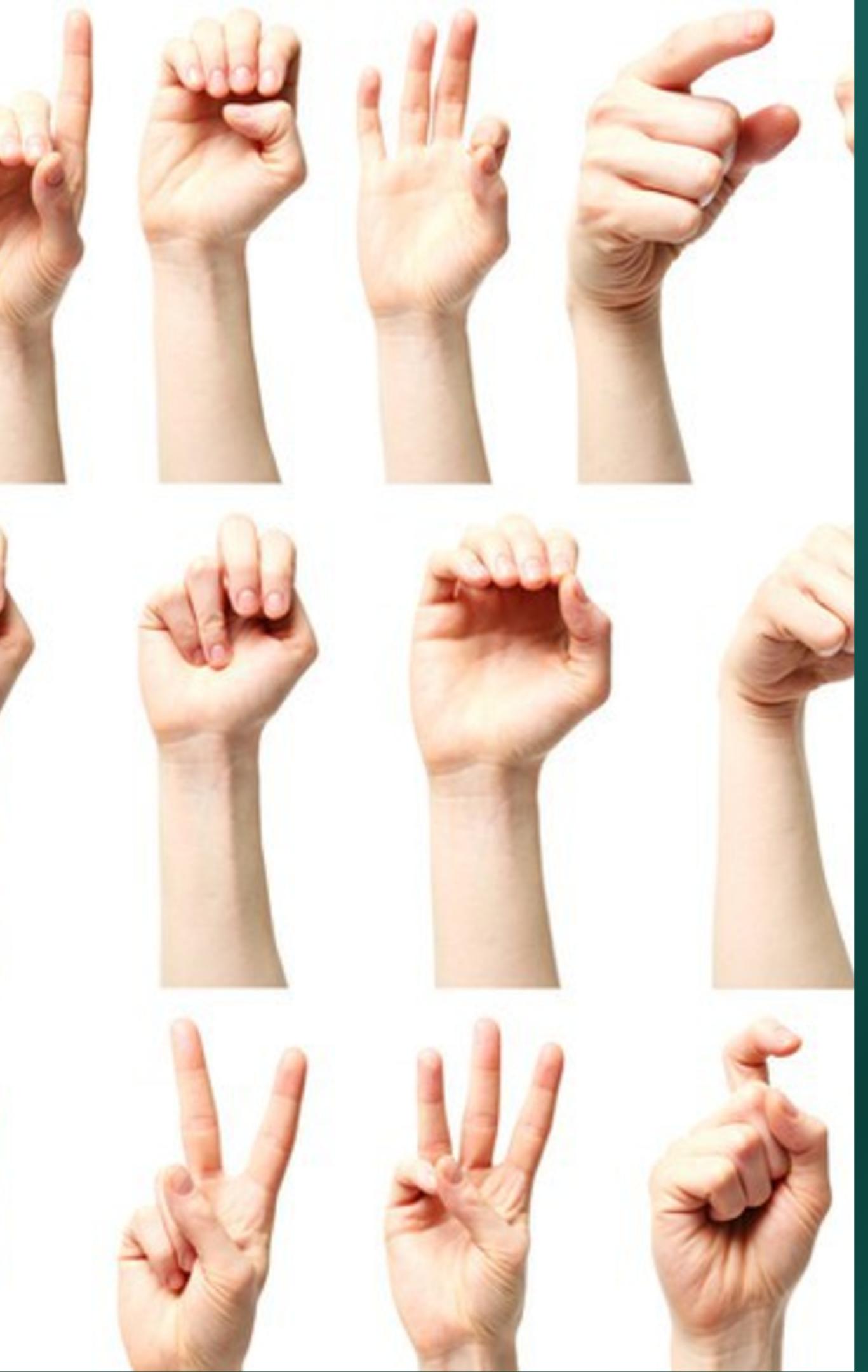
```
_id: ObjectId('67bf2be0ea47ae95e6da0d59')
timestamp : 2025-02-26T14:57:36.658+00:00
model_name : "CNN_MoroccanSignLanguage_v1"
overall_accuracy : 0.97
training_epochs : 15
class_metrics : Array (32)
below_average_performers : Array (10)
confusion_patterns : Array (10)
false_positive_leaders : Array (5)
0: Object
_id : "gaaf"
count : 8
1: Object
_id : "ra"
count : 8
2: Object
_id : "haa"
count : 6
3: Object
_id : "seen"
count : 6
4: Object
_id : "ha"
count : 5
total_classes : 32
```

Résultats et discussion

Résultats et discussion

Détection en temps réel





Conclusion

En résumé, ce projet a atteint son objectif principal qui était de développer un modèle CNN précis et efficace pour la reconnaissance de la langue des signes marocaine.

Bien que des améliorations restent possibles, les résultats obtenus démontrent clairement le potentiel de l'apprentissage profond pour relever les défis complexes de ce domaine.

Nous espérons que ce travail contribuera à améliorer l'accessibilité pour les personnes sourdes et malentendantes au Maroc et dans le monde arabophone, et à promouvoir une communication plus inclusive.