Comment comprendre facilement la langue des signes grâce au machine learning et au computer vision

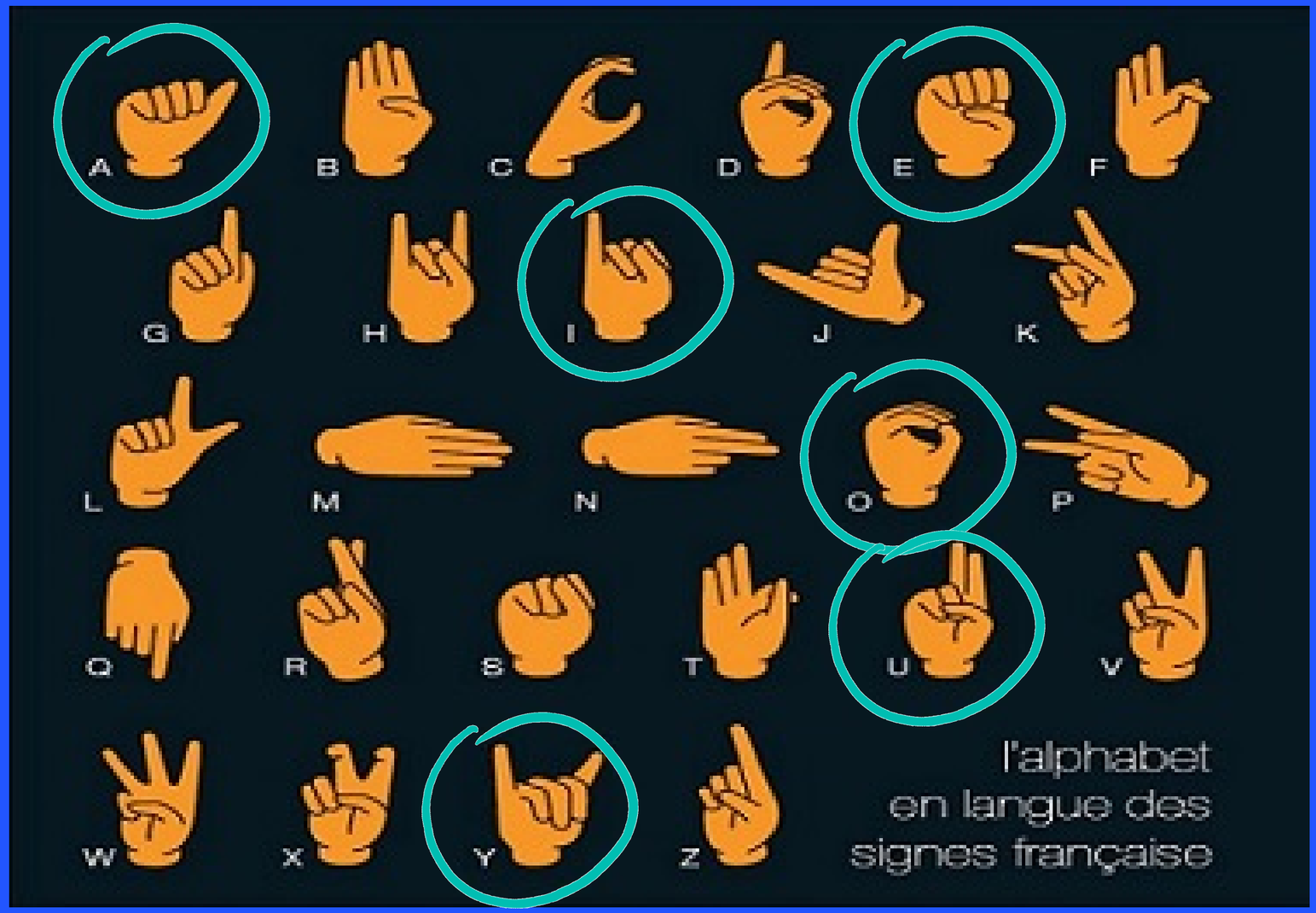
Romane Lollier, An Hoang, Agathe Mathis

- Problématique
- Jeu de données
- Modèle
- Ajustement du modèle
- Ajustement de jeu de donnees
- Resultat et conclusion

Problématique

Reconnaissance de la LSF → Alphabet LSF → Voyelles LSF

Jeu de données

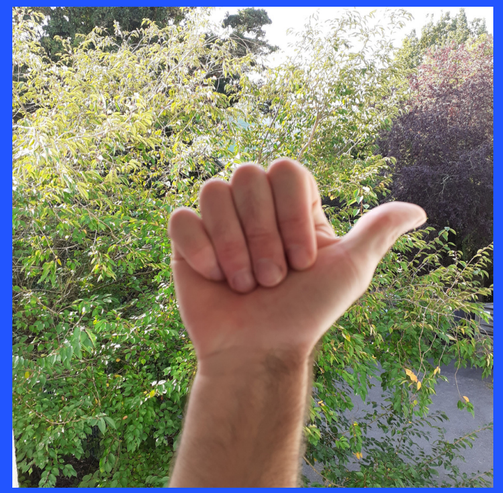
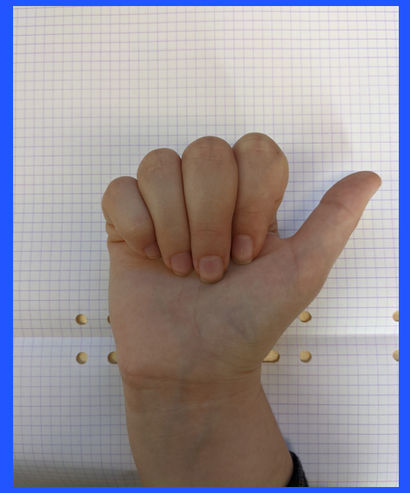


Jeu de données

Photos de mains en variant les fonds, les personnes et la luminosité

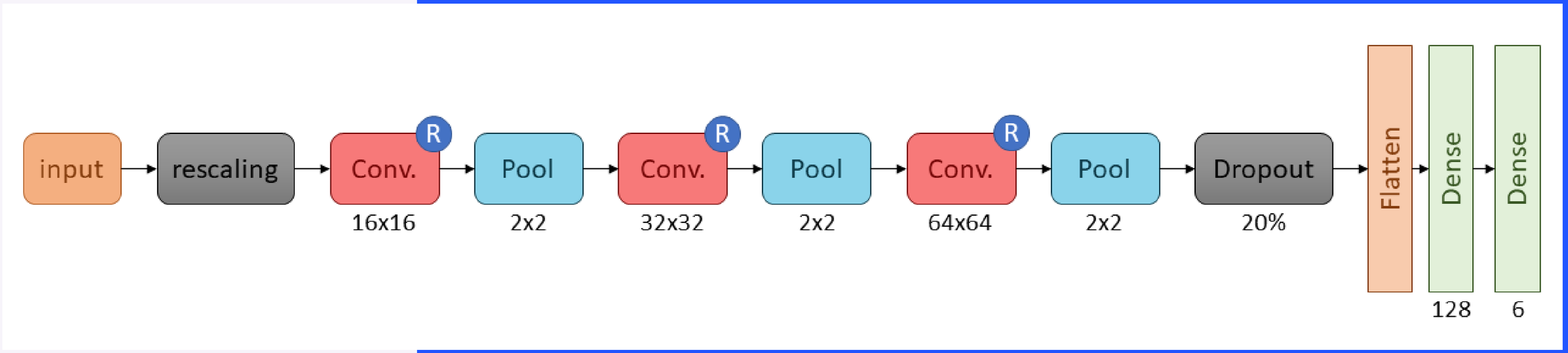
Exemple : Lettre A



Formatage des images :
- fond neutre, intérieur, extérieur
- 180x180
- 6 classes: ~ 80 image/class
- Données : matrice en 3 dimensions (largeur, hauteur et couleurs)

Conception du modèle CNN



Conception du modèle

```
Using 325 files for training.
Found 406 files belonging to 6 classes.
Using 81 files for validation.
```

Essai du modèle

Ajustement du modèle

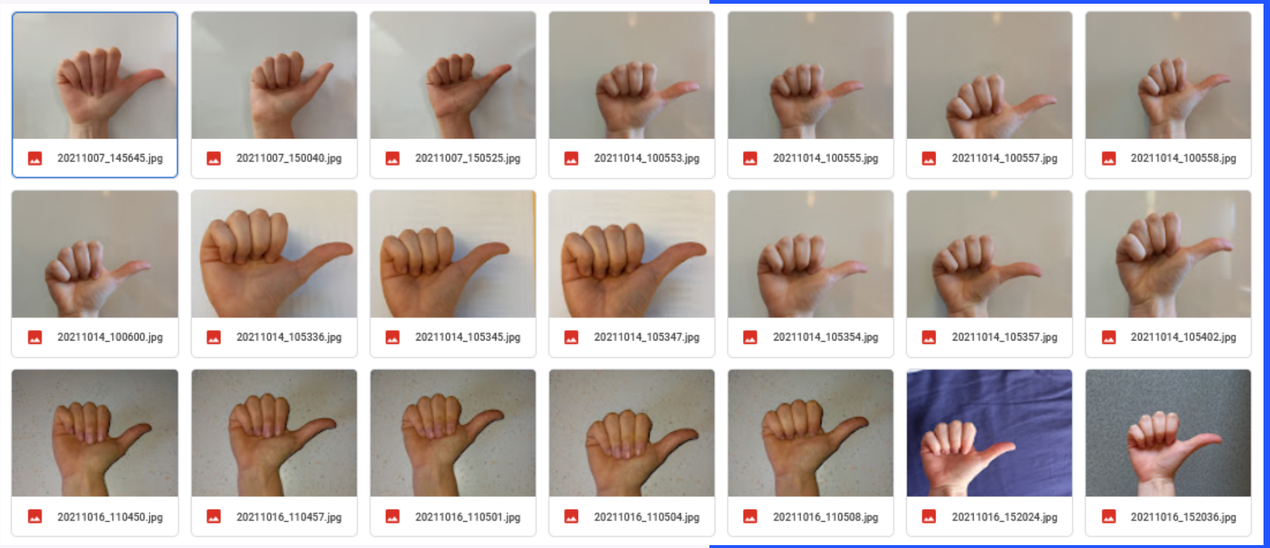
Ajout de convolutions

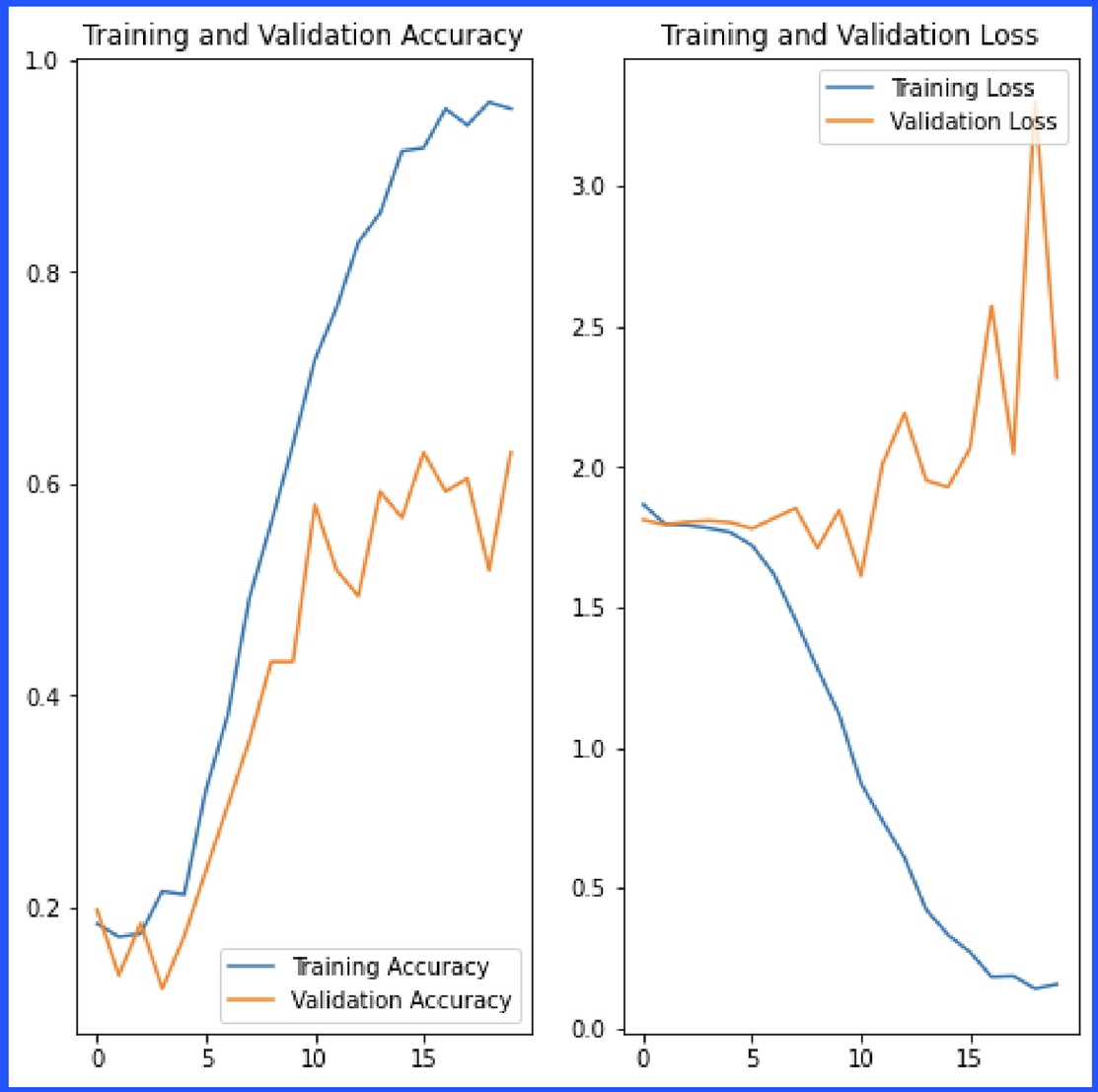
```
model = Sequential([
  layers.Rescaling(1./255, input_shape=(image_height, image_width, 3)),
  data_augmentation,
  layers.Conv2D(32, 3, padding='same', activation='relu'),
  layers.Conv2D(32, 8, padding='same', activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(64, 8, padding='same', activation='relu'),
  layers.Conv2D(32, 8, padding='same', activation='relu'),
  layers.Conv2D(64, 2, padding='same', activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(64, 3, padding='same', activation='relu'),
  layers.Conv2D(64, 3, padding='same', activation='relu'),
  layers.MaxPooling2D(),
  tf.keras.layers.GlobalAveragePooling2D(),layers.Dropout(0.2),
  layers.Flatten(),
  layers.Dense(128, activation='relu'),
  layers.Dense(num_classes)
])
```

Légère amélioration de l'accuracy mais temps de traitement excessivement long

Test du modèle

Ajout de photos





Ajustement du modèle

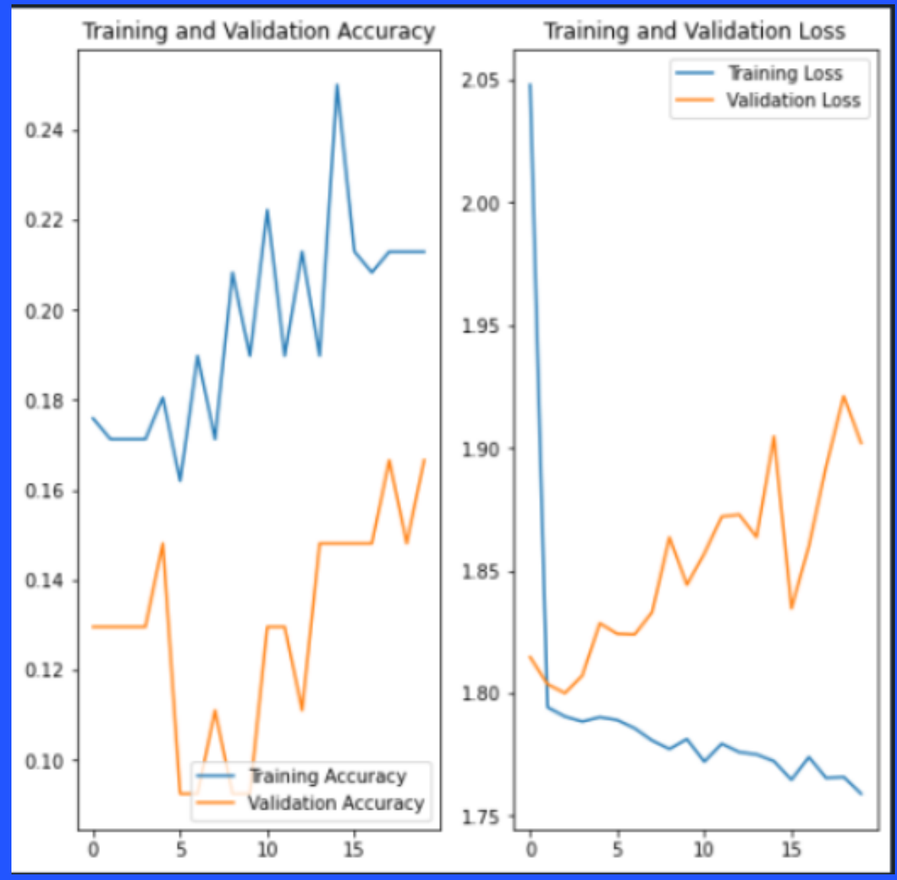
Modification des photos

- Rotation de 20°, retournement horizontal et réduction du contraste de 20%

Ajustement du modèle

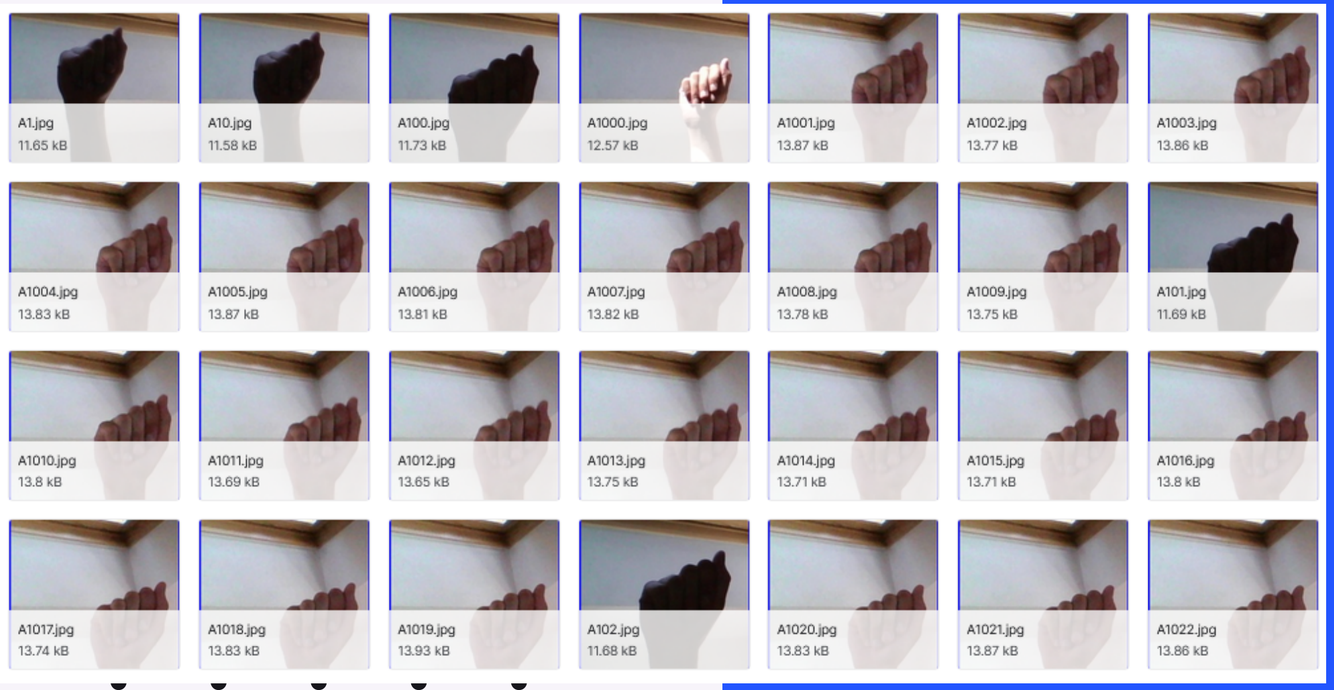
Modification des photos

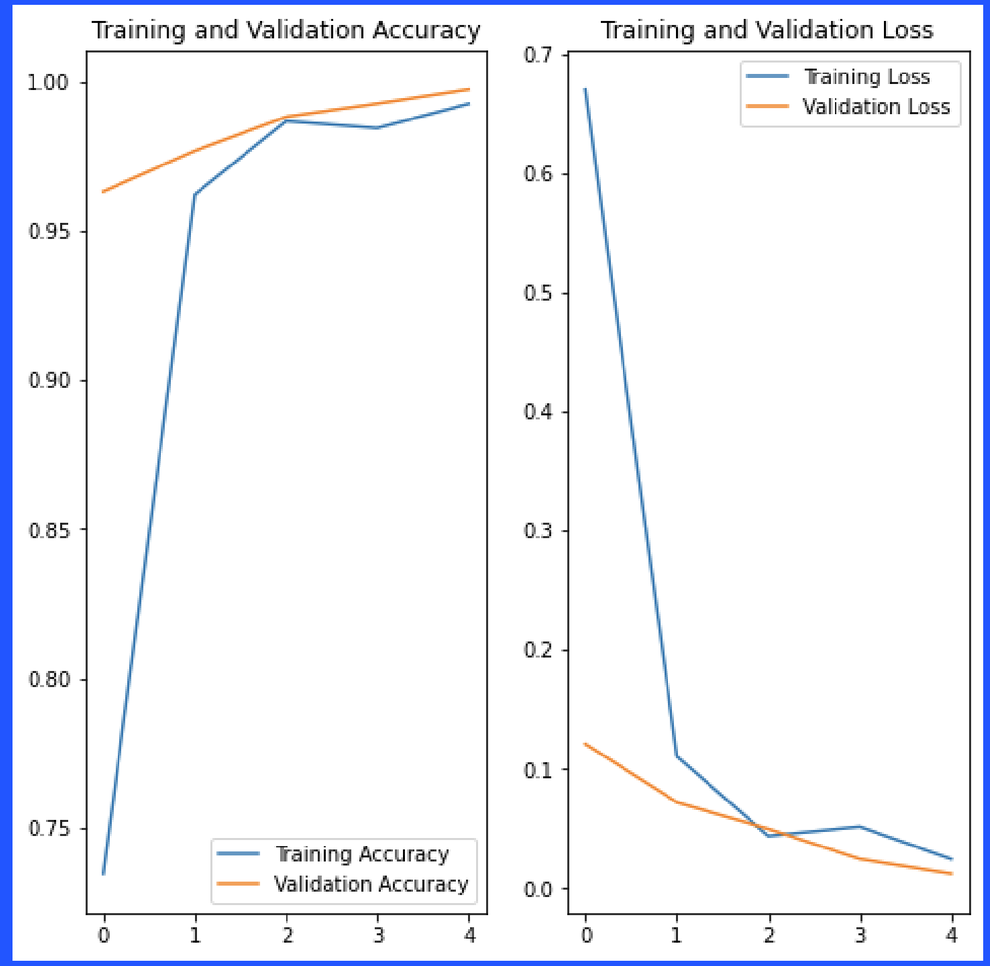
- Coupage et collage aléatoire



Application à un autre dataset

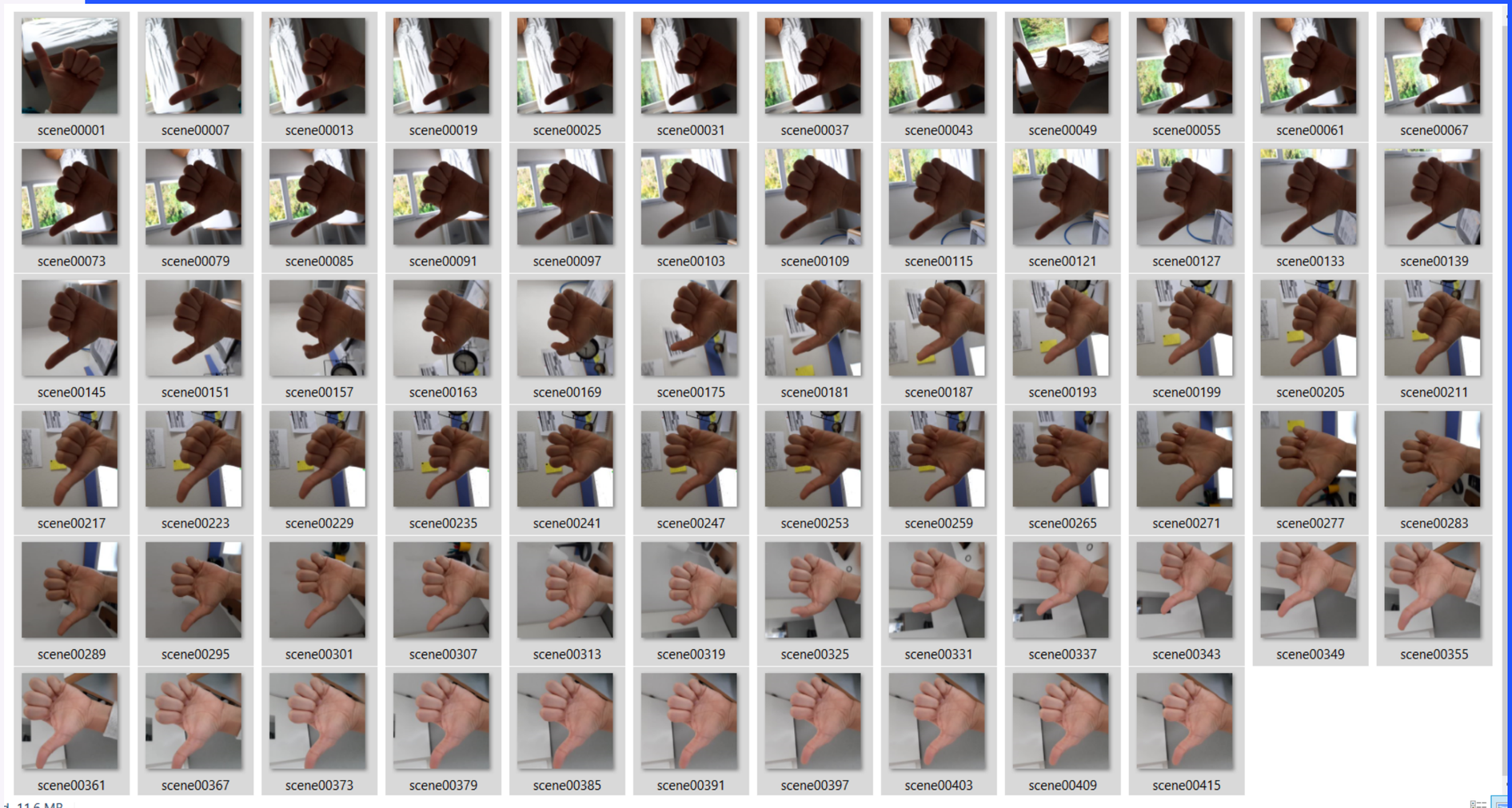
- American signs language
- ~ 3000 photos / lettre
- photos moins uniformes



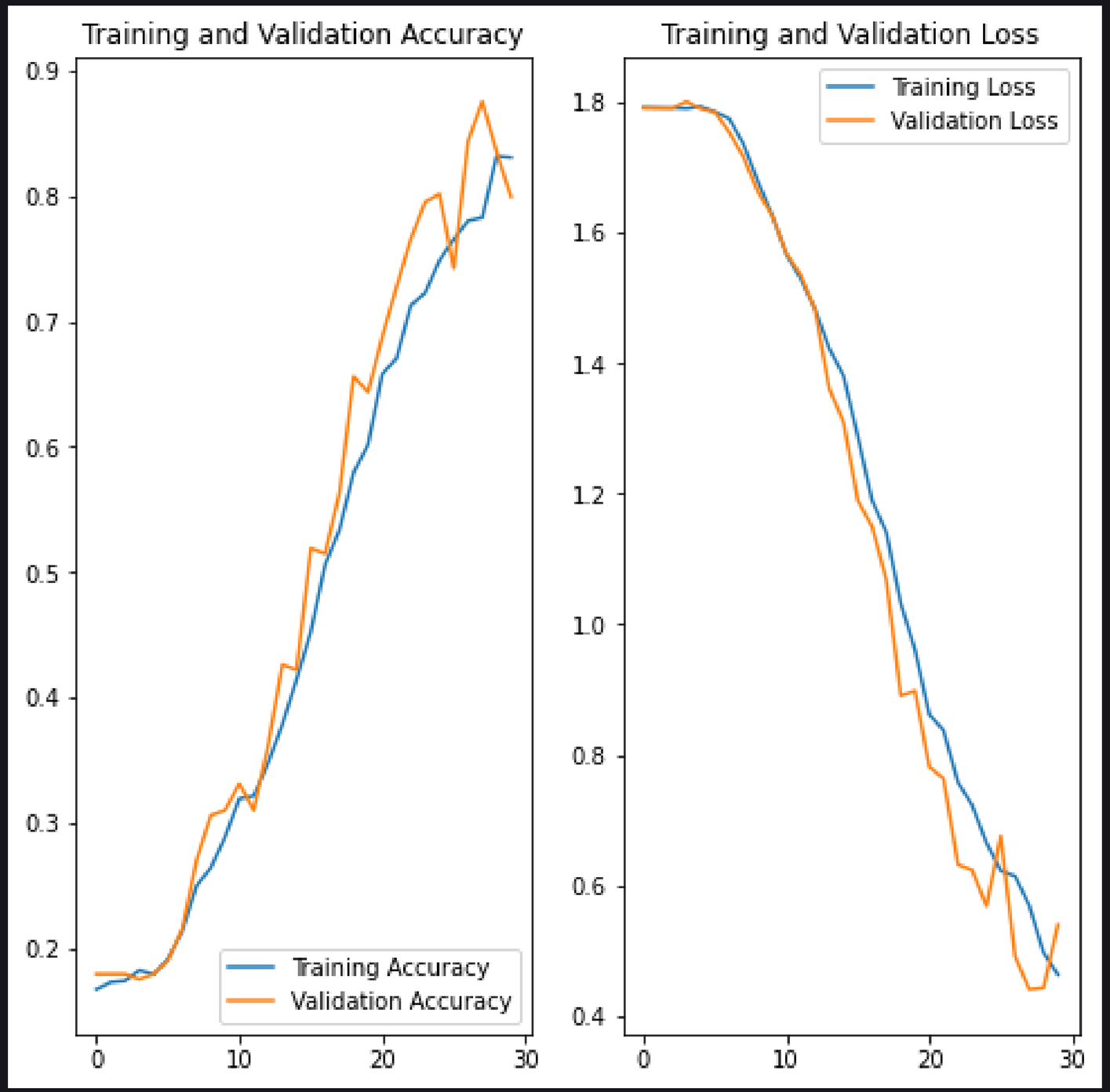


Amélioration du dataset original

Conversion des "frames" de la vidéo en photo individuelles

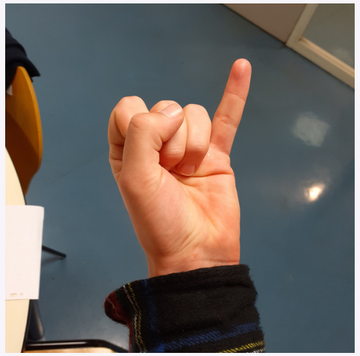


Application du modèle au nouveau dataset



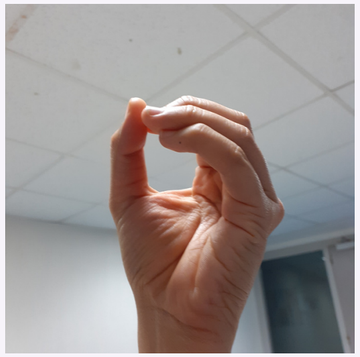
E 63.7%

A - 62.5%

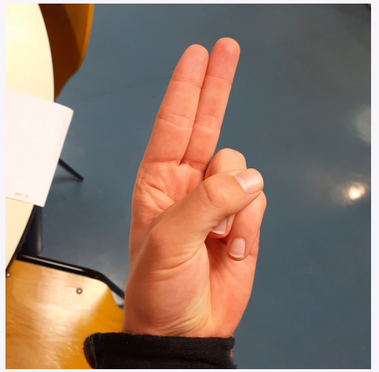


Y - 74.4%

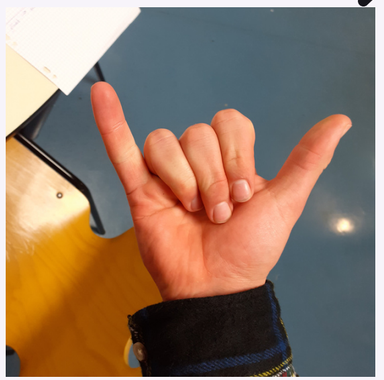
I



0 - 89.4%



U - 91.1%



A - 69.3%

Y

Conclusion