

Problématique

1. La segmentation sémantique avec l'imagerie infrarouge présente des défis spécifiques tels que le faible contraste, le bruit élevé et la sensibilité aux variations des conditions lumineuses
2. Les architectures UNet existantes ont tendance à surajuster, sont intensives en calcul et luttent pour maintenir la précision dans des conditions de faible visibilité [1]

Objectifs

- Le principal objectif est de modifier l'architecture UNet traditionnelle pour mieux s'adapter à l'imagerie infrarouge, en se concentrant spécifiquement sur les tâches de segmentation sémantique.
- Visant à améliorer la précision et l'efficacité informatique de la segmentation sémantique dans les images infrarouges, en particulier pour la détection de masques faciaux dans diverses positions.
- L'intégration de la convolution séparable dans l'architecture UNet réduit la complexité du modèle et améliore la vitesse de traitement sans compromettre les performances.

Bases de données et labellisation

- L'ensemble de données a été développé par notre équipe de recherche dans le cadre d'un sous-projet intitulé 'Partie Expérimentale'. Il comprend des images infrarouges capturées à l'aide d'une caméra FLIR.
- Les images ont été collectées auprès de différentes personnes, incluant une variété de zones et de positions du visage pour assurer la diversité (Fig.1).
- L'ensemble de données inclut des scénarios avec différents types de masques faciaux portés par les sujets, reflétant les conditions réelles.
- L'étiquetage de l'ensemble de données a été réalisé à l'aide de l'outil VGG Image Annotator, dont les détails sont illustrés dans la Fig.2.



FIGURE 1 – Exemple d'images de la base de données

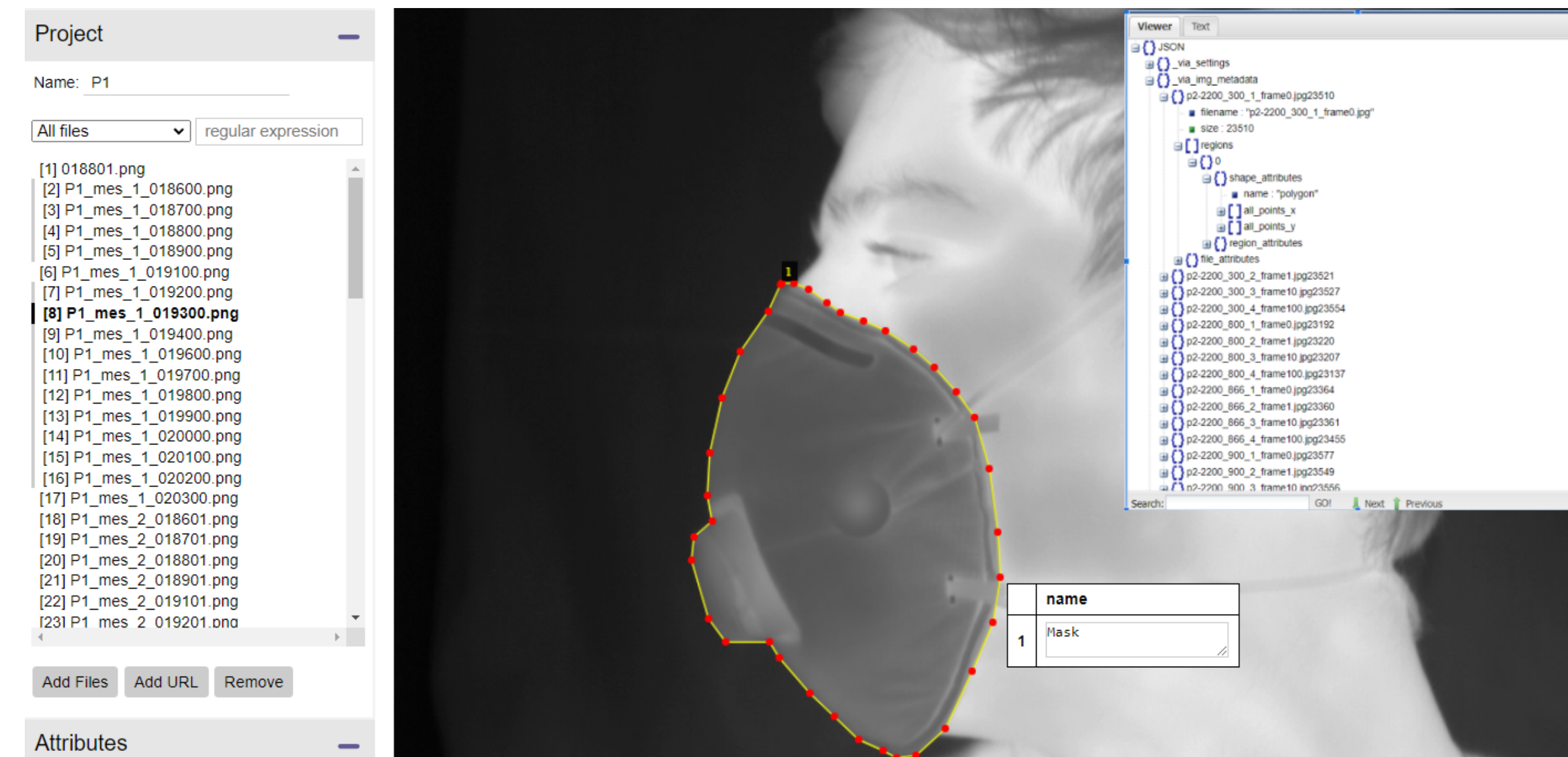


FIGURE 2 – Interface de l'annotateur d'images vgg

Modèle

- Utilise une structure UNet améliorée avec des blocs CNN et des convolutions séparables en profondeur, conçue pour gérer efficacement les tâches de segmentation sémantique dans l'imagerie infrarouge.
- Intègre des connexions directes et une régularisation par dropout au sein de chaque bloc CNN pour améliorer la stabilité de l'apprentissage et permettre des configurations de réseau plus profondes sans augmenter le coût computationnel
- Comprend un chemin de contraction pour le sous-échantillonnage et l'extraction de caractéristiques, ainsi qu'un chemin expansif avec des convolutions transposées pour une localisation précise et un suréchantillonnage.

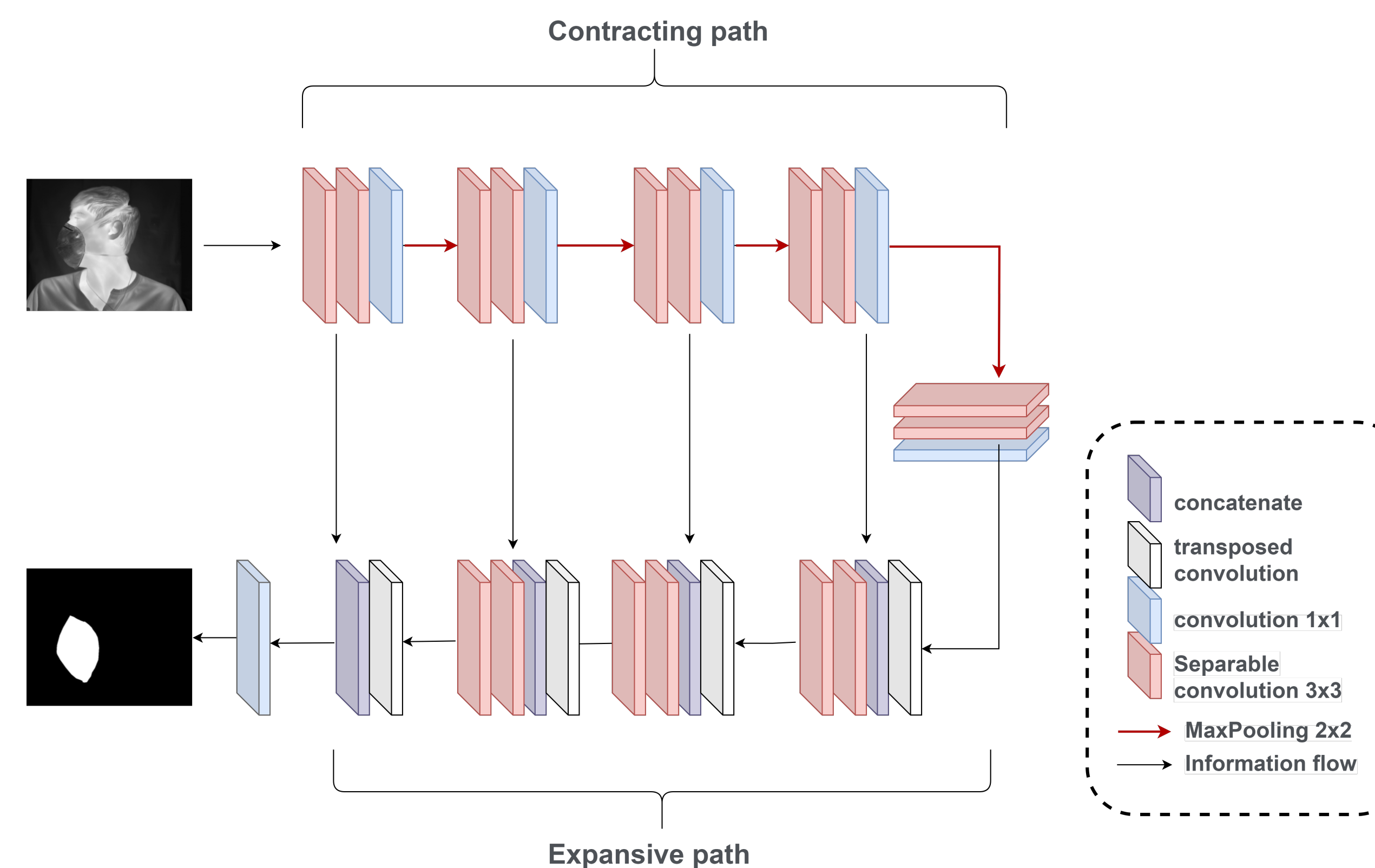


FIGURE 3 – Architecture du modèle proposé

Résultats

TABLE 1 – Comparaison entre la méthode proposée et les méthodes existantes

Method	Accuracy	mIoU	Training Time
Proposed Model	0.98	0.96	441s
Unet	0.98	0.95	660s
Unet++	0.94	0.78	500s
Mask R-CNN+ResNet [2]	0.90	0.89	-

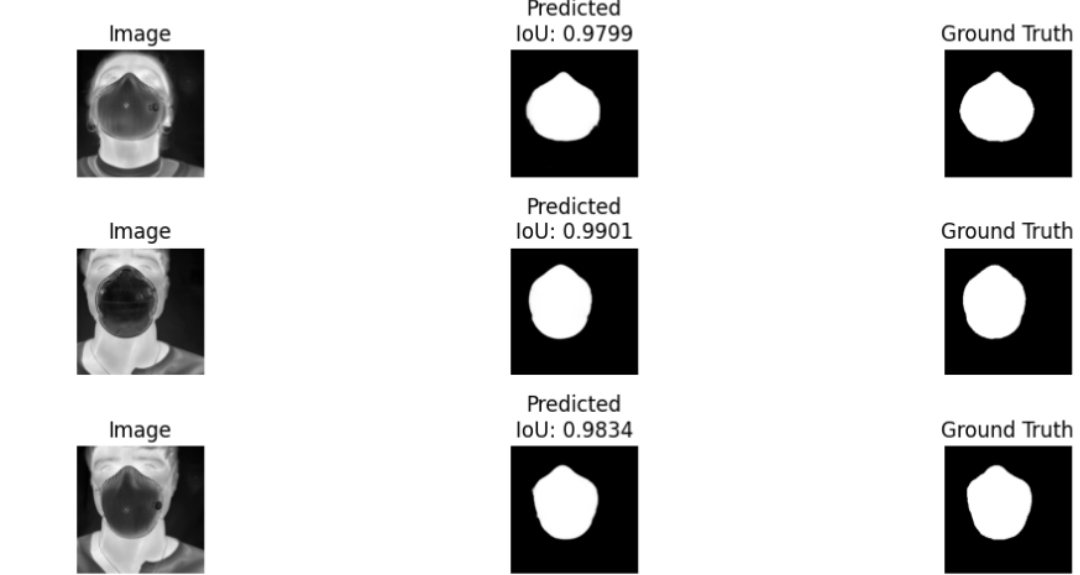


FIGURE 4 – Résultats de la segmentation

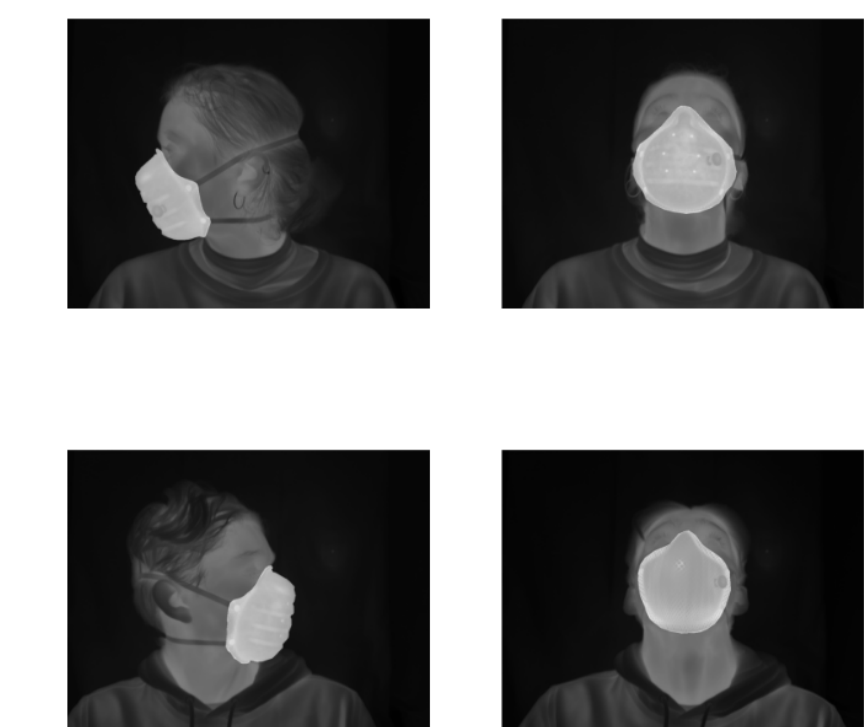


FIGURE 5 – Visualisation de la segmentation

Conclusion

- **Réduction de complexité** : L'intégration des convolutions séparables dans le modèle U-Net réduit le nombre d'opérations nécessaires, le rendant plus rapide et plus léger, idéal pour les applications en temps réel.
- **Utilisation flexible** : Le modèle fonctionne bien même avec de petites quantités de données et réduit la probabilité d'erreurs dues au surajustement.
- **Installation simplifiée** : Grâce à sa structure allégée, le modèle peut être installé et utilisé plus facilement sur de nombreux types d'appareils, facilitant son déploiement généralisé.
- **Travaux futurs – Suivi du masque en temps réel** : Ce modèle pourrait être développé pour suivre le masque en temps réel, même lorsque le porteur est en mouvement. Cela améliorerait significativement la capacité à détecter les fuites autour du masque

Références

- [1] Zülfiye Kütük and Görkem Algan. Semantic segmentation for thermal images : A comparative survey. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 286–295, 2022.
- [2] Min Zhang, Kai Xie, Yu-Hang Zhang, Chang Wen, and Jian-Biao He. Fine segmentation on faces with masks based on a multistep iterative segmentation algorithm. *IEEE Access*, 10 :75742–75753, 2022.