

A Deep Learning Solution for Infected Malaria's Cells Detection

Joan-Rufin Toukap

National Advanced School Of Engineering, MSP 1

Problématique

- En Afrique, beaucoup de personnes souffrent de maladies tropicales comme la Diarrhée, la typhoïde, la leishmaniose, lèpre **ainsi que le paludisme qui est le plus fréquent et également le plus mortel.**
- Le paludisme chaque année touche plus de **200 millions d'Africains et cause la mort de plus de 450 000 d'entre eux.**
- La principale cause de ce triste bilan annuel découle du fait que cette maladie à cause de la précarité et du sous-développement qui règne sur notre continent, n'est la plus part du temps pas détectée, ni soigné au bon moment. Ce qui laisse généralement la possibilité au parasite (**plasmodium**) de se développer et d'envahir le corps. Dès lors il devient difficile voir même impossible de sauver l'hôte.
- Ainsi pour résoudre ce problème, nous avons proposé un model de **computer vision** qui permet de détecter à partir d'images de cellules du frottis sanguin, la présence éventuelle du parasite du paludisme.

Motivations

- Permettre a tous les africains de se dépister facilement et régulièrement du paludisme afin de promouvoir la santé et le bien être de tous.

base de données

- Nous avons télécharger un Dataset de **KAGGLE contenant 26000** images de cellules du frottis sanguin de plusieurs hôtes infectées ou pas par le parasite du paludisme, réparties en proportion égale et qui ont été modifiées pour faciliter le processus.
- À partir de celui-ci nous avons aléatoirement créé deux Dataset : **l'entraînement(80%), le test(20%).**

Architecture

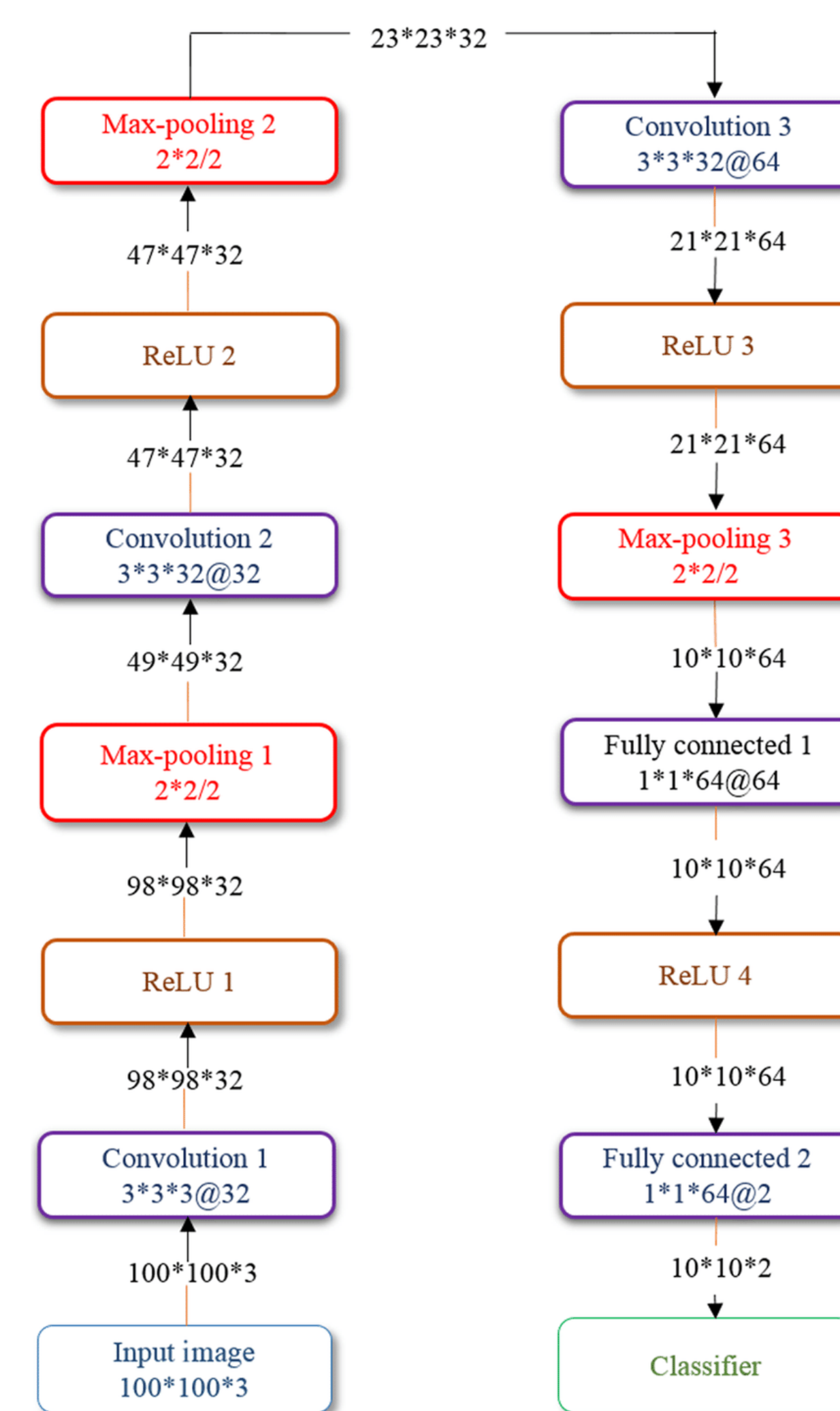


Figure 1:Architecture du Model

Mode Opératoire

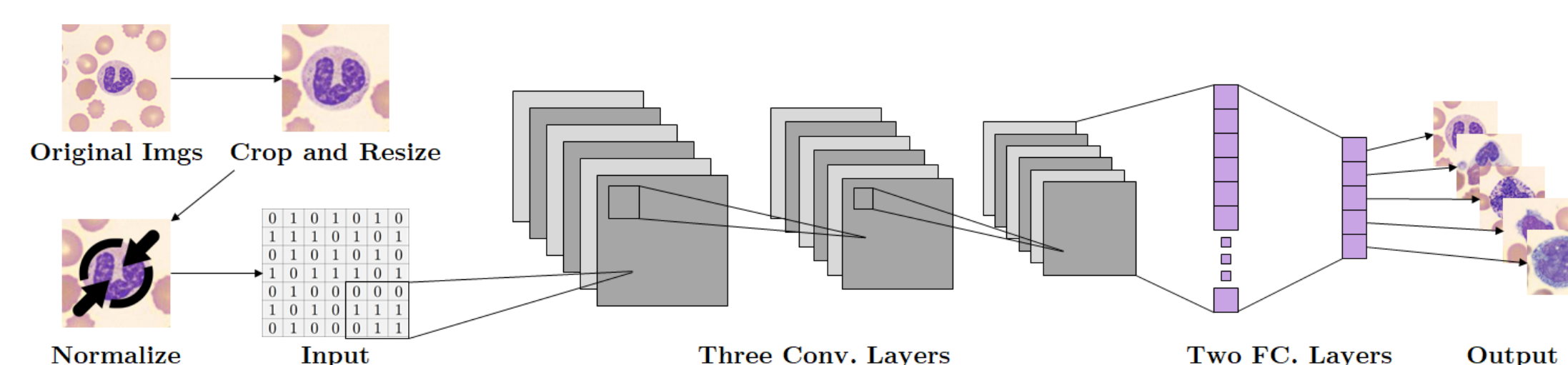


Figure 2:Vue d'Ensemble

- pour concevoir ce model, nous avons utiliser deux libraires de vision par ordinateur conguent par facebook: torchvision pour redimensionner et retoucher les images, et Pytorch pour mettre en place les différentes couches convolutives et entièrement connectées présentées plus haut.
- L'ensemble est relié à une focntion softmax (calcule la probabilité exponentielle des éléments d'une matrice) qui nous retourne une matrice de poids de 2 éléments le premier correspondant à la classe non infectée et le second a celle infectée. Ainsi le model fait une prediction vers la classe representée par le poids le plus grand.

Entraînement, Test Et Résultats

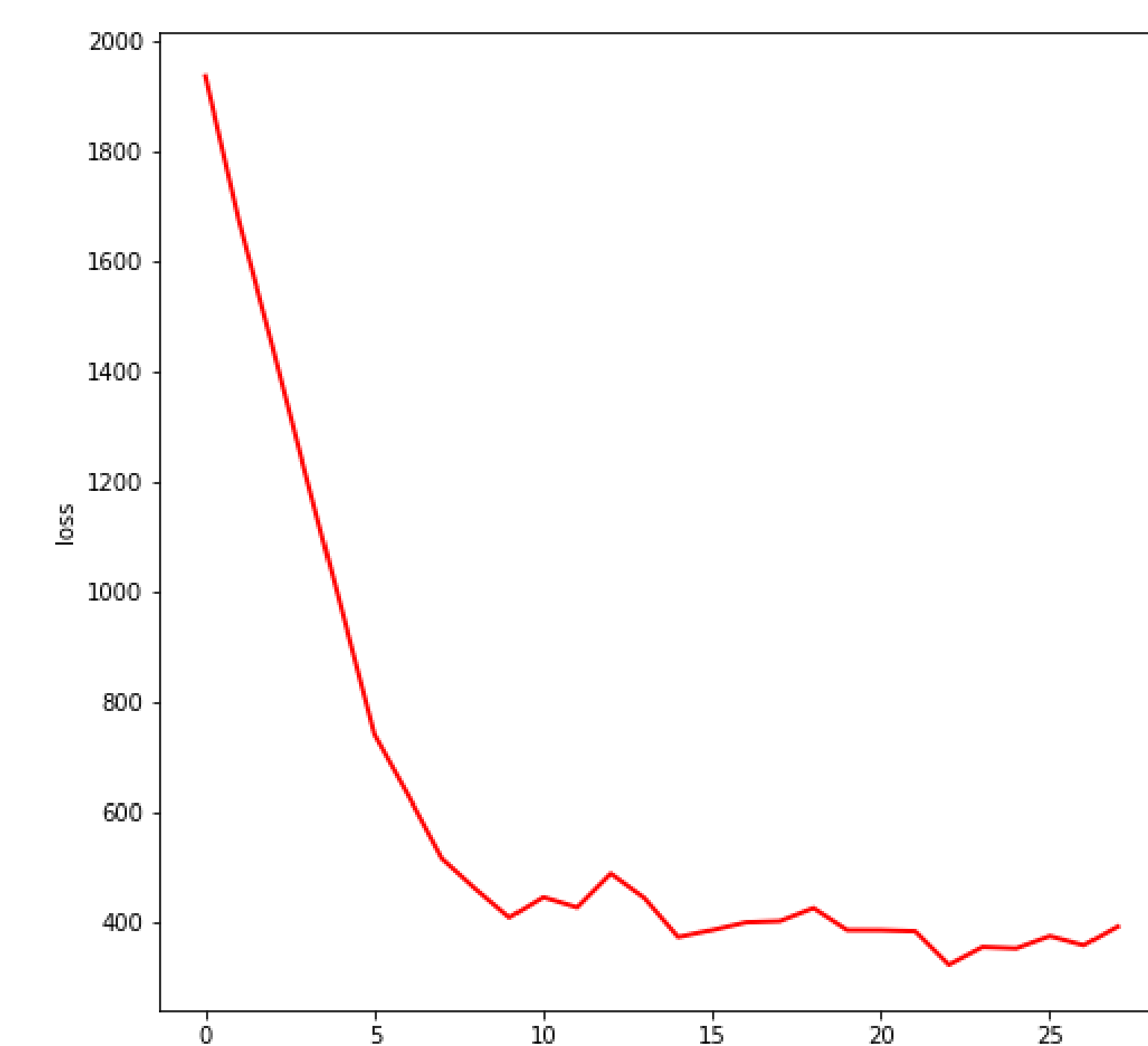


Figure 3:Evolution de la Fonction de Coût.

```
train_Accuracy = 96.6%
confusion_matrix: [[10681  317]
 [ 431 10617]]
test_Accuracy = 95.5%
confusion_matrix: [[2668  113]
 [ 136 2595]]
program parameters save to: .\save_params
```

Figure 4:Pourcentages de Précision Et Matrices de confusion.

La matrice de confusion permet pour une catégorie d'images, de montrer le nombre correctement classées et classées ailleurs après prédiction par le model. Exemple: 10681 represente les nombre d'images d'entraînement non infectées correctement classées tandis que 317 representent les images de cette même catégorie mais classées comme images de cellules infectées

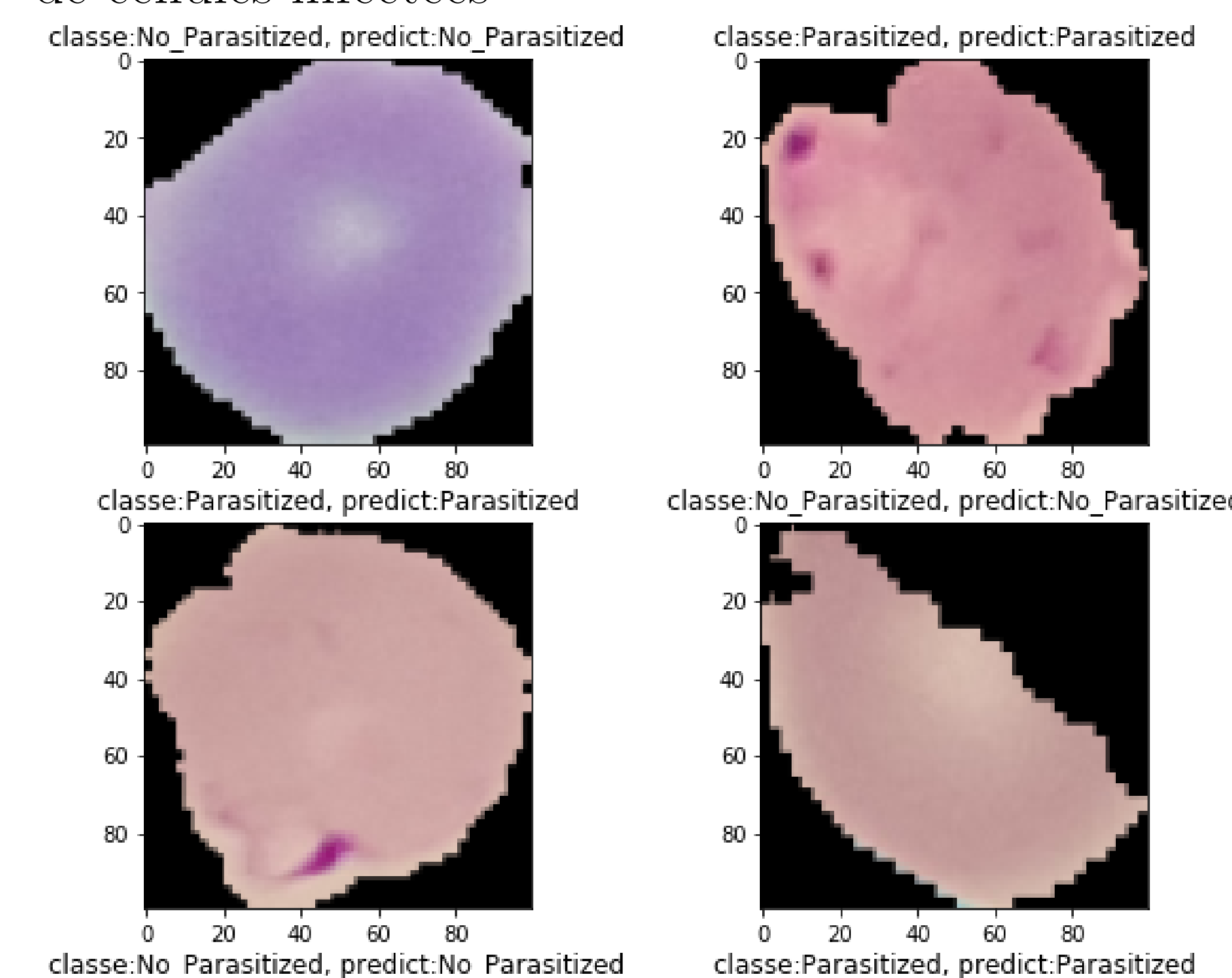


Figure 5:Quelques Exemples de prédictions de Test.

Conclusion et Améliorations Futures

- Avec ce model nous mettons l'intelligence et les datas au service d'une vision futuriste de la santé.
- Desormais, il sera plus facile pour nos valeureux laboratins de dépister le parasite du paludisme dans le sang et donc de sauver des vies.
- 1 Nous continuons à rechercher des bases de données afin d'améliorer le projet pour y inclure la reconnaissance de plusieurs autres maladies.
- 2 De fabriquer des algorithme de segmentation ceci dans le but de pouvoir localiser les parties de cellules infectées et de pouvoir mieux orienter les recherches de maladie dans le sang.
- 3 Enfin l'idée derrière ce projet c'est de pououvoir créer un microscope intelligent fonctionnant sur ces algorithmes et capable d'aider les medecins à aller plus vite dans la recherche des maladies chez leur patients.

Travail Simimaire

- <https://www.kaggle.com/mrudhuhas/malaria-cell-images-classification-cnn> Fait par kaggle sur l'architecture CNN mis en place grâce à Tensorflow(GOOGLE)

Contact Information

- tel: +237 6 55 09 11 71
- toukaprufin@gmail.com