

Epilab / Détecteur de cellules

Certification “Gérer un projet en mobilisant les méthodes agiles”

07/06/2023

Rapport rédigé par :
✶ LECROART, Antoine

SOMMAIRE

Résumé exécutif	3
Introduction	3
Contexte du projet	3
Analyse du projet	4
Compréhension du besoin	4
Réponse technique	4
L'organisation Agile	4
La réalisation technique.....	5
Enjeux de la mise en oeuvre	5
Justification des choix	5
Les choix Agile.....	5
Les choix Techniques.....	6
Problèmes rencontrés	6
Les problèmes Agiles	6
Les problèmes techniques	7
Solutions apportées.....	7
Évaluation du projet	7
Bilan du projet.....	7
Améliorations envisagées	8
Conclusion	8
Apprentissages, perspectives pour le projet professionnel.....	8

Résumé exécutif

Durant ma formation en développement d'Intelligence Artificielle au GRETA, j'ai pu effectuer un stage de trois semaines chez Epilab, une entreprise située à Palaiseau, qui développe un test pour diagnostiquer la tuberculose.

Cette entreprise utilise des microbilles magnétiques pour capturer les cellules et ainsi répliquer l'effet d'une centrifugeuse à moindre coût. Afin de tester l'efficacité de ces billes, il est nécessaire d'analyser un grand nombre d'images issues de tests, processus qui serait grandement facilité par l'IA.

J'ai donc réalisé un programme capable de détecter et de compter ces cellules en mobilisant des méthodes agiles.

Introduction

Contexte du projet

L'entreprise Epilab fait de la Recherche & Développement afin de créer un test portatif pour détecter la tuberculose. Son domaine d'activité est donc la Biotechnologie.

Afin de créer un test portatif pour détecter la tuberculose, il est nécessaire de "capturer" les cellules à étudier (les regrouper). La méthode classique consiste à centrifuger la solution contenant les cellules, ce qui les fait s'accumuler au fond de la fiole, pour ensuite les prélever. Une centrifugeuse est coûteuse et nécessite de l'électricité pour fonctionner, il fallait donc trouver une solution à faible coût et utilisable dans n'importe quel environnement.

Ainsi, Epilab utilise une technique qui consiste à ajouter des microbilles magnétiques dans la solution, qui vont se fixer aux cellules. Il suffit ensuite d'aimanter la solution, ce qui va capturer les cellules.

Epilab cherche actuellement à appliquer ce même processus aux cellules cancéreuses issues du poumon, et réalise de nombreuses expériences de test.



Figure 1. Schéma expérimental d'un protocole de capture de cellules eucaryotes par des billes microbilles magnétiques.

Pour analyser les résultats de ces tests, les chercheurs doivent compter les cellules capturées et non-capturées en utilisant une technique appelée "comptage sur cellules de Malassez". Le comptage sur cellules de Malassez est effectué sur plusieurs échantillons (jusqu'à 70 parfois), rendant ce processus très long et fatigant, car l'observateur doit analyser les cellules directement à l'oculaire du microscope. Ce dernier, émettant une lumière forte et claire, entraîne la fatigue des yeux au cours du temps de manipulation.

L'IA est nécessaire pour cette entreprise afin de compter les cellules pour simplifier et optimiser grandement le travail des chercheurs.

Analyse du projet

Compréhension du besoin

Les besoins de ce projet ont été établis dans un cahier des charges où les besoins ont été divisés en plusieurs objectifs, certains primaires et d'autres secondaires, par ordre d'importance.

Objectifs primaires :

- Compter des cellules avec une fiabilité de 95% minimum (5% d'erreur) par image.
- Pouvoir analyser tout type d'image (avec ou sans billes, matrice complexe ou non).
- Facile d'utilisation (interface utilisateur ou un protocole utilisation clair et précis).
- Modifiable ultérieurement (partage du code).

Objectifs secondaires :

- Calcul de l'écart-type du résultat (les échantillons sont toujours faits en triplicata, il est donc possible de déterminer l'écart-type sur un triplicata).
- Calcul d'un ratio de capture : nombre de cellules capturées (comparaison culot et surnageant).
- Calcul du recouvrement des cellules par les billes magnétiques.
- Résultats des calculs sous forme csv pour être analysées sur Microsoft Excel.

Réponse technique

L'organisation Agile

Dans le but de réaliser ce projet de la meilleure façon, nous avons mis en place une méthode Scrum allégée en raison de l'échelle de temps réduite et étant le seul membre de l'équipe de développement :

- le client était l'ensemble de l'équipe de chercheurs, à qui profitera le programme réalisé,



- ma tutrice a joué le rôle de product-owner en rédigeant un cahier des charges qui fut découpé en user-stories en début de stage dans un backlog (sur Click-Up),
- J'ai ainsi joué le rôle de Scrum-master/dev en développant l'application et en animant les reviews, rétrospectives et daily-meetings.

Chaque semaine représentait un sprint avec une review et rétrospective le vendredi. Nous avons aussi effectué des daily-meetings quotidiens pour aborder les users-stories qui seront traitées dans la journée.

La réalisation technique

Pour répondre au besoin du client, j'ai décidé de réaliser une interface graphique avec PyQt, mettant en place un modèle IA YOLOv8. Le programme a ensuite été compilé en exécutable avec PyInstaller.

L'application permettait de charger les images ordonnées selon des marqueurs dans leur nom. Puis les images pouvaient être traitées une par une ou en lot de six pour la détection. On pouvait ensuite analyser ces détections pour en tirer des conclusions sur l'efficacité du test. Enfin, un tableau regroupait les informations chiffrées (nombre de cellules, écart-type, ratios et concentration). Ces résultats étaient exportables en format xlsx ou csv.

Enjeux de la mise en oeuvre

Justification des choix

Les choix Agile

Ce projet se déroulant sur trois semaines, nous avons choisi de réaliser un sprint chaque semaine. Au préalable, nous avons élaboré un backlog, où les objectifs furent disséqués en tâches et sous-tâches à accomplir.

Lors d'un court entretien les lundis matin, nous avons évalué quelles tâches traiter en priorité pour le reste de la semaine. Puis chaque jour, je notifiais ma tutrice de mon avancement, des problèmes rencontrés et des tâches qui seront traitées dans la journée (à la manière d'un daily meeting).

Les vendredis, nous avons effectué des reviews où je présentais le travail accompli, mais aussi des backlogs groomings car il fut nécessaire parfois de remanier, supprimer ou ajouter certaines tâches. Nous avons aussi abordé les questions d'organisation et de communication à la manière d'une rétrospective.

De manière générale, nous avons opté pour une approche scrumban, car au fur et à mesure de l'avancement, certains points devaient être modifiés.

Les choix Techniques

Étant donné la durée courte de ce projet, il était nécessaire d'utiliser des outils simples et déjà connus. Ainsi, je me suis orienté vers une interface graphique PyQt, en partie réalisé sur QtDesigner, et un modèle IA Yolo d'ultralytics.

PyQt est une librairie largement utilisée et connue dans le développement d'applications. Elle est diverse et permet une grande liberté dans la création d'applications. De plus, avec l'utilisation de QtDesigner, il est assez simple de réaliser un template de l'application qui sera transformé en code par la suite.

Il existe de nombreuses méthodes et modèles IA pour accomplir l'objectif du projet. Néanmoins, il m'était nécessaire de le réaliser en peu de temps. La librairie ultralytics propose des modèles IA simples et rapides à utiliser et entraîner dans le cadre de la détection d'objets (ici, les objets sont les cellules). Je savais que Yolo était utilisé pour la détection d'objets sur des images satellites, donc capable de détecter beaucoup de petits objets sur une image. Le modèle Yolo existe en plusieurs versions (v3, v5, v7,...) et en plusieurs sous-modèles (n, s, m, l,...). J'ai choisi le modèle v8 x, modèle le plus récent et puissant. La vitesse de détection est impactée par la taille du modèle, mais dans le cadre du projet, quelques secondes de plus pour effectuer la détection n'a pas d'impact négatif.

Étant important que l'application soit simple à utiliser, je l'ai compilé en exécutable avec Pyinstaller afin de faciliter l'installation du logiciel pour Epilab, le client.

Problèmes rencontrés

Les problèmes Agiles

Dans la globalité, il n'y a pas eu de problèmes d'organisation du point de vue de l'équipe, le travail de développement n'étant effectué que par moi-même.

Bien que les méthodes Agiles étaient inconnues de l'entreprise, nous avons pu adopter une communication concise et efficace entre l'équipe (moi), le product-owner (ma tutrice) et le client (Epilab). Les problèmes techniques qui sont apparus ont pu ainsi être efficacement abordés, et des mesures correctionnelles rapidement appliquées.



Les problèmes techniques

Tout d'abord, la capture des images était effectuée avec un smartphone placé dans l'objectif du microscope, donnant des images avec un cadrage légèrement différent à chaque fois, ainsi qu'une déformation optique de l'image (effet fish-eye). De plus, chaque chercheur utilisait son smartphone personnel, les images étaient donc légèrement différentes en résolution et en couleurs pour chaque smartphone.

Puis, lors de la préparation des images, il s'avérait que les cellules étaient trop peu visibles, ce qui a impacté ma capacité à préparer les données d'entraînement et les performances de l'IA.

Solutions apportées

Au terme de la première review, nous avons conclu qu'il était nécessaire de standardiser l'acquisition des données. Epilab a donc investi dans une caméra spécialement conçue pour le microscope, permettant ainsi d'avoir le même cadrage et échelle de couleur sur l'ensemble des données, ainsi qu'une meilleure résolution et l'absence de déformation optique.

Nous avons aussi établi qu'il serait nécessaire d'appliquer une coloration aux cellules, afin qu'elles soient plus visibles. En effet, pour entraîner l'IA, il est nécessaire d'avoir un jeu de données où les objets à identifier sont labellisés, ce qui va permettre à l'IA d'apprendre la nature de ce qu'elle doit détecter. Il était donc nécessaire de colorer les cellules afin que je puisse préparer les meilleures données possibles et aider l'IA à faire de bonnes détections.

Évaluation du projet

Bilan du projet

Ce projet fut un succès sur de nombreux points. La mise en place des méthodes Agiles a eu un impact très positif sur la structure du travail et la résolution des problèmes, rendant ainsi possible un aboutissement conséquent sur une si courte période.

L'application fonctionne, elle permet de traiter six images à la fois (deux images par échantillon, et trois échantillons pour établir une moyenne), de calculer la concentration en cellules sur chaque image, de calculer les écart-types et ratios de captures. Ces résultats peuvent ensuite être enregistrés en format xlsx ou csv. Ces points correspondent à l'ensemble des users-stories créées.

Néanmoins, la nécessité de colorer les cellules a grandement impacté la réalisation du projet. Bien que ce point fut abordé assez tôt, lors de la première review, l'équipe de chercheurs a dû expérimenter longuement pour établir un protocole fiable, sans y parvenir. Il existe plusieurs types de cultures sur lesquelles le programme devait fonctionner, et la coloration n'a marché que



sur un seul. L'application est donc fonctionnelle pour ce type de culture, mais pas sur les autres. Par conséquent, l'un des objectifs primaires n'a pas été accompli.

Améliorations envisagées

Malgré la réussite de la partie développement, le projet n'a pas pu aboutir en raison d'un manque de temps dans l'acquisition des données. Afin qu'Epilab puisse améliorer le programme pour qu'il soit fonctionnel sur tout type de cultures, je leur ai fourni un protocole détaillé pour qu'ils puissent eux-mêmes entraîner l'IA et donc l'améliorer, une fois qu'un protocole de coloration fiable sera trouvé.

D'un point de vue Agile, la mise en place de la méthode Scrum(-ban) aurait été bien plus efficace si ce projet avait été réalisé par une équipe de développement et non par ma simple personne. Néanmoins, l'adapté à notre cas s'est montré tout de même largement bénéfique pour la détection rapide de problèmes et ainsi adapter le projet.

Conclusion

Apprentissages, perspectives pour le projet professionnel

Ce stage m'a beaucoup appris sur le développement d'applications IA et dans la mobilisation des méthodes Agiles. Mes connaissances dans le domaine de vision par ordinateur, ainsi quand dans les interfaces graphiques, se sont grandement améliorées et je tire une certaine expérience sur l'organisation de projets et sur la communication entre parties.

J'ai aussi beaucoup appris sur le domaine du Biomédical et de la recherche scientifique. Souhaitant m'orienter dans ce domaine, ce stage constitue un réel atout pour la suite de ma carrière professionnelle et je suis confiance plus confiant dans ma capacité à travailler en équipe avec la méthode Scrum.

