

Rapport de Projet

OCR PRO MAX

December 8, 2024

Créé par :

Alexis Chafaï

Lorys Févry

Otto Debie

Baptiste Mandry-Delimoges

Table des matières

1	Introduction	3
2	Lorys, détection des éléments	4
2.1	Présentation et Motivation	4
2.2	État d'avancement	5
2.3	Aspects techniques	7
3	Baptiste, découpage et sauvegarde des lettres	9
3.1	Mon Parcours	9
3.2	Le projet, et ses contraintes	12
3.3	Les bénéfices moraux du projet	15
4	Otto, réseau de neurones	17
4.1	Présentation et Motivation	17
4.2	État d'avancement	19
4.3	Aspects techniques	20
5	Alexis, traitement de l'image	22
5.1	Présentation et Motivation	22
5.2	État d'avancement	23
5.3	Aspects techniques	24
6	Tableau d'avancement	26
7	Conclusion	27

1 Introduction

Ce rapport a pour but de décrire les réalisations de notre groupe pour cette première soutenance du projet d'OCR du S3. Pour rappel OCR en français est traductible par reconnaissance optique de caractères. L'objectif de notre projet est donc de coder un algorithme d'OCR appliqué à une grille de mots mêlés afin qu'il puisse la résoudre en retrouvant tout les mots dedans. Plusieurs étapes sont nécessaires à la construction de ce projet. Nous en avons identifié trois principales que nous nous sommes repartis de la manière suivante : Alexis, s'occupe de la partie interface avec la fenêtre et le traitement de l'image, Otto s'occupe du réseau de créer le réseau de neurones et de son entraînement enfin Lorys et Baptiste s'occupent de la reconnaissance de la grille chacun travaillant une partie différente de l'autre. Dans la suite de ce rapport, nous détaillons chacun notre travail ainsi que notre ressenti à propos de ce projet mais aussi plus généralement de notre rapport à l'informatique.

2 Lorys, détection des éléments

2.1 Présentation et Motivation

Motivé par le domaine de l'intelligence artificielle (IA), j'ai un fort intérêt pour les systèmes de reconnaissance d'images et les algorithmes de traitement de données visuelles. Ce projet, impliquant un logiciel de type OCR (Optical Character Recognition) pour résoudre une grille de mots cachés, m'a immédiatement attiré, car il représente un défi dans la création de systèmes capables d'analyser et d'interpréter des informations visuelles, un domaine où l'IA peut apporter de nombreuses solutions innovantes.

Dans cette perspective, j'ai pris en charge la tâche de détection de la grille, étape cruciale qui vise à identifier les contours et les cases de la grille de lettres sur l'image d'entrée. Cette tâche, bien que complexe, se rapprochait de mes intérêts pour l'IA, car elle nécessitait de comprendre la structure de la grille et d'analyser ses motifs visuels. Au-delà des compétences techniques, ce travail m'a également permis de développer des compétences en recherche de méthodes et en utilisation de la bibliothèque SDL, que j'ai explorée en profondeur.

2.2 État d'avancement

La première avancée majeure du projet a été l'implémentation de la binarisation des images. Ce processus, essentiel en traitement d'images, consiste à convertir chaque pixel en noir ou blanc en fonction d'un seuil prédéfini. La binarisation permet de simplifier les données en réduisant l'image à deux niveaux de valeurs, ce qui facilite la détection des structures comme les lettres et les grilles. Le choix d'un seuil ajustable s'est révélé crucial pour s'adapter aux variations de luminosité et de contraste dans les images d'entrée. Des tests ont été réalisés pour ajuster ce seuil en fonction des conditions spécifiques des images fournies, ce qui a conduit à une binarisation optimale dans la majorité des cas.

Un élément clé de cette avancée a été l'intégration d'un algorithme de propagation, conçu pour explorer les pixels connectés et regrouper les zones cohérentes dans l'image binarisée. Cet algorithme fonctionne en identifiant des régions continues de pixels noirs (ou blancs, selon le fond), les considérant comme des entités distinctes, comme des cases ou des lettres. La propagation a permis une détection efficace des cases et des lettres dans les grilles. Elle a également été perfectionnée pour exclure les petites zones non pertinentes ou les zones trop grandes qui ne correspondent pas aux dimensions habituelles des lettres. Ce développement a constitué une avancée majeure dans l'identification des structures d'intérêt.

Une fois les zones d'intérêt détectées, une étape de coloriage des régions sur l'image originale a été mise en œuvre. Cette fonctionnalité permet d'encadrer les régions détectées avec des bordures colorées (par exemple, en bleu), facilitant ainsi la validation visuelle des résultats. Cette méthode a servi d'outil de débogage et de validation, permettant d'identifier rapidement les cas où l'algorithme pourrait échouer ou mal détecter des lettres ou des cases. Cette visualisation a également permis d'identifier des zones où des ajustements de paramètres étaient nécessaires, notamment pour le seuil de binarisation ou les critères de taille des régions.

Une autre avancée significative a été la mise en place d'un processus de standardisation des régions détectées au format 28x28 pixels. Ce format est un standard largement utilisé dans les systèmes de reconnaissance de caractères,

notamment dans les modèles d'apprentissage automatique comme les réseaux de neurones. Chaque région détectée (par exemple, une lettre) est redimensionnée en conservant ses proportions tout en s'adaptant à la taille fixe de 28x28 pixels. Cette uniformité garantit que les données extraites peuvent être directement utilisées dans des modèles OCR ou dans des systèmes de reconnaissance de texte existants. Cette étape de redimensionnement a également été optimisée pour réduire les distorsions, ce qui est crucial pour préserver les caractéristiques visuelles des lettres.

Enfin, un effort particulier a été consacré au prétraitement des images avant la binarisation. La conversion des images en niveaux de gris (grayscale) constitue une étape essentielle pour uniformiser l'image et réduire les variations de couleur ou de luminosité. Cela permet de simplifier l'image et d'améliorer la précision de la binarisation. En convertissant chaque pixel en une seule composante d'intensité lumineuse, la conversion en niveaux de gris facilite également l'identification des contours des lettres et des lignes de la grille. Cette étape, bien que préliminaire, a un impact direct sur la qualité des résultats des étapes suivantes.

2.3 Aspects techniques

La binarisation repose sur l'utilisation d'un seuil ajustable pour convertir les pixels en noir ou blanc. La flexibilité offerte par cet ajustement permet de s'adapter aux variations spécifiques des images d'entrée, notamment en termes de luminosité et de contraste. L'implémentation utilise des bibliothèques de traitement d'images pour accéder directement aux pixels et appliquer des transformations globales ou locales. L'un des défis techniques a été de choisir un seuil optimal pour garantir une séparation nette entre les lettres et le fond, tout en minimisant les erreurs dues aux ombres ou aux variations de l'arrière-plan.

L'algorithme de propagation a été conçu pour explorer les connexions entre pixels afin de regrouper les zones pertinentes. Techniquement, l'algorithme utilise une pile ou une file pour parcourir les pixels adjacents et déterminer la taille et les limites de chaque région détectée. Chaque région est ensuite caractérisée par ses coordonnées minimales et maximales (`min_x`, `min_y`, `max_x`, `max_y`) ainsi que par ses dimensions (largeur et hauteur). Pour garantir une détection fiable, des contraintes géométriques ont été ajoutées : les zones trop petites ou trop grandes, ou celles dont le rapport largeur/hauteur est anormal, sont exclues. Ces ajustements ont permis d'améliorer significativement la précision de l'algorithme.

Le coloriage des zones détectées a été implémenté pour faciliter la visualisation et le contrôle des résultats. Chaque région validée par l'algorithme de propagation est encadrée sur l'image originale avec une bordure de couleur (par exemple, bleue). Cela est réalisé en modifiant directement les pixels de l'image originale en utilisant des outils de manipulation pixel par pixel. Cette fonctionnalité a non seulement permis de valider les résultats de l'algorithme, mais elle a également servi à identifier rapidement les erreurs potentielles ou les zones où des ajustements étaient nécessaires.

La mise en forme des zones détectées au format 28x28 pixels a été réalisée à l'aide d'une technique de redimensionnement proportionnel. L'algorithme prend en compte les proportions des régions détectées et applique une interpolation pour redimensionner chaque région tout en minimisant les distorsions.

Ce processus garantit que les lettres restent lisibles et que leurs caractéristiques visuelles sont préservées. Ce format uniformisé est ensuite utilisé pour alimenter des modèles de reconnaissance de caractères ou des systèmes OCR.

Le prétraitement des images inclut une conversion en niveaux de gris avant la binarisation. Cette conversion utilise une méthode simple mais efficace pour calculer une intensité lumineuse moyenne pour chaque pixel, basée sur ses composantes rouge, vert et bleu. En réduisant les variations inutiles de couleur, cette étape améliore la qualité des contours détectés lors de la binarisation. Les lettres et les lignes de la grille sont ainsi mieux distinguées du fond, ce qui a un impact direct sur la qualité des résultats des étapes suivantes.

La manipulation des images a été réalisée principalement à l'aide de SDL, une bibliothèque spécialisée dans le traitement des graphismes. Bien que SDL ne soit pas conçue spécifiquement pour le traitement d'images avancé, elle offre des outils flexibles pour manipuler les pixels d'images BMP. Cela a permis d'implémenter des fonctions personnalisées pour la binarisation, le coloriage, et le redimensionnement. Les fonctions SDL ont été intégrées dans un pipeline cohérent, où chaque étape est optimisée pour préserver la qualité des données extraites.

Enfin, des fonctionnalités supplémentaires ont été intégrées pour manipuler les grilles, comme l'inversion des colonnes ou la réorganisation des données. Ces outils permettent d'adapter les données extraites aux exigences spécifiques des systèmes cibles, garantissant ainsi une compatibilité maximale avec les frameworks OCR.

Cette combinaison d'avancées fonctionnelles et techniques a permis de construire un pipeline robuste et flexible pour la détection et le traitement des lettres dans des grilles. Chaque étape, de la binarisation au redimensionnement, a été soigneusement conçue pour maximiser la précision et la fiabilité, tout en permettant une intégration future dans des systèmes OCR avancés. Les ajustements réalisés tout au long du projet assurent une adaptabilité aux variations des images d'entrée et posent une base solide pour les prochaines étapes.

3 Baptiste, découpage et sauvegarde des lettres

3.1 Mon Parcours

Depuis mon plus jeune âge, l'informatique a exercé sur moi une fascination profonde et presque magnétique. L'idée qu'un simple ensemble de commandes puisse produire des résultats concrets et tangibles m'a toujours émerveillé. À mes yeux, un ordinateur n'a jamais été un simple appareil électronique, mais une porte ouverte sur un univers de possibilités infinies, où la créativité et la logique s'unissent pour donner vie à des solutions inattendues et ingénieuses.

Chaque interaction avec un ordinateur représentait une opportunité de plonger dans cet univers mystérieux. Je me souviens encore de mes premières explorations numériques, moments empreints d'un mélange d'émerveillement et de curiosité. Tout a commencé par des gestes simples et innocents : appuyer sur les touches du clavier pour observer leurs effets, changer les fonds d'écran ou personnaliser les paramètres du système d'exploitation. Bien que ces tâches puissent paraître anodines, elles m'ont appris à comprendre que l'ordinateur n'est pas seulement une machine, mais un outil dont la puissance réside dans la manière dont on l'utilise.

Mon premier ordinateur était un modeste appareil, bien loin des machines puissantes que l'on trouve aujourd'hui. Je me souviens encore du bruit distinctif du modem 56k se connectant à Internet, une sorte de symphonie technologique qui annonçait un monde inconnu. Explorer ce monde était un peu comme entrer dans une bibliothèque infinie où chaque clic ouvrait une nouvelle porte. À l'époque, mes centres d'intérêt étaient simples : jouer à des jeux éducatifs, explorer des encyclopédies numériques ou expérimenter avec des logiciels de dessin. Mais ces petites expériences ont eu un impact profond, jetant les bases d'une fascination qui n'a cessé de croître.

Une quête d'apprentissage et d'innovation Cette réflexion approfondie peut être illustrée par des exemples concrets et des anecdotes qui rendent chaque concept plus tangible et relatable. À travers des expériences variées, ces idées peuvent être reliées à des défis ou des opportunités réels, ce qui permet de mieux

comprendre leur portée dans différents contextes.

En avançant dans mon parcours, j'ai pris conscience que comprendre l'informatique ne se résume pas à savoir utiliser des logiciels ou coder des programmes simples. Il s'agit aussi d'adopter une manière de penser, une méthodologie pour résoudre des problèmes et repousser les limites du possible. Cette approche m'a motivé à explorer les multiples facettes de cette discipline en perpétuelle évolution.

Cette réflexion approfondie peut être illustrée par des exemples concrets et des anecdotes qui rendent chaque concept plus tangible et relatable. À travers des expériences variées, ces idées peuvent être reliées à des défis ou des opportunités réels, ce qui permet de mieux comprendre leur portée dans différents contextes.

L'un des aspects les plus fascinants de l'informatique réside dans sa capacité à combiner créativité et rigueur. Alors que certains domaines privilégient l'une ou l'autre, l'informatique exige un équilibre subtil entre les deux. Prenons, par exemple, le développement de logiciels : concevoir une interface utilisateur nécessite de la créativité pour garantir une expérience agréable, tandis que coder les fonctionnalités sous-jacentes exige une précision quasi mathématique. Cette dualité a toujours exercé sur moi une attraction particulière, car elle reflète un défi intellectuel constant.

Exploration des concepts techniques Cette réflexion approfondie peut être illustrée par des exemples concrets et des anecdotes qui rendent chaque concept plus tangible et relatable. À travers des expériences variées, ces idées peuvent être reliées à des défis ou des opportunités réels, ce qui permet de mieux comprendre leur portée dans différents contextes. Rapidement, mon intérêt s'est orienté vers des concepts plus techniques et abstraits. J'ai voulu comprendre pourquoi certains algorithmes sont plus efficaces que d'autres, comment fonctionnent les bases de données, ou encore de quelle manière l'intelligence artificielle (IA) change le monde qui nous entoure. Ces réflexions ont enrichi mon parcours et m'ont poussé à explorer des domaines plus spécialisés.

Rapidement, mon intérêt s'est orienté vers des concepts plus techniques et abstraits. J'ai voulu comprendre pourquoi certains algorithmes sont plus efficaces que d'autres, comment fonctionnent les bases de données, ou encore de quelle manière l'intelligence artificielle (IA) change le monde qui nous entoure. Ces réflexions ont enrichi mon parcours et m'ont poussé à explorer des domaines plus spécialisés.

Par exemple, comprendre les algorithmes m'a conduit à étudier des structures de données comme les arbres binaires et les tableaux associatifs. Ces concepts, bien qu'abstraits, ont des applications pratiques fascinantes. Un arbre binaire, par exemple, permet de rechercher rapidement des informations dans une base de données, tandis qu'un tableau associatif est essentiel pour optimiser des processus complexes, comme le routage des paquets sur Internet.

Mon intérêt pour l'informatique a véritablement pris une nouvelle dimension lors d'un stage effectué durant ma classe de troisième. Ce stage, réalisé au sein d'une startup spécialisée dans le deep learning, m'a permis de m'immerger dans un domaine d'une sophistication inouïe. Cette expérience a représenté bien plus qu'un simple stage d'observation : elle a marqué un tournant décisif dans mon parcours, élargissant mon horizon technique et intellectuel.

Pendant ce stage, j'ai été initié aux principes fondamentaux du deep learning, une branche fascinante de l'intelligence artificielle inspirée du fonctionnement du cerveau humain. J'ai découvert les réseaux neuronaux artificiels, ces structures complexes capables de traiter des quantités massives de données pour en extraire des patterns significatifs. Le simple fait de comprendre comment ces réseaux, organisés en couches successives, parviennent à analyser, interpréter et prédire des résultats a complètement transformé ma perception de l'informatique.

Un des moments marquants a été l'étude d'un modèle capable de détecter des objets dans des images. Par exemple, j'ai observé comment un réseau entraîné pouvait identifier une voiture ou un piéton sur une photo en une fraction de seconde. Cette capacité d'analyse instantanée m'a impressionné non seulement par sa complexité, mais aussi par les applications concrètes qu'elle rend possibles dans des domaines tels que la sécurité, la santé ou encore les transports.

Comprendre les mécanismes du deep learning m'a également amené à réfléchir à ses implications. Si cette technologie peut être utilisée pour diagnostiquer des maladies ou améliorer la sécurité routière, elle soulève aussi des questions éthiques majeures. Par exemple, comment garantir que les algorithmes ne reproduisent pas les biais présents dans les données? Cette réflexion a enrichi mon intérêt pour l'informatique en y ajoutant une dimension philosophique.

Une autre étape clé de mon parcours a été la réalisation d'un projet d'OCR (Optical Character Recognition). Ce projet, effectué dans le cadre d'un travail collaboratif, m'a permis de développer mes compétences techniques tout en affrontant des défis stimulants.

3.2 Le projet, et ses contraintes

Le choix du langage C pour ce projet a représenté une double opportunité : celle de comprendre les bases de la programmation à bas niveau tout en apprenant à gérer les contraintes qu'elle impose. Contrairement à des langages

modernes comme Python, le C offre une maîtrise fine des ressources matérielles, ce qui en fait un outil particulièrement adapté pour des projets nécessitant des optimisations poussées.

Pour moi, travailler avec le C signifiait plonger dans la gestion explicite de la mémoire, un aspect à la fois intimidant et fascinant. Par exemple, lorsque nous avons développé des fonctions pour lire et manipuler des fichiers bitmap (.bmp), j'ai dû comprendre comment les données sont stockées dans la mémoire. Cette expérience m'a offert un regard approfondi sur l'efficacité des algorithmes et sur l'importance des choix techniques.

Dans le cadre de ce projet OCR, nous avons exploité plusieurs bibliothèques essentielles. Parmi elles, SDL2 (Simple DirectMedia Layer) a joué un rôle central dans la manipulation des images. Cette bibliothèque, bien connue pour sa polyvalence et sa robustesse, nous a permis de simplifier de nombreuses tâches complexes liées au traitement des fichiers bitmap. Par exemple, grâce à SDL2, nous avons pu charger, redimensionner et modifier des images sans avoir à gérer manuellement chaque pixel. Cela nous a offert une abstraction précieuse, permettant de concentrer nos efforts sur la logique fondamentale de l'OCR.

Nous avons également utilisé d'autres bibliothèques standard du langage C, telles que `math.h`, pour effectuer des calculs complexes. Une partie spécifique du projet consistait à redresser des caractères inclinés dans une image scannée. Cette tâche a nécessité l'application de fonctions trigonométriques afin de corriger les distorsions. Ce défi, bien que technique, m'a permis de comprendre l'interconnexion entre la théorie mathématique et son application pratique.

Le projet ne s'est pas déroulé sans embûches. L'un des principaux défis a été la gestion explicite de la mémoire. Contrairement à des langages de haut niveau comme Python ou Java, où la mémoire est gérée automatiquement, le langage C exige de l'utilisateur qu'il alloue et libère manuellement les ressources. Cette contrainte, bien qu'intimidante au départ, s'est avérée être une opportunité d'apprentissage inestimable. Par exemple, lors de la manipulation de grandes images, une mauvaise gestion de la mémoire entraînait des fuites, ce qui provoquait des ralentissements et des crashes. Grâce à une recherche minutieuse et

à l'étude de la documentation, j'ai pu comprendre et implémenter des solutions efficaces, comme l'utilisation appropriée des pointeurs et des structures de données dynamiques.

Un autre défi majeur a été l'interprétation correcte des données bitmap. Ces fichiers contiennent des informations codées en binaire, ce qui les rend difficiles à déchiffrer sans outils adaptés. J'ai appris à décoder manuellement ces fichiers, ligne par ligne, pour extraire les informations nécessaires, comme la couleur des pixels ou leur position. Ce processus m'a offert une vue rare et fascinante sur la manière dont les ordinateurs traitent les images.

Internet a été une ressource clé tout au long de ce projet. Dans le domaine de l'informatique, la capacité à rechercher des informations précises est une compétence cruciale. Que ce soit pour résoudre des bugs ou comprendre des concepts complexes, les forums, tutoriels et documentations en ligne sont des outils indispensables.

L'un des sites les plus utiles a été Stack Overflow, une plateforme où des milliers de développeurs partagent leurs connaissances. Par exemple, lors de l'intégration de SDL2, nous avons rencontré un problème lié à la configuration de l'environnement de développement. Une réponse détaillée sur ce forum nous a permis de surmonter cet obstacle en quelques heures, ce qui aurait pris des jours autrement. Ce type de collaboration numérique reflète la beauté de la communauté informatique, où les connaissances sont partagées librement pour le bénéfice de tous.

Au-delà des forums, les documentations officielles des bibliothèques, comme celle de SDL2, se sont révélées précieuses. Elles m'ont appris à lire et interpréter des spécifications techniques, une compétence que je considère désormais essentielle pour tout développeur.

3.3 Les bénéfices moraux du projet

Mener un projet ambitieux tel que cet OCR tout en jonglant avec les responsabilités scolaires a été un défi d'équilibre. Entre les cours, les devoirs et d'autres engagements académiques, trouver du temps pour coder et expérimenter demandait une organisation rigoureuse. J'ai appris à prioriser mes tâches, en utilisant des outils tels que des plannings hebdomadaires et des listes de priorités. Cette discipline organisationnelle m'a non seulement aidé à respecter les délais du projet, mais a également renforcé mes compétences en gestion du temps, qui sont précieuses dans tous les aspects de la vie.

Les moments de frustration étaient inévitables, surtout lorsque des bugs persistaient malgré des heures de débogage. Cependant, j'ai découvert que ces défis étaient des opportunités d'apprentissage déguisées. Par exemple, après avoir passé une journée entière à résoudre un problème lié à la lecture des fichiers bitmap, j'ai ressenti une immense satisfaction lorsque la solution a enfin fonctionné.

L'un des aspects les plus gratifiants de ce projet a été la collaboration avec mes camarades. Travailler en équipe m'a permis de découvrir les forces de chacun et d'apprendre des approches différentes. Par exemple, un membre de notre équipe était particulièrement talentueux pour optimiser les algorithmes, tandis qu'un autre avait une excellente compréhension des bases mathématiques nécessaires pour redresser les images.

Collaborer efficacement exige une communication claire et régulière. Nous avons utilisé des outils comme GitHub pour gérer notre code et Slack pour nos discussions. Ces plateformes ont facilité le partage des idées et la résolution rapide des problèmes. De plus, elles m'ont initié à des pratiques professionnelles, telles que la gestion des versions et le travail en mode asynchrone.

Comme tout projet technique, celui-ci n'a pas été exempt de difficultés. Certaines étapes, comme l'implémentation des algorithmes de reconnaissance, ont nécessité une créativité et une ingéniosité considérables. Nous avons parfois

dû combiner des solutions issues de différentes sources ou inventer nos propres méthodes pour surmonter les obstacles.

Chaque échec rencontré lors de ce projet a été une leçon déguisée. Par exemple, une tentative de compresser les données pour accélérer le traitement des images a initialement conduit à des erreurs imprévues. Cependant, en analysant les causes du problème, nous avons découvert une méthode encore plus efficace pour atteindre notre objectif.

En rétrospective, ce projet et les expériences qui l'ont précédé ont marqué mon parcours d'une manière indélébile. Ils m'ont permis de développer des compétences techniques approfondies, de renforcer ma capacité à travailler en équipe et de comprendre les exigences des projets ambitieux.

Plus important encore, ces expériences m'ont appris à rêver grand et à voir les défis comme des tremplins vers la croissance personnelle et professionnelle. Je suis convaincu que ces apprentissages me seront précieux tout au long de ma carrière d'ingénieur, et je suis impatient de continuer à explorer, innover et contribuer au monde fascinant de l'informatique.

4 Otto, réseau de neurones

4.1 Présentation et Motivation

Depuis mon plus jeune âge, l'intelligence artificielle m'a toujours captivé, non seulement par son caractère innovant, mais aussi par ses capacités extraordinaires à apprendre, à évoluer et à s'adapter en analysant d'importants volumes de données. Cette fascination initiale s'est transformée en un véritable engouement au fil du temps, à mesure que je découvrais l'étendue des possibilités offertes par cette discipline, qui mêle science et technologie pour résoudre des problèmes complexes.

L'IA, par son potentiel à simuler certains aspects de la cognition humaine, m'a constamment poussé à approfondir mes connaissances. Ainsi, lorsque j'ai eu l'opportunité de m'impliquer dans ce projet, mon intérêt s'est immédiatement manifesté. Ce dernier combine des techniques avancées de reconnaissance de lettres, un domaine qui repose sur des principes fondamentaux d'apprentissage automatique, un sujet qui me passionne profondément.

Ce projet m'a offert une occasion unique de m'immerger dans des problématiques concrètes liées à l'IA, en mettant en pratique des concepts fondamentaux tout en explorant des approches innovantes. Il s'est agi pour moi d'un exercice particulièrement enrichissant, car il m'a permis de mieux comprendre les subtilités des mécanismes qui permettent à une machine de reconnaître, traiter et comprendre des caractères écrits. La complexité de cette tâche réside à la fois dans la logique rigoureuse nécessaire à sa mise en œuvre et dans la créativité indispensable pour concevoir des solutions adaptées. Ce défi, en croisant réflexion méthodique et innovation technique, a renforcé ma motivation à approfondir mes compétences dans ce domaine.

Dans ce cadre, j'ai eu la chance d'être impliqué activement dans toutes les étapes du développement, en particulier la conception et la mise en œuvre d'un réseau de neurones spécialement adapté aux besoins du projet. Cette démarche a exigé une analyse approfondie et méthodique, englobant des aspects techniques variés : l'architecture du réseau, le choix des algorithmes d'apprentissage, et l'optimisation des hyperparamètres. Chacune de ces étapes a joué un rôle cru-

cial dans la garantie de résultats précis et exploitables. Cette expérience a consolidé mes compétences techniques en modélisation, entraînement de modèles et analyse des performances, tout en me permettant de mieux appréhender les défis propres à l'intelligence artificielle. De plus, elle a encouragé une réflexion structurée et novatrice face à des problématiques complexes, alimentant ainsi ma passion pour un domaine en constante évolution.

4.2 État d'avancement

Le projet que nous avons entrepris visait à concevoir une intelligence artificielle capable d'identifier une lettre donnée à partir d'une image en entrée. Ce défi technique, aussi complexe qu'ambitieux, a nécessité des ajustements importants par rapport aux modèles théoriques initialement étudiés. Cela a mis en évidence la complexité de la tâche et souligné l'opportunité de repousser les limites des réseaux neuronaux.

Nous avons bâti ce projet en nous appuyant sur les bases solides posées lors de notre première phase de travail, tout en introduisant des modifications significatives pour répondre aux exigences spécifiques. Parmi ces ajustements, une refonte complète de l'architecture du réseau a été réalisée pour l'adapter au mieux à la reconnaissance des caractères. Cette adaptation s'est accompagnée d'un travail minutieux d'optimisation des paramètres et de gestion des données complexes, contribuant à une efficacité accrue.

Parallèlement, un effort particulier a été consacré à l'optimisation du code, réduisant les temps de calcul de manière notable et rendant le modèle plus performant dans des conditions réelles. Une fonctionnalité clé du projet réside dans la mise en œuvre d'un système avancé de sauvegarde. Ce dernier permet de conserver l'état interne du réseau, notamment les poids et biais appris, sous forme de fichiers externes. Cela garantit non seulement une continuité dans le développement et l'entraînement, mais aussi une grande flexibilité pour reprendre le travail ou déployer le modèle sans perte d'informations cruciales. En outre, cette fonctionnalité permet d'utiliser un réseau déjà entraîné pour traiter une seule image et fournir instantanément la lettre correspondante, une avancée qui simplifie les processus de validation et d'expérimentation.

Cependant, malgré les avancées techniques réalisées, certaines limitations ont été identifiées. Le temps disponible pour entraîner le modèle s'est révélé insuffisant pour atteindre des performances optimales. Bien que le réseau soit fonctionnel, les résultats obtenus – avec un taux de réussite actuel d'environ 33 % – demeurent en deçà des objectifs fixés.

Ces performances soulignent la nécessité d'un entraînement prolongé, avec davantage d'itérations et une exploration plus approfondie des hyperparamètres pour améliorer les prédictions et réduire les erreurs. Ce constat met en évidence l'importance cruciale d'investir du temps et des ressources supplémentaires pour renforcer la robustesse du modèle.

En somme, bien que le projet repose sur des bases solides et des choix techniques judicieux, les résultats actuels indiquent qu'un travail supplémentaire est nécessaire pour atteindre le niveau de performance souhaité. Ces efforts incluent un entraînement prolongé, une optimisation fine et l'exploration de nouveaux paramètres pour maximiser le potentiel de notre intelligence artificielle.

4.3 Aspects techniques

Le développement de cette intelligence artificielle repose sur une conception rigoureuse et méthodique, mettant en avant une architecture de réseau de neurones multicouches soigneusement structurée. La définition précise du réseau,

basée sur une liste configurable précisant le nombre de neurones dans chaque couche, permet une flexibilité exceptionnelle dans l'adaptation de l'architecture aux besoins spécifiques de la reconnaissance de caractères. Cette modularité a été déterminante pour ajuster le modèle et maximiser son efficacité dans le cadre de cette tâche complexe. En outre, la gestion des biais à travers une matrice dédiée et des poids via listes de matrices garantit la fiabilité des calculs et des interactions entre les couches.

L'entraînement du réseau s'appuie sur la méthode éprouvée de la descente de gradient stochastique, qui s'est avérée particulièrement adaptée aux contraintes et objectifs de notre projet. Nous avons opté pour une approche basée sur l'utilisation de mini-lots, divisant les données en sous-ensembles pour équilibrer efficacité et précision. Cette stratégie permet d'accélérer les calculs tout en assurant une convergence fiable des résultats. Grâce à cette méthodologie, nous avons pu gérer efficacement les ressources disponibles et maintenir des temps de calcul acceptables, même en présence de données volumineuses.

La fonction d'activation choisie pour ce projet est la ReLU (Rectified Linear Unit), reconnue pour ses performances supérieures dans les réseaux neuronaux modernes. Contrairement à la fonction sigmoïde, initialement testée, ReLU s'est montrée mieux adaptée à nos besoins, notamment grâce à sa capacité à atténuer le problème du gradient évanescent et à accélérer l'entraînement. Cette fonction, qui retourne le maximum entre l'entrée et zéro, a permis une convergence rapide et stable, offrant des performances globales améliorées. Ce choix illustre l'importance d'adapter les paramètres et les méthodologies en fonction des résultats observés pour maximiser l'efficacité du modèle.

En conclusion, les aspects techniques de ce projet reposent sur des choix méthodologiques réfléchis et des implémentations soigneusement calibrées. La flexibilité architecturale, l'utilisation de la descente de gradient avec mini-lots, et l'adoption de ReLU comme fonction d'activation ont permis de construire un socle solide pour le développement de l'intelligence artificielle. Cependant, les performances actuelles démontrent que des améliorations sont encore nécessaires, notamment à travers un entraînement plus poussé, une exploration approfondie des hyperparamètres et une optimisation continue, afin de pleinement exploiter le potentiel du modèle et répondre aux exigences de la reconnaissance de caractères.

5 Alexis, traitement de l'image

5.1 Présentation et Motivation

J'ai commencé réellement à m'intéresser à l'informatique un peu avant ma seconde à travers la programmation. J'ai tout de suite accroché l'aspect challengeant poussant toujours à réfléchir un problème et sa résolution par soi même avec une possibilité et une variation énorme des réponses. J'ai surtout été ébahi devant toutes ces lignes de mots qui donnaient un résultat qui n'avait rien à voir avec du texte, ces lignes qui peuvent interagir directement sur ce qu'affichait l'ordinateur et sur quoi et comment je manipule ce qui est à l'écran. Après cette découverte, à mon entrée en seconde, j'étais sûr d'une chose : j'allais prendre la spécialité Numérique et Sciences de l'Informatique jusqu'en terminale et j'allais en faire mes études post-bac. Cette envie s'est renforcée lorsqu'en seconde j'ai assisté au cours de SNT qui ne sont ni plus ni moins une initiation à l'informatique et à l'immensité des domaines auxquels elle a de l'impact. En voyant tout c'est champs d'action dans autant de domaine me faisait dire que si j'aimais l'informatique sans pour autant savoir quoi précisément j'avais le temps car il y a tellement à découvrir. D'autant plus que ce domaine est en constante évolution depuis maintenant plusieurs années ne serait-ce que sur l'intelligence artificielle. Voilà ce qui m'a attiré en informatique, à la fois la variété des domaines d'application mais aussi les possibilités et les résultats impressionnants de complexité avec des outils paraissant facile à manipuler.

Aujourd'hui, je suis toujours dans une mentalité de découverte et de curiosité afin d'être sûr de ne pas me fermer des portes. C'est pourquoi je n'ai pas de domaine que je repousse automatiquement. J'essaye de toucher un peu à tout pour me faire un avis sur tout. Dans ce projet je m'attaque à quelque chose que je n'avais jamais fait avant ça me permet de continuer dans ma dynamique de tout faire avant de choisir pour ne passer à côté de mon potentiel futur domaine préféré. La partie technique n'est pas le seul aspect du projet sur lequel j'ai besoin de travailler mais aussi le travail en groupe. Ce travail permet de développer chez moi des qualités essentielles que ce soit pour un ingénieur ou pour moi en tant qu'humain. Par exemple, Otto, Baptiste et Lorys sont des personnes avec qui j'avais assez peu ou pas de relations avant ce projet et me mettre avec eux m'a permis de mieux les connaître, en bien évidemment, ce qui a augmenté ma confiance en moi et en mes relations avec les autres. Cet exemple n'est qu'un parmi une multitude d'améliorations humaines et techniques qui font de moi, je pense, un meilleur futur ingénieur en informatique, un meilleur collègue de travail de groupe et une meilleure personne globalement.

5.2 État d'avancement

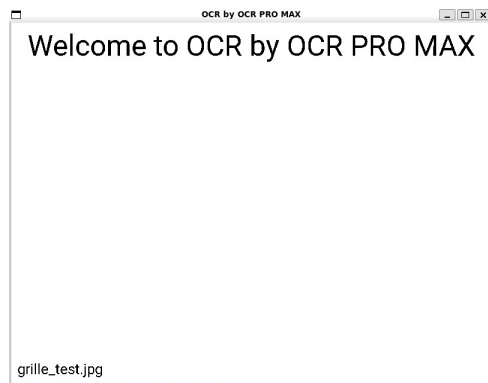
De mon côté, je m'occupe de la partie graphique du projet, c'est à dire tout ce qui touche à l'interface et donc aux interactions que l'utilisateur aura avec notre application. Pour cette première soutenance, je devais donc produire le code pour avoir une fenêtre de laquelle l'utilisateur peut charger son image depuis le dossier duquel l'utilisateur travail. Au chargement de cette dernière, un filtre noir et blanc sera appliqué pour faciliter le travail de reconnaissance de notre algorithme de résolution car cela augmentera le contraste de l'image ce qui permettra de mieux identifier les lettres ou la grille dans une image initialement de mauvaise qualité. Enfin l'utilisateur doit pouvoir faire tourner son image afin que cette dernière soit dans le sens adéquat. Cette dernière étape s'inscrit toujours dans cette dynamique de simplifier à la fois le travail de notre algorithme et l'algorithme en lui-même puisqu'il n'a besoin de reconnaître les lettres que dans un seul sens, le bon.

Tout d'abord en voyant les tâches qui m'ont été attribuées je pensais que cela allait être rapide car nous avons un tp sur comment manipuler des images et une fenêtre et que je m'en étais pas mal sorti. Cependant, à cause de cela je m'y suis pris un peu tard et j'ai pu faire à peine la moitié de ce que j'avais à faire en terme de tâche mais beaucoup moins en terme de temps de travail. En effet les premières étapes de création de fenêtre ainsi que de chargement d'image ont été les plus simples. J'ai modulé l'interface afin de rendre l'expérience utilisateur la plus agréable possible sans pour autant y passer trop de temps dessus en ajoutant un titre en rendant la taille de l'image adaptative aux changements de taille de fenêtre. Cependant les premières difficultés sont arrivées lors de l'implémentation du filtre noir et blanc. En effet, même si j'avais l'algorithme et le code parcourant chaque pixel de l'image afin d'assombrir ou clarifier le pixel afin d'augmenter le contraste. En effet au moment du chargement de l'image, au bout d'un petit moment, le programme s'arrêtait sur une "Segmentation Error". Cette erreur est en général produite lorsqu'on manipule un pointeur dans un espace mémoire qui ne lui est pas alloué. J'ai passé au moins un jour entier sur cette erreur passant en revue chaque ligne en faisant à chaque fois des petites modifications, sans trouver pour autant l'origine de cette erreur. J'ai donc pris l'initiative de laisser ma fonction commentée et de rester à la version où l'utilisateur peut simplement charger une image.

5.3 Aspects techniques

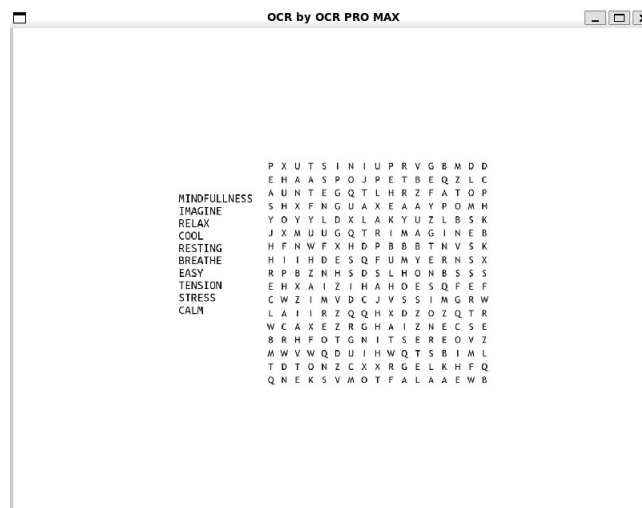
Dans cette partie je vais parler de choix que j'ai fait dans mon travail, dans le dossier Interface, afin d'avoir une idée à la fois précise et concise de mon cheminement de pensée ainsi que de son fonctionnement. Le dossier est composé d'un fichier main où tout les autres fichiers seront appelés pour exécuter entièrement le code, cela permet si nous avons besoin plus tard de créer de nouveaux fichiers dans ce dossier d'avoir un fichier qui résume l'exécution pour y voir plus clair. Il y a aussi un dossier Roboto contenant la police utilisée fournie gratuitement par Google j'ai choisi cette police car dans mes recherches c'est celle qui me semblait être la plus neutre. Enfin le plus intéressant est le fichier interface.c car c'est lui qui contient toute mon avancée. Je détaille son contenu plus bas.

Tout d'abord, j'importe différentes bibliothèques, certaines sont quasi indispensables pour coder en C telles que `stdio.h` et `string.h`, cependant d'autres sont moins courantes et pourtant indispensables dans la conception du projet, ce sont les différentes bibliothèques SDL. La première, `SDL.h` permet d'accéder aux fonctions de base de manipulation de fenêtre avec SDL elle permet par exemple d'ouvrir et de fermer une fenêtre. La deuxième, `SDL_image.h` comme son nom l'indique donne accès à des fonctions de traitement d'image. C'est grâce à cette bibliothèque que l'utilisateur peut charger son image et faire varier sa taille. Enfin `SDL_ttf.h` est liée aux options d'écriture au sein même de la fenêtre. Plus précisément c'est avec ses fonctions qu'on va pouvoir manipuler des polices d'écritures afin d'écrire dans la fenêtre.



- Fenêtre au lancement avec entrée du nom de l'image

Après ça le fichier est divisé en trois parties. La première composée essentiellement d'initialisation de pointeurs de texture, de fenêtre ou de police. La deuxième, la plus importante contient le code exécutable tant que la fenêtre est ouverte. C'est à dire qu'il y a la partie qui gère l'écriture du chemin d'accès à l'image, celle qui affiche l'image dans la fenêtre tout en conservant les proportions des éléments de la fenêtre en fonction de la taille de cette dernière. La dernière est finalement la conclusion du code puisqu'elle supprime tous les éléments qui étaient présent lors de la fermeture de la fenêtre.



- Grille chargée

6 Tableau d'avancement

Après la présentation globale de notre projet nous avons pu former le tableau d'avancement ci-dessous:

	Prévu pour le 08/12	Réalisé
Interface graphique et image	100%	25%
Découpage grille	100%	100%
Détection des éléments	100%	85%
Réseau de neurones	100%	60%

Table 1: Avancement des tâches

7 Conclusion

Ce projet a été pour nous une véritable aventure au cours de laquelle nous avons dû faire face à de nombreux défis, parfois complexes et inattendus. Tout au long de cette expérience, nous avons réussi à avancer sur plusieurs aspects clés, ce qui nous a permis d'explorer des domaines que nous ne maîtrisions pas auparavant. Ces découvertes ont enrichi notre expertise et nous ont permis de poser des bases solides pour la suite du projet. Cependant, certains points restent encore en cours d'analyse et nécessitent des ajustements pour que nous puissions progresser davantage et parvenir à un résultat final conforme au cahier des charges initialement établi.

Bien que nous ayons accompli des avancées significatives, nous devons reconnaître que la situation actuelle est encore en dessous attentes. Cette situation s'explique non seulement par un manque de temps, mais également par des lacunes au niveau de notre organisation, tant sur le plan individuel que collectif. Ces difficultés organisationnelles ont principalement résulté d'un manque de communication efficace au sein du groupe, ce qui a parfois ralenti nos efforts et créé des malentendus. Nous avons pris conscience de ces faiblesses et travaillons activement à les corriger pour améliorer notre efficacité à l'avenir.

Malgré ces obstacles, cette expérience a été extrêmement formatrice. Nous avons non seulement acquis des compétences techniques précieuses en programmation, mais aussi développé des aptitudes essentielles en gestion de groupe et en collaboration.