



# RAPPORT DE STAGE

Stage d'analyse de la réalisation d'une mission professionnelle

Du 6 janvier au 31 janvier 2025

---

**Réentraînement du modèle YOLOv8 pour la  
détection des troncs d'olivier**

---

**PELLENC**

Stagiaire en IA

Imane MOUMOUN

Tuteur entreprise	:	Thierry STEIN
Encadrants académiques	:	Jean-Philippe TOPPIA Thibaut METAILLER

# Remerciements

---

Tout d'abord, je souhaite exprimer ma sincère gratitude à M. Jean-Philippe TOPPIA et M. Thibaut METAILLER, mes encadrants académiques, pour leurs conseils précieux qui ont grandement contribué à l'élaboration de ce rapport de stage.

Je tiens également à remercier chaleureusement M. Simon BARBEAU, PDG de PELLENC, pour m'avoir offert l'opportunité de réaliser ce stage au sein de l'entreprise.

Je suis profondément reconnaissante à M. Thierry STEIN, mon tuteur de stage, pour m'avoir accueillie au sein de son équipe et pour m'avoir permis de développer des compétences pratiques dans un environnement dynamique.

Je remercie également M. Haitam AOUAD IDRISSE, ingénieur en IA et systèmes embarqués, pour avoir partagé son expertise et pour m'avoir accompagné tout au long de ce stage, contribuant ainsi à ma progression tant sur le plan technique que professionnel.

Enfin, je souhaite exprimer ma reconnaissance à l'ensemble des ingénieurs du bureau d'études pour leur accueil chaleureux, leur positivité et leur bienveillance, qui m'ont permis de me sentir pleinement intégrée et ont grandement enrichi mon expérience.

# Sommaire

---

Introduction	1
I. Présentation de l'entreprise d'accueil	2
1. Présentation générale du groupe PELLENC	2
1.1. Historique	2
1.2. Gamme de produits	3
1.2.1. Division « AGRI »	4
1.2.2. Division « GREEN & CITY TECHNOLOGY »	4
1.3. Les innovations iconiques de PELLENC	5
1.4. Chiffres clés	6
2. Organisation du siège social à Pertuis	8
2.1. Structure de l'entreprise d'accueil	8
2.2. Mode managérial	10
2.3. Aménagement des locaux	10
2.4. Communication au sein de l'entreprise	11
3. L'équipe R&T et ses spécificités	12
II. Missions et réalisations	15
1. Fondements théoriques	15
1.1. IA, apprentissage automatique et vision par ordinateur	15
1.2. Processus d'apprentissage automatique	15
1.3. Apprentissage profond	16
1.4. Fonctionnement d'un CNN	18
1.5. Le modèle YOLO	19
2. Présentation du projet de stage	21
2.1. Besoins et objectifs	21
2.2. Déroulement du stage	22
3. Mes missions	23
3.1. Prise en main de l'environnement de développement	23
3.2. Collecte des données	24

3.3. Construction du dataset annoté	25
3.4. Filtrage du dataset	28
3.5. Réentraînement du modèle YOLOv8	28
3.6. Évaluation du modèle	29
3.6.1. Métriques d'évaluation	29
3.6.2. Résultats et discussion	31
4. Retour d'expérience	33
III. L'IA et l'agriculture: une révolution entre opportunités et défis	36
1. Contexte et problématique	36
2. L'IA au service de l'agriculture	36
2.1. Évolutions technologiques du domaine agricole	36
2.2. Opportunité de l'IA dans l'agriculture	37
2.2.1. Le rôle de l'IA dans l'optimisation de la production agricole	37
2.2.2. L'agriculture de précision et la gestion durable des ressources grâce à l'IA	38
2.2.3. Automatisation et robotisation des tâches agricoles grâce à l'IA	39
3. Les limites et défis de l'IA dans l'agriculture	41
3.1. Une fracture numérique au détriment des petits exploitants	41
3.2. Dépendance aux données	42
3.3. Des risques pour l'emploi et les savoir-faire traditionnels	43
4. Vision future de l'Agriculture Intelligente	44
4.1. Vers une collaboration homme-machine pour une agriculture plus durable	44
4.2. Encadrer l'IA pour une agriculture éthique, équitable et régulée	45
4.3. Scénarios pour l'agriculture de demain	46
Conclusion	48
Glossaire	49
Bibliographie	51
Annexes	53

# Introduction

---

Au cours de ma première année du cursus ISMIN à l'école des Mines de Saint-Étienne, j'ai eu l'occasion d'effectuer un stage d'analyse d'entreprise d'une durée de 4 semaines au sein de l'entreprise PELLENC située à Pertuis dans le Vaucluse.

PELLENC est une entreprise spécialisée dans le domaine des technologies agricoles. Fondée en 1973, elle s'est rapidement imposée comme un acteur majeur sur le marché grâce à son engagement constant dans la recherche et le développement de solutions agricoles innovantes. Son siège social, basé à Pertuis, abrite également ses principaux centres de production et de recherche.

J'ai eu l'opportunité de décrocher ce stage grâce au forum entreprise organisé par l'association FEI au sein de mon école. C'est à cette occasion que j'ai rencontré M. Thierry Stein, responsable du service Recherche & Technologie (R&T) chez PELLENC. Lors de notre échange, j'ai pu lui faire part de mon intérêt marqué pour l'intelligence artificielle (IA) et de mon souhait de rejoindre son équipe. Après l'étude de ma candidature, j'ai reçu une réponse favorable début décembre, suivie d'un entretien avec l'équipe que j'allais intégrer. Cet échange m'a permis de mieux appréhender les objectifs du stage, en particulier leur besoin spécifique d'un modèle de réseau de neurones pour la détection de troncs d'oliviers. Donc si j'ai choisi ce stage c'est pour confirmer mon projet professionnel, qui est de devenir ingénieure spécialisée en intelligence artificielle embarquée.

Le présent rapport retranscrit à la fois mon expérience complète au sein du bureau d'études de PELLENC, ainsi que les réflexions que j'ai pu avoir pendant mon stage. Dans un premier temps, je présenterai le fonctionnement ainsi que le contexte dans lequel l'entreprise PELLENC évolue. Ensuite, j'explicitai les missions que j'ai réalisées ainsi que les résultats obtenus. Enfin, j'analyserai le rôle de l'IA dans l'agriculture, en explorant les opportunités qu'elle offre, les défis qu'elle soulève et les perspectives qu'elle dessine pour l'avenir de l'agriculture.

# I. Présentation de l'entreprise d'accueil

---

## 1. Présentation générale du groupe PELLENC

### 1.1. Historique

PELLENC est un groupe français d'envergure internationale qui est l'un des leaders mondiaux dans les domaines de la viticulture, l'oléiculture et l'arboriculture. Fondée en 1973 par Roger PELLENC, l'entreprise a établi son siège social à Pertuis dans le Vaucluse. L'histoire de la création et du développement de l'entreprise peut se résumer comme suit :

- 1973** : Démarrage de l'activité par Roger PELLENC seul, dans un hangar agricole. L'histoire de PELLENC commence dans les années 70, lorsque Roger PELLENC, ancien professeur au lycée technique d'Avignon met au point dans son garage des machines pour la viticulture avec son ami, Jean Paul Motte.
- 1974** : Création de la SARL (Société A Responsabilité Limitée) par PELLENC et MOTTE. Lancement du premier outil pour la viticulture : l'écimeuse
- 1983** : Passage de la société du statut de SARL à celui de S.A. (Société Anonyme). Jusqu'en 1983, l'activité prospère et de nombreuses machines viennent compléter la gamme proposée par la société.
- 1989** : Augmentation du capital qui passe à 764 684 €. La société est alors organisée en quatre départements produits, deux services fonctionnels, un bureau d'études et un service commercial.
- 1991** : Premier élargissement à l'international (Espagne et Australie).
- 1993** : Lancement de la première Machine à Vendanger.
- 1998** : La société PELLENC compte 130 employés. Augmentation du capital à 3 884 596 €.
- 2000** : Création de PELLENC Slovakia (PSRO), première filiale de production.
- 2004** : Création de la première batterie au lithium-ion.
- 2008** : Face à l'essor de la viticulture dans les pays émergents, le développement international du groupe s'est accentué par la création de filiale.
- 2013** : Lancement du produit vedette : La MAV Optimum.
- 2014** : Le groupe SOMFY rachète la totalité des parts de PELLENC.
- 2017** : Départ de Roger PELLENC comme directeur de l'entreprise.

- 2018** : PELLENC SA devient PELLENC S.A.S (Société par Action Simplifiée).
- 2019** : Lancement de la deuxième version de l'Optimum.
- 2022** : PELLENC a investi quasiment 10 M€ pour renforcer ses équipements industriels et compte désormais huit usines en Europe.
- 2023** : PELLENC se distingue et se voit attribuer trois médailles de bronze lors de la 31e édition du SITEVI (salon international des équipements et savoir-faire pour les productions vigne-vin, olive, fruits-légumes)

Aujourd'hui, le groupe PELLENC possède 21 filiales réparties à travers le monde ( figure 1) : la plupart d'entre elles se concentrent autour de la France en Europe occidentale ; mais aussi dans des pays comme la Chine, l'Australie, les Etats Unis, l'Afrique du Sud, etc. 9 sont des centres de production, tandis que les autres sont plutôt dédiées aux aspects commerciaux et SAV (Service Après-Vente).



Figure 1 : Filiales PELLENC dans le monde

## 1.2. Gamme de produits

Il existe deux divisions principales (ou Business Unit) dans cette société. La division « AGRI » qui regroupe les machines liées à la viticulture, à la viniculture et à l'oléiculture et la division « Green and City Technology » qui regroupe les outils électroportatifs principalement pour l'arboriculture et l'entretien des espaces verts.

### 1.2.1. Division « AGRI »

La division « AGRI » est spécialisée dans la conception et la fabrication de machines agricoles destinées aux secteurs de la viticulture, de la viniculture et de l'oléiculture. Pour la viticulture (culture de la vigne), PELLENC propose du matériel de taille, d'écimage, de vendange, de travail du sol et de pulvérisation. En viniculture (élaboration du vin), la gamme inclut des équipements tels que les machines à vendanger, les trieuses optiques de raisins et les égreneurs de grappes. Pour l'oléiculture (culture et exploitation des oliviers pour produire des olives de table ou de l'huile d'olive), l'entreprise offre des machines spécifiques comme les secoueurs équipés de bras vibrants pour la récolte. La figure 2 montre trois exemples de machines agricoles de la marque PELLENC : Une machine à vendanger ( image à gauche ), une tailleuse de vigne (image du centre) et une machine de récolte des olives ( image de droite).



Figure 2 : Produits de gamme AGRI

### 1.2.2. Division « GREEN & CITY TECHNOLOGY »

La division GCT « Green & City Technology » de PELLENC est spécialisée dans le développement et la commercialisation d'outils électroportatifs professionnels destinés à l'entretien des espaces verts et urbains. Ces outils ont la particularité de fonctionner grâce à des batteries « sac-à-dos » au lithium-ion, offrant ainsi des solutions écologiques, performantes et adaptées à divers besoins tels que la tonte, le débroussaillage, le désherbage, la taille, l'élagage et le nettoyage.



La figure 3 montre trois exemples de ces outils : une tronçonneuse électrique alimentée par une batterie dorsale (image à gauche) , un sécateur électrique (image au centre) et un élagueur sur perche alimenté par une batterie dorsale (image à droite)

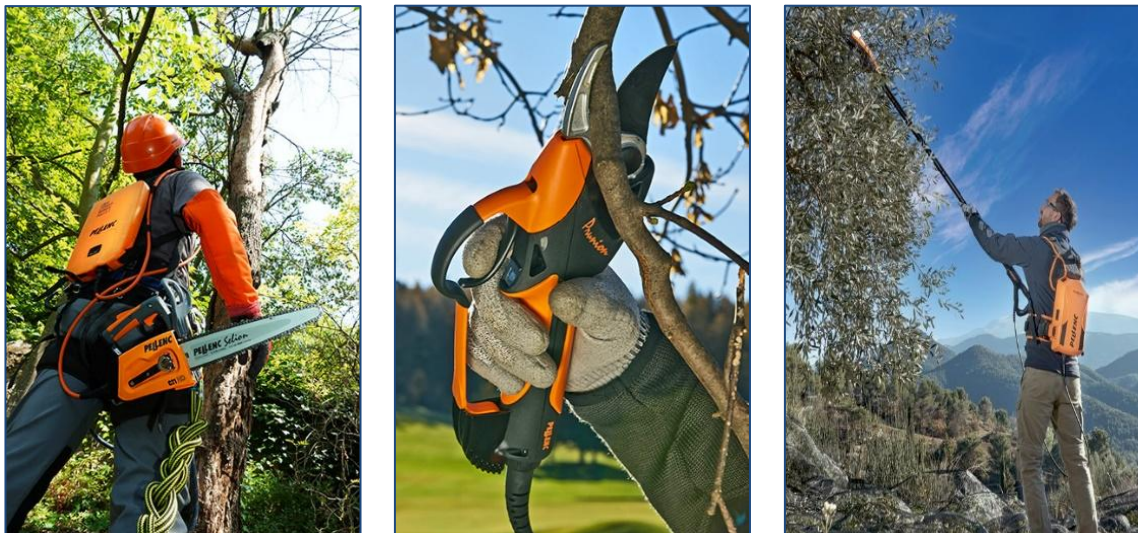


Figure 3 : Produits de gamme GCT

### 1.3. Les innovations iconiques de PELLENCE

Depuis plus de 50 ans, PELLENCE s'affirme comme un véritable moteur d'innovations dans les secteurs de la viticulture, de l'arboriculture et de l'entretien des espaces verts. L'entreprise a marqué l'histoire de la technologie agricole avec des avancées emblématiques, comme le sécateur hydraulique (1976) et électrique (1987), révolutionnant les travaux de taille, ou encore les machines à vendanger (1993), qui ont facilité la récolte mécanique des raisins. En 1983, elle invente la première solution de mécanisation pour la prétaille des vignes palissées, avant de lancer FIXION en 1995, le premier outil dédié à l'attachage des sarments. Visionnaire, PELLENCE a également été pionnière en 2004 avec l'intégration des solutions lithium-ion sur un outil électroportatif pour l'agriculture. Le concept multifonction, introduit dans les années 1990, et l'acquisition de PERA en 2014, ont permis de couvrir l'ensemble des besoins de la vigne à la cave, tandis que la plateforme numérique PELLENCE Connect optimise la gestion et la traçabilité des opérations. Toujours à la pointe, l'entreprise a présenté en 2023 le RX-20, un véhicule autonome conçu pour réaliser des tâches répétitives comme le désherbage mécanique.

D'ailleurs, la reconnaissance internationale de ces innovations est illustrée par les dix médailles d'or décernées par le SITEVI entre 1983 et 2009, ainsi que les trois médailles de bronze que l'entreprise

a reçu lors de l'édition 2023 des SITEVI Innovation Awards pour son convoyeur pneumatique AIR 3, également pour le RX-20 et enfin pour le SECURION. Par son esprit novateur et son expertise technologique, PELLENC s'impose aujourd'hui comme un acteur clé de la transformation durable de l'agriculture.

### 1.4. Chiffres clés

La société PELLENC est active sur plusieurs marchés, notamment dans l'arboriculture et l'entretien des espaces verts et urbains, mais c'est surtout dans le secteur de la viticulture qu'elle réalise une grande partie de son chiffre d'affaires. La figure 4 présente l'évolution de ce chiffre d'affaires de PELLENC sur trois dernières décennies (2003, 2013 et 2023). En effet, ce chiffre a connu une forte croissance passant de 133 millions d'euros en 2013 à 311 millions d'euros en 2023. Bien qu'une baisse d'activités ait été observée cette année, en raison des tendances du marché, le bilan global reste positif. Aujourd'hui, l'entreprise compte 1900 collaborateurs, dont 200 spécialisés en recherche et développement, un axe stratégique auquel elle consacre 6 % de son chiffre d'affaires annuel. Cette orientation vers l'innovation se traduit également par un total de 1346 brevets déposés en 2023, confirmant son dynamisme technologique.

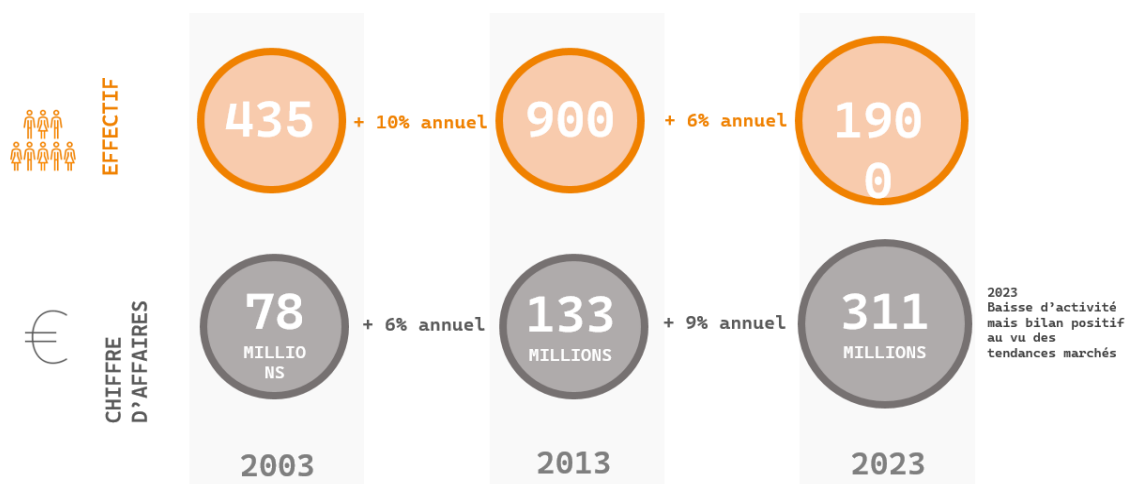


Figure 4 : Évolution du chiffre d'affaires de l'entreprise

Sur le marché des machines à vendanger, PELLENC se positionne comme un acteur important avec 31 % de parts de marché (figure 5), cependant elle fait face à des concurrents de premier plan. New Holland, leader incontesté, domine ce secteur avec près de 50 % des parts depuis cinq ans. Fondée en 1885 en Pennsylvanie, cette entreprise, spécialisée dans les engins agricoles et de chantier, a été rachetée par le groupe Fiat en 1999 et est aujourd'hui basée à Turin, en Italie. Un autre concurrent

notable est Grégoire, qui détient environ 17 % du marché. Créée en 1972 en France, cette entreprise est entièrement dédiée à la construction de machines à vendanger. Au-delà des machines à vendanger, PELLENC se distingue dans d'autres segments clés. Dans le marché des sècheurs électriques, elle occupe 24 % des parts, se classant ainsi derrière Electrocoup, leader avec 60 % (image de gauche de la figure 6). Dans le domaine des attacheurs, PELLENC domine largement avec 78 % des parts de marché, laissant loin derrière Infaco, qui en détient 11 % (image de droite de la figure 6). Cette polyvalence et cette expertise dans plusieurs secteurs permettent à PELLENC de maintenir une position compétitive face à des acteurs établis, tout en consolidant sa réputation d'innovateur et de spécialiste dans le domaine viticole et agricole.

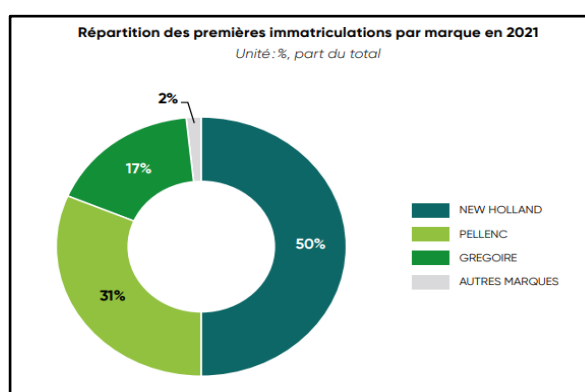


Figure 5 : Parts du marché des machines de vendanges en 2021

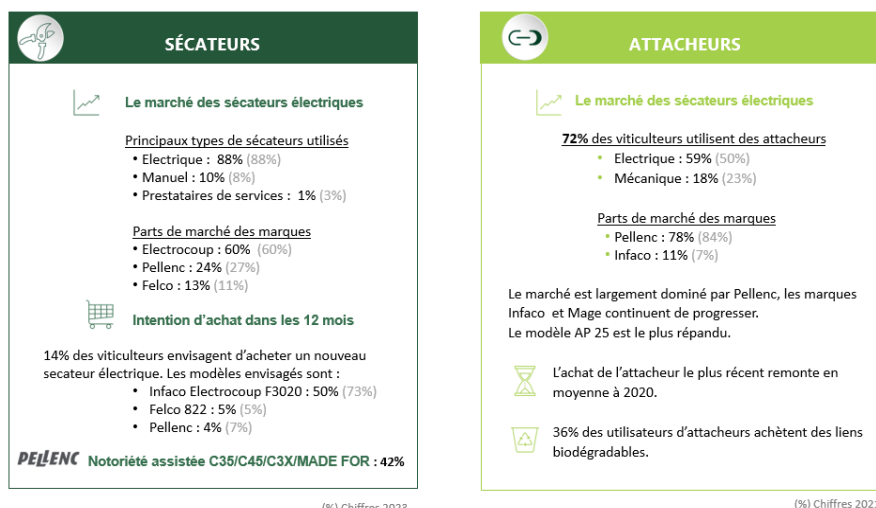


Figure 6 : Étude réalisée auprès de 400 viticulteurs interrogés par téléphone entre juillet et août 2024

## 2. Organisation du siège social à Pertuis

### 2.1. Structure de l'entreprise d'accueil

J'ai effectué mon stage au siège social du groupe PELLENC à Pertuis, Vaucluse (figure 7). L'effectif de cette société s'élève à environ 450 salariés. PELLENC S.A.S est une des rares entreprises Vauclusiennes à se développer à l'international. Elle a su de plus s'imposer comme leader mondial dans ses principaux domaines d'activité.



Figure 7 :Siège social du groupe PELLENC

Le site de Pertuis abrite les principales directions stratégiques du groupe PELLENC : Direction générale, financière, des ressources humaines, recherche et développement, marketing et industrielle.

Afin d'avoir un aperçu de la structure du siège de PELLENC à Pertuis, l'organigramme de la figure 8 présente l'organisation du site en différentes divisions et services. L'ensemble est placé sous la supervision de la Direction Générale, qui encadre la Direction Financière, la Direction Recherche & Développement, ainsi que la Direction des Ressources Humaines, de la Communication Interne et des Services Généraux.

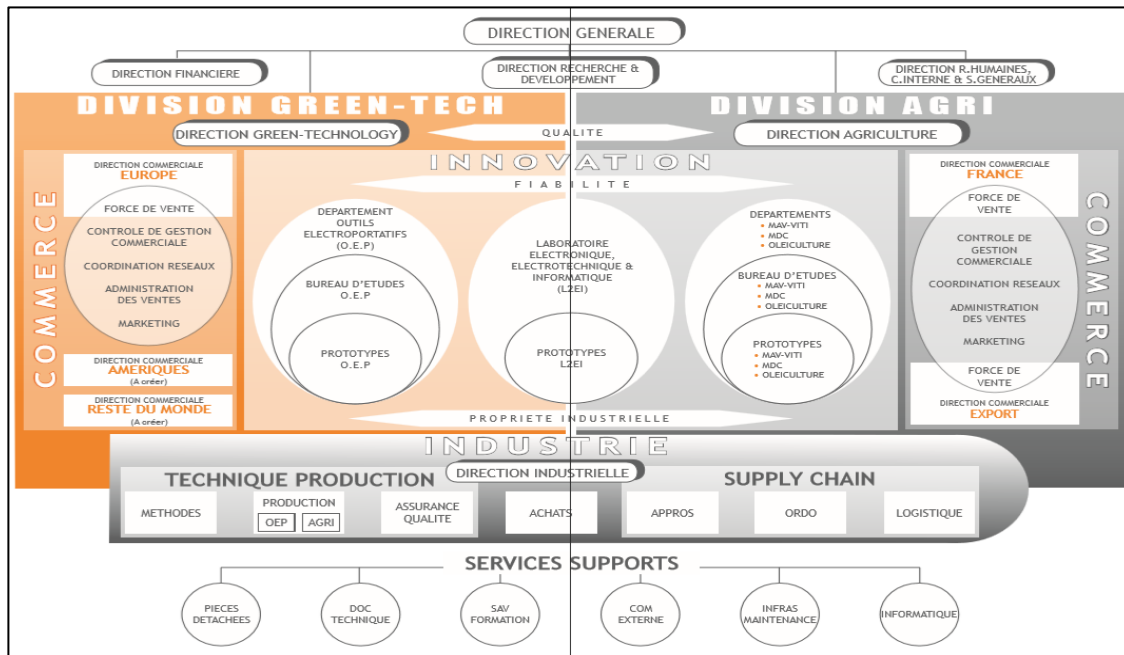


Figure 8 : Organigramme PELLENC S.A.S

L'organisation repose sur deux grandes divisions : « GREEN-TECH » et « AGRI ». La Division « GREEN-TECH », pilotée par la Direction Green-Technology, comprend une branche commerciale organisée en plusieurs pôles. Elle intègre aussi un volet technique représenté par le département des Outils Électroportatifs (O.E.P.) qu'est composé d'un bureau d'études et une section dédiée aux prototypes. De son côté, la division « AGRI », gérée par la Direction Agriculture, possède sa propre structure commerciale. Cette division comprend également plusieurs départements techniques spécialisés (MAV-VITI, MDC et Oléiculture), chacun disposant de son bureau d'études et de sa section prototypes. Un laboratoire dédié à l'électronique, l'électrotechnique et l'informatique (L2EI) vient compléter le pôle technologique des divisions « GREEN-TECH » et « AGRI ».

Ainsi, le département Recherche et Développement (R&D) est structuré en équipes ayant des missions distinctes. Certaines équipes se concentrent sur le développement, c'est-à-dire la conception industrielle des produits, tandis que d'autres se consacrent à la recherche et à l'évaluation de nouvelles technologies. Cette répartition des tâches permet à l'entreprise de combiner innovation et industrialisation, en explorant des solutions technologiques tout en assurant leur faisabilité et leur intégration dans des produits finis.

La Direction Industrielle, regroupe les fonctions essentielles associées à la technique Production et la Supply Chain. L'ensemble est soutenu par plusieurs Services Supports, qui assurent le bon fonctionnement de l'organisation.

### 2.2. Mode managérial

Depuis sa création en 1973 sous la direction de Roger PELLENC, l'entreprise a adopté une organisation pyramidale, un modèle structurant qui a favorisé son développement en alliant innovation et cohérence, tout en renforçant sa réputation de fiabilité. Cette approche organisationnelle a permis à l'entreprise de croître de manière structurée et efficace. En 2022, Simon BARBEAU, ancien dirigeant du groupe Brandt, où il avait mené avec succès une relance économique sur quatre ans, a été nommé PDG du groupe PELLENC par les actionnaires du groupe SOMFY. Sous sa direction, le siège social de Pertuis s'inscrit pleinement dans ce modèle hiérarchique vertical, où la Direction Générale centralise les décisions stratégiques avant de les décliner au sein des divisions et services spécialisés. D'après mes cours de management à l'École des Mines de Saint-Étienne et mes propres constatations, il est clair que PELLENC fonctionne avec un style de management directif.

### 2.3. Aménagement des locaux

Dans le bâtiment administratif du siège social, le rez-de-chaussée est occupé par les unités de production, de logistique et de prototypage. Tout le processus, de la conception initiale à la fabrication des prototypes, est réalisé en interne. Cette proximité entre les phases de conception et de production constitue un atout majeur, permettant à PELLENC de consolider son leadership en innovation et de répondre aux exigences spécifiques de sa clientèle.

Au premier étage, on retrouve les services des ressources humaines, du marketing et de l'administration des ventes. Le deuxième étage, où j'ai passé la majeure partie de mon stage, regroupe les bureaux d'étude R&D ainsi que la direction générale. Les bureaux d'études sont aménagés en open-space, favorisant la communication entre les équipes, mais pouvant parfois nuire à la concentration en raison du bruit ambiant. Par ailleurs, les chemins menant aux bureaux traversent les zones de stockage de la chaîne de production ainsi que les passages de gros engins, ce qui peut engendrer des problèmes de sécurité et de circulation.

J'ai été particulièrement frappé par la faible représentation des femmes dans le domaine de la recherche et de l'ingénierie, où elles ne constituent que 8 % des ingénieurs R&D. Cette inégalité est si prononcée qu'elle se manifeste même dans l'aménagement des locaux : à l'étage des bureaux d'études R&D, il y a deux toilettes pour hommes contre une seule pour femmes. En revanche, j'ai constaté que cette tendance s'inverse dans d'autres secteurs comme les ressources humaines (RH) ou le marketing, où les femmes sont souvent majoritaires. Dès lors, cette observation m'a conduit à



réfléchir sur les enjeux de parité et à prendre conscience des déséquilibres persistants dans le monde professionnel.



Figure 9 : Atelier de production PELLENC

## 2.4. Communication au sein de l'entreprise

Au cours de mon stage chez PELLENC, j'ai pu constater que la communication au sein des équipes des bureaux d'études est particulièrement fluide, grâce à la mise en œuvre de plusieurs pratiques collaboratives. Chaque employé dispose notamment d'un téléphone sans fil, ce qui simplifie considérablement les échanges internes.

Les équipes organisent des réunions quotidiennes, appelées « daily », qui durent environ 15 minutes. Ces rencontres permettent de faire le point sur l'avancement des projets et de discuter des éventuelles difficultés rencontrées. Par ailleurs, l'adoption de la méthode Agile « Scrum » constitue un véritable atout pour l'organisation et la gestion des projets. Cette méthodologie favorise non seulement une

structuration rigoureuse des tâches, mais aussi une communication continue entre les différents pôles. De plus, à la fin de chaque phase du projet (sprint), les équipes mènent une rétrospective de sprint. Cette étape essentielle permet de vérifier si les objectifs fixés ont été atteints. Dans le cas contraire, les raisons des écarts sont identifiées, et des pistes d'amélioration sont proposées pour optimiser les processus futurs.

Cependant, j'ai également constaté des limites, notamment en ce qui concerne la communication entre les différents niveaux hiérarchiques. En effet, lors d'une discussion informelle avec des monteurs-assembleurs, qui occupent des postes équivalents à ceux d'opérateurs, ces derniers ont exprimé leur frustration face à un manque de communication avec leurs supérieurs. Ce problème semble s'accroître pendant les périodes de forte activité, comme les vendanges, où la cadence de travail devient plus soutenue. Ces opérateurs ont également mis en lumière une disparité dans la gestion des horaires. Travaillant à l'heure, ils disposent de peu de flexibilité et doivent respecter des horaires stricts, notamment en reprenant leur poste immédiatement après la pause déjeuner à 13h30.

En revanche, les ingénieurs des bureaux d'étude bénéficient d'un système forfaitaire, qui leur offre une plus grande liberté dans l'organisation de leur temps de travail. Par exemple, ils peuvent aménager leurs horaires hebdomadaires pour atteindre le quota de 35 heures par semaine avec plus de souplesse.

### **3. L'équipe R&T et ses spécificités**

J'ai effectué mon stage au sein du service Recherche et Technologie (R&T), dirigé par M. Thierry STEIN, qui était également mon tuteur de stage. La figure 10 représente l'organigramme du service R&T. Ce service a pour mission d'évaluer le potentiel des nouvelles technologies en tenant compte des contraintes industrielles ainsi que des besoins spécifiques des différents domaines d'activité de PELLENC, tels que la viticulture, l'arboriculture et l'entretien des espaces verts.



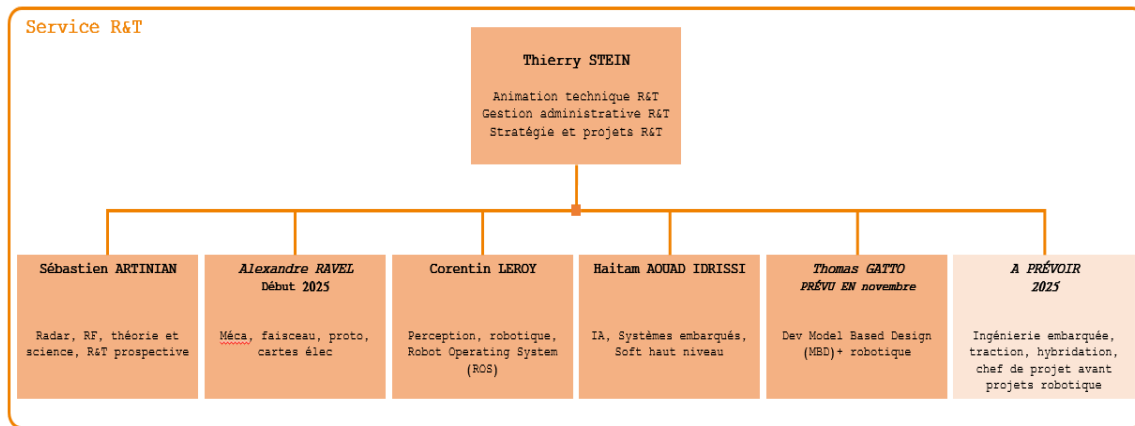


Figure 10 : Équipe R&amp;D

La figure 11 représente le diagramme R&D du groupe PELLENC. Ce diagramme met en évidence l'interconnexion entre les différents services et leur contribution à la stratégie globale de l'entreprise permettant une transition efficace entre la recherche et le développement. Au sommet du diagramme, la Direction Générale (DG) supervise l'ensemble du fonctionnement, tandis que le CODIR R&T (Recherche et Technologie) du groupe joue un rôle clé dans la gestion et l'orientation des projets de recherche et de technologie. Directement rattachée à cette structure, la Direction Innovation est responsable du développement technologique et de la collaboration avec des organismes et partenaires de recherche externes. Deux services essentiels lui sont associés : le Service PI (Propriété Intellectuelle) et le Service R&T, qui travaillent ensemble pour protéger et faire évoluer les innovations.

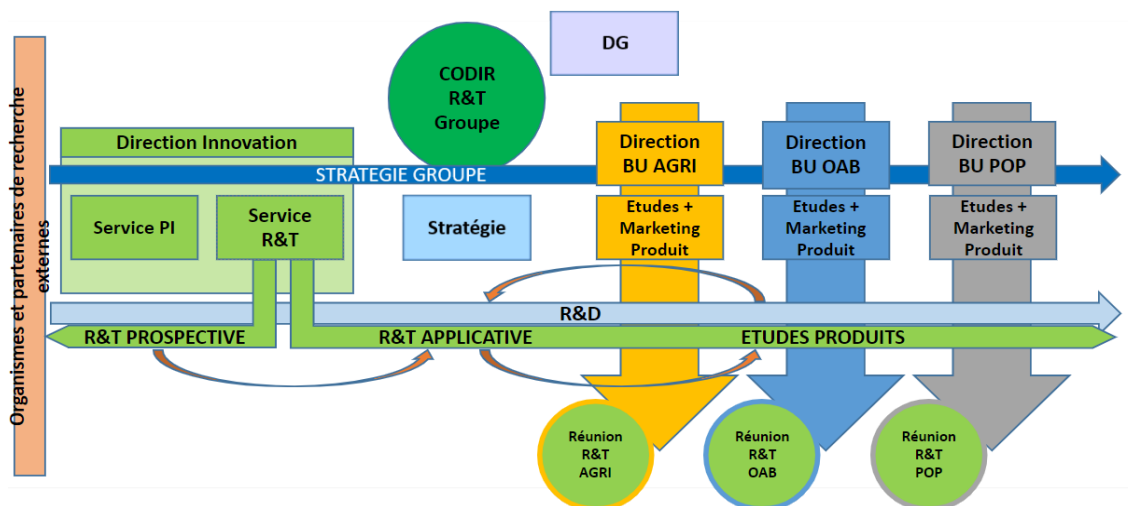


Figure 11 : Diagramme R&amp;D

Le schéma de la figure 11 illustre le lien entre la R&D et les études produits menées par les trois Business Units (BU AGRI, BU OAB et BU POP). Chaque BU est chargée des études et du marketing produit, assurant ainsi l'adéquation entre les innovations technologiques et les besoins du marché. La coordination entre ces entités est renforcée par des réunions R&T spécifiques à chaque BU, permettant de suivre l'avancement des projets et d'adapter les stratégies en conséquence.

Le service R&T fonctionne en collaboration avec la R&D étant donné qu'il étudie principalement des solutions qui pourraient ensuite être confiées à la R&D afin d'étudier leur implémentation. Dans ce contexte, deux approches de recherche sont menées par le service R&T: la première approche est la R&T Prospective, qui se concentre sur des travaux exploratoires. Ces derniers sont souvent confrontés à de nombreux verrous technologiques et incertitudes, ce qui rend leur aboutissement incertain et nécessite des délais plus longs estimés à plus de cinq ans. La deuxième approche est la R&T Applicative, qui vise à transformer rapidement certaines avancées technologiques en solutions concrètes et développer un POC (Proof of Concept), qui sera ensuite pris en charge par les bureaux d'études des différentes BU pour une éventuelle industrialisation (avec un délai moyen estimé entre 3 et 5 ans).

Par ailleurs, le service R&T partage le même open-space que la R&D, il ressemble en fait plus à une équipe qu'à un réel service. Cependant, il réside une grande différence entre les deux en termes de style de management et les attendus spécifiques. Dans la R&D, des réunions directives sont tenues chaque jour afin de fixer les objectifs attendus et associer une mission à chacun de ses membres. À l'inverse, dans le service R&T, étant basé sur la recherche et l'étude de technologies sur de longues périodes avec moins de résultats à court terme à présenter, et ayant un effectif réduit, les réunions tenues sont des réunions informatives et l'ensemble des membres du service travaillent ensemble.

Enfin, la R&T produit chaque année des rapports détaillés sur l'ensemble des projets menés, permettant à l'entreprise de bénéficier des crédits d'impôt recherche. Ces crédits permettent de déduire 30 % des investissements en recherche sous forme d'amortissement, ce qui contribue à financer les salaires de la moitié des ingénieurs en R&D.

## II. Missions et réalisations

---

### 1. Fondements théoriques

#### 1.1. IA, apprentissage automatique et vision par ordinateur

L'intelligence artificielle (IA) est un domaine qui est dédié à la création de systèmes capables d'imiter certaines fonctions cognitives humaines. L'objectif principal de l'IA est de permettre aux machines d'exécuter de manière autonome des tâches qui, en général, nécessitent une intervention humaine. Ces tâches incluent, entre autres, la perception, le raisonnement, l'apprentissage, la résolution de problèmes, et la prise de décision.

L'apprentissage automatique ou Machine Learning (ML) est une branche de l'IA et le fondement de la plupart des solutions d'IA. Il permet aux ordinateurs d'apprendre à partir de données sans suivre des instructions prédéfinies pour effectuer une tâche spécifique. Les systèmes de Machine Learning sont conçus pour identifier des modèles et des relations à partir des données et utiliser ces modèles pour accomplir diverses tâches, telles que la classification, la prédiction ou la recommandation.

La vision par ordinateur (Computer Vision) est un domaine spécifique de l'IA qui utilise le ML pour permettre aux machines d'interpréter et d'analyser des images ou des vidéos de la même manière que le font les êtres humains. La vision par ordinateur englobe un large éventail d'applications, allant de la reconnaissance faciale à la détection d'objets, en passant par la segmentation d'images, la reconstruction 3D, la surveillance vidéo, etc. Elle est utilisée dans de nombreux domaines tels que la médecine, l'automobile, la sécurité, la robotique, l'industrie, les médias, etc.

#### 1.2. Processus d'apprentissage automatique

L'apprentissage automatique consiste à développer un modèle à l'aide d'un algorithme de minimisation. Cet algorithme ajuste progressivement les paramètres du modèle afin de réduire au maximum l'erreur entre les prédictions et les valeurs réelles. En optimisant une fonction de coût, il permet au modèle d'améliorer sa précision et sa capacité à généraliser sur de nouvelles données. Une fois que le modèle a été correctement entraîné, il est en mesure d'exécuter la tâche pour laquelle il a été conçu.

L'apprentissage automatique comprend deux types : supervisé et non supervisé. Dans l'apprentissage supervisé, le modèle est entraîné sur des données étiquetées, où chaque entrée a une sortie connue, ce qui lui permet d'apprendre par correction d'erreur. À l'inverse, l'apprentissage non supervisé

fonctionne sur des données non étiquetées, où le modèle identifie des structures ou des regroupements sans indication préalable. A titre d'exemple, un algorithme d'apprentissage automatique supervisé utilisé pour classifier des photos en fonction du nombre d'humains présents sur ces dernières va initialement recevoir des milliers d'images de personnes, celles-ci seront alors appelées données d'entraînement. En effet, pendant la phase d'entraînement, le modèle va apprendre à classifier des images dont on connaît déjà la nature, elles sont dites labellisées ou étiquetées. Une fois l'entraînement terminé, la performance du modèle est évaluée à l'aide de données de test. Cet ensemble de test contient des données pour lesquelles les étiquettes sont connues mais qui n'ont pas été exposées au modèle auparavant. Ainsi, si le modèle se révèle performant sur l'ensemble de test, cela indique qu'il est fiable.

La figure 12 illustre le processus de classification en apprentissage automatique supervisé. Tout commence par l'entrée des données, constituées de différentes formes (carrés, ronds, triangles), qui servent à entraîner un modèle. Pendant cette phase d'apprentissage, l'algorithme ajuste ses paramètres pour associer correctement chaque forme à une catégorie. Une fois le modèle entraîné, il est capable de prédire la classe de nouvelles formes en les comparant à celles déjà apprises. Enfin, la sortie du modèle montre le résultat de cette classification, où les objets sont identifiés et triés selon leur catégorie.

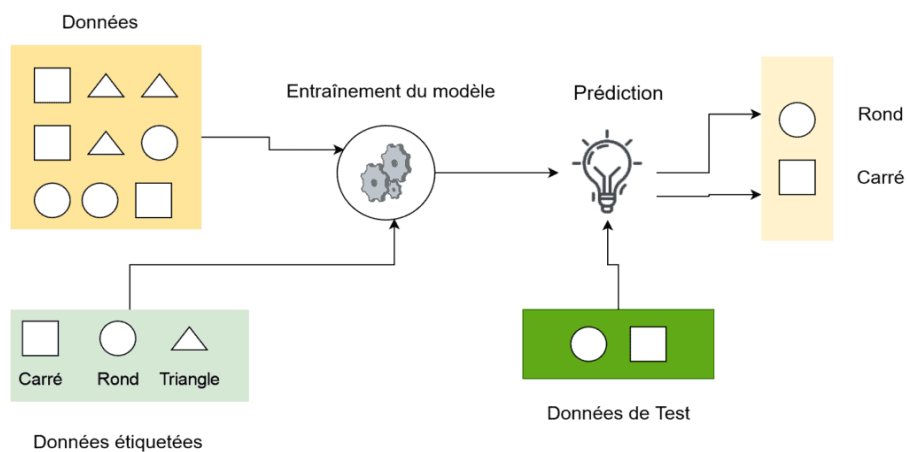


Figure 12 : Étapes de l'apprentissage automatique d'un modèle

### 1.3. Apprentissage profond

L'apprentissage profond ou Deep Learning est l'une des technologies principales du Machine Learning. Avec le Deep Learning, nous parlons de systèmes capables de mimer les actions du cerveau

humain grâce à des réseaux de neurones artificielles. La figure 13 illustre la hiérarchie entre intelligence artificielle (IA), apprentissage automatique et deep learning.

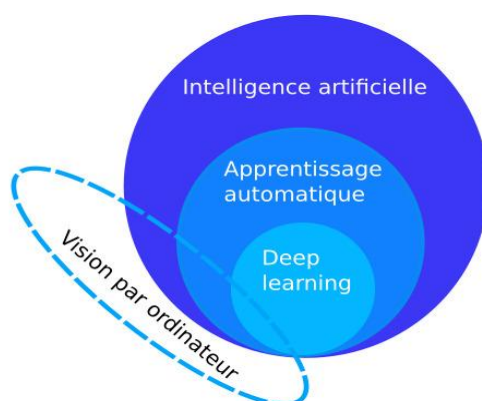


Figure 13 : Deep Learning et Machine Learning

Les neurones présents dans le cerveau humain fonctionnent selon un schéma simple (figure 14) : Un neurone va recevoir des signaux excitateurs via les synapses qui se situent aux terminaisons des dendrites. Ainsi, les dendrites constituent les entrées du neurone et les axones leurs sorties vers d'autres neurones. Le corps cellulaire se charge d'activer plus ou moins le neurone, en fonction des intensités électriques des signaux excitateurs. Cette activation se produit lorsque le potentiel d'action du neurone dépasse un certain seuil, que l'on appelle « potentiel d'activation ».

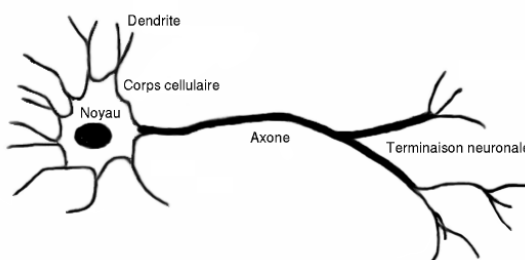


Figure 14 : Schéma d'un neurone biologique

Dans le cas d'un modèle d'Intelligence Artificielle, le fonctionnement est très similaire. En effet, les réseaux de neurones sont un ensemble de neurones, organisés en plusieurs couches et reliés entre eux. En pratique, les neurones de la première couche vont transmettre des informations simples aux neurones de la couche suivante. Ce qui aura une influence sur leur degré d'activation et ainsi de suite. La dernière couche nous fournira alors des informations sur le résultat.

La figure 15 représente un réseau de neurones profond, composé de plusieurs couches. Les neurones bleus forment la couche d'entrée, où les données initiales sont introduites. Les neurones noirs appartiennent aux couches cachées, où les calculs et transformations des données sont effectués.

Enfin, les neurones verts constituent la couche de sortie, qui produit le résultat final. Ce type d'architecture permet aux modèles d'apprentissage profond d'extraire des caractéristiques complexes et d'effectuer des prédictions précises.

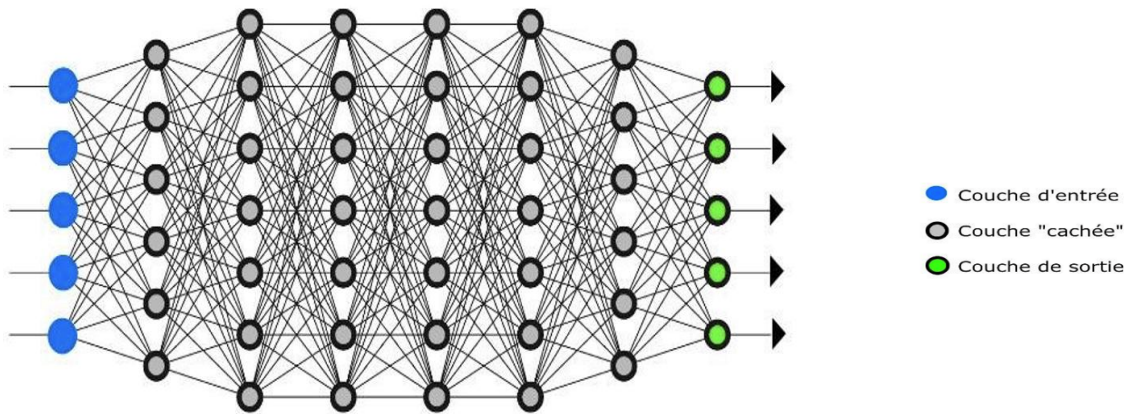


Figure 15 : Architecture des réseaux de neurones profond

### 1.4. Fonctionnement d'un CNN

Les réseaux de neurones convolutifs (Convolutional Neural Networks : CNNs) sont une sous-catégorie des réseaux neuronaux profonds, principalement utilisés pour l'analyse des images visuelles. Inspirés par la manière dont le cortex visuel humain traite les stimuli visuels, les CNNs ont transformé le domaine de la vision par ordinateur en permettant des tâches telles que la classification des images, la détection des objets, et la segmentation d'images. Ces réseaux CNNs se distinguent des réseaux classiques par l'utilisation de couches de convolution, qui permettent d'extraire automatiquement des caractéristiques importantes (bords, textures, formes) à différents niveaux d'abstraction.

La figure 16 illustre le fonctionnement d'un réseau CNN pour la classification d'images. L'entrée est une image qui passe par des couches de convolution et de pooling, où les caractéristiques sont extraites et réduites. Ensuite, ces données sont vectorisées et envoyées vers une couche entièrement connectée, qui associe les caractéristiques aux différentes catégories possibles (ex. oiseau, chien, chat). Le modèle produit des probabilités pour chaque classe, permettant ainsi une classification multi-catégorie.

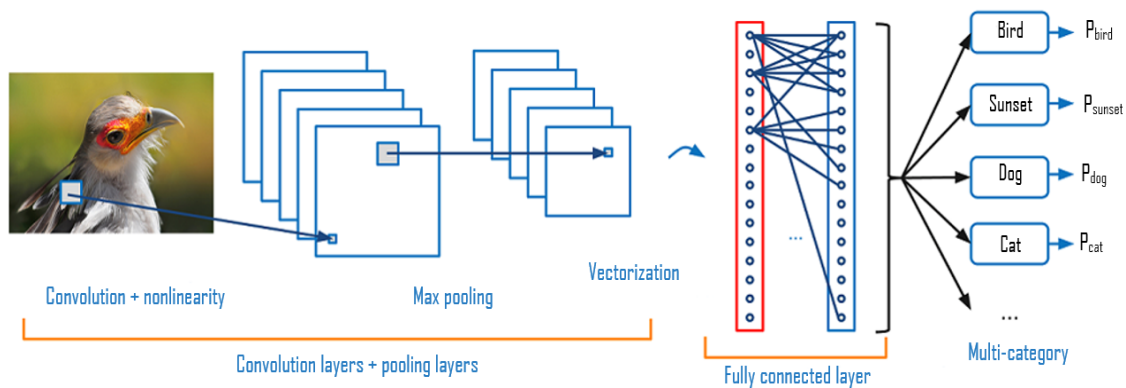


Figure 16 : Architecture d'un CNN

Les couches de convolution, fondamentales dans les CNN, se composent de filtres représentés par des matrices de valeurs, appelées cartes de caractéristiques (feature maps). Lorsqu'une image est traitée par une couche de convolution, celle-ci applique un filtre pour produire une carte de caractéristiques. Par exemple, une image de 5x5 pixels traitée avec un filtre 3x3 générera une carte de 3x3. Ces filtres sont conçus pour détecter des éléments spécifiques dans l'image.

Après chaque couche de convolution, les données passent généralement par une couche de pooling. Celle-ci agit comme une réduction dimensionnelle en divisant les cartes de caractéristiques en blocs et en conservant, par exemple, la valeur maximale de chaque bloc (max-pooling). Cela réduit la taille des cartes tout en préservant les informations essentielles, ce qui permet de diminuer la complexité tout en conservant les éléments les plus significatifs.

Une fois les données traitées par les couches de convolution et de pooling, elles atteignent la couche entièrement connectée. Celle-ci combine les caractéristiques extraites pour classer l'image et génère un vecteur de taille  $N$ , où  $N$  représente le nombre de classes dans le problème de classification. Chaque élément du vecteur indique la probabilité que l'image appartienne à une classe spécifique. Par exemple, pour une classification entre chiens et chats, un vecteur de sortie  $[0.8, 0.2]$  signifie que l'image a 80 % de chances de représenter un chien et 20 % de chances de représenter un chat.

### 1.5. Le modèle YOLO

Une percée dans le domaine de vision par ordinateur est survenue avec le développement du modèle YOLO (You Only Look Once). Conçu comme un modèle qui a révolutionné l'approche de la détection d'objets dans les images. En effet, les modèles de détection traditionnels impliquaient souvent un processus en deux étapes : d'abord identifier les régions d'intérêt, puis classer ces régions. En contraste, YOLO a innové en prédisant à la fois les classifications et les boîtes englobantes

(bounding boxes) en une seule passe à travers le réseau neuronal, accélérant considérablement le processus et améliorant les capacités de détection en temps réel.

La figure 17 illustre le fonctionnement d'un modèle YOLO, il est représenté par les principales étapes suivantes :

1. **Division en grille** : L'image d'entrée est divisée en une grille de taille  $S \times S$ , où chaque cellule est responsable de la détection d'objets dans sa région.
2. **Prédiction des boîtes englobantes**: Chaque cellule prédit plusieurs boîtes englobantes (rectangles entourant des objets) , avec un score de confiance indiquant la probabilité qu'un objet soit présent dans cette région.
3. **Carte de probabilité des classes** : Un réseau de neurones estime les probabilités d'appartenance des objets détectés à différentes classes (ex. chien, vélo, voiture).
4. **Sélection des détections finales** : Après l'application de techniques comme la suppression non maximale (Non-Maximum Suppression : NMS) pour éliminer les prédictions redondantes, les objets finaux sont détectés avec leurs boîtes de délimitation colorées.

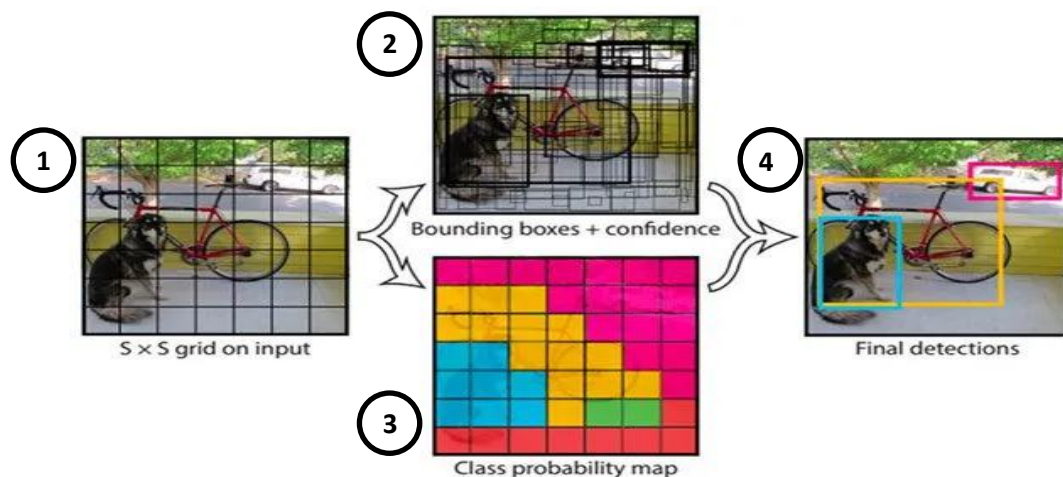


Figure 17 : Fonctionnement du modèle YOLO

La suppression non maximale NMS fonctionne comme suit :

1. Les boîtes englobantes appartenant à une même classe avec une probabilité inférieure à un seuil défini, par exemple 0.6, sont supprimées.
2. La boîte englobante ayant la plus haute probabilité de détection pour la classe est sélectionnée.



3. Les boîtes englobantes qui chevauchent celle sélectionnée, avec un IoU (Intersection over Union, voir partie métrique d'évaluation) supérieur à un seuil sont supprimées, car elles représentent probablement le même objet.
4. On réitère les étapes précédentes.

Le chevauchement entre les boîtes englobantes est évalué à l'aide de l'IoU : Si l'IoU entre deux boîtes englobantes dépasse un seuil de 0.5, elles sont considérées comme représentant le même objet. Une fois l'algorithme exécuté, seules les boîtes englobantes les plus pertinentes sont conservées, réduisant ainsi le nombre initial de détections. Par exemple sur l'image 4 de la figure 17, au lieu d'une centaine de boîtes englobantes initiales, seules trois subsistent pour caractériser trois objets : le vélo, le chien et la voiture.

## **2. Présentation du projet de stage**

### **2.1. Besoins et objectifs**

L'utilisation de machines agricoles pour la récolte des olives s'est fortement développée, permettant d'améliorer l'efficacité et de réduire les coûts de main-d'œuvre. Afin de cibler correctement les troncs d'oliviers et réduire les dommages aux arbres, l'entreprise PELLENC utilisait un système basé sur un nuage de points LiDAR pour la détection des troncs.

Avant mon arrivée chez PELLENC, une réunion a été organisée au sein de l'équipe R&T pour définir les objectifs et le déroulé de mon stage. L'objectif principal retenu était de déterminer si l'utilisation de modèles d'intelligence artificielle, en particulier basés sur la vision par ordinateur, permet d'automatiser davantage cette tâche avec une grande précision.

L'enjeu pour la R&D serait d'intégrer un modèle intelligent dans des calculateurs embarqués qui seront placés sur les machines à vendanger Optimum (mavO), afin de les rendre plus autonomes, offrant ainsi un gain de temps, une meilleure adaptation de ces machines et une gestion plus efficace des exploitations oléicoles.

L'objectif principal de mon projet de stage est de développer un modèle pour la détection des troncs d'oliviers dans des images et des vidéos afin d'optimiser le fonctionnement des machines agricoles de récolte. Plus précisément, il s'agit de constituer une base de données d'images annotées contenant des troncs d'oliviers sous différentes conditions, d'entraîner un modèle capable de les identifier avec précision, et d'optimiser ses performances en ajustant les hyperparamètres et en améliorant la qualité des données d'apprentissage. Une fois le modèle entraîné, il sera testé et évalué.

Pendant ce stage d'un mois, mon rôle a consisté à soutenir l'équipe R&T de l'entreprise dans l'exploration de l'IA, un enjeu particulièrement stratégique en raison de son besoin d'expertise dans ce domaine.

### 2.2. Déroulement du stage

Pour mener à bien mon projet de stage dans les délais, un plan d'action a été établi lors de la réunion d'intégration avec l'équipe R&T. Le déroulement de ce stage, d'une durée de quatre semaines (du 6 au 31 janvier 2025) est décrit dans la suite de cette section.

**La première semaine** a été consacrée à la découverte de l'entreprise, ainsi qu'à la maîtrise de l'environnement de travail et des outils mis à ma disposition, notamment l'environnement de développement sous Docker. J'ai également profité de cette période pour renforcer mes compétences en intelligence artificielle à travers l'étude d'articles scientifiques spécialisés, afin de répondre efficacement aux besoins techniques du projet qui m'a été confié. Une attention particulière a été portée à la version YOLOv8, reconnue pour ses performances de détection rapide et précise.

**La deuxième semaine** a été dédiée à plusieurs sessions de brainstorming pour élaborer un protocole d'acquisition de données, définissant les méthodes de capture et de sauvegarde des données. Ce protocole a ensuite été testé sur le parking de l'entreprise, où des oliviers étaient plantés. Pour cette expérimentation, une caméra Intel RealSense D435 (caméra de profondeur stéréo) a été montée sur une poussette, permettant de varier les positions et les angles de prise de vue. Ce test nous a permis d'évaluer la faisabilité du protocole pour mieux optimiser notre temps lors de la visite de la parcelle agricole.

Au cours de **la troisième semaine**, j'ai réalisé l'annotation manuelle de l'ensemble de données composé d'images et de vidéos capturées sur le parking, en ciblant spécifiquement la détection des troncs d'oliviers. Cette étape cruciale a permis de constituer une première base de données labellisée, indispensable pour l'apprentissage du modèle. Une fois l'étiquetage terminé, j'ai implémenté et entraîné le modèle YOLOv8. Par la suite, j'ai évalué les résultats obtenus en suivant un protocole d'évaluation, préalablement défini, afin de mesurer l'efficacité du modèle en termes de précision de détection.

**La quatrième semaine** a été marquée par la visite d'une parcelle agricole à Bournissac à Noves, où nous avons réalisé des acquisitions de données sur différents types d'oliviers en suivant le protocole préalablement établi. Nous avons utilisé plusieurs capteurs, notamment la camera RealSense, une

caméra ZED, un LiDAR et un radar. Dans un second temps, j'ai documenté l'ensemble des données collectées dans un fichier indexé au format JSON, décrivant les détails de chaque capture (état de l'environnement, lieu, heure, position des capteurs, etc.), afin d'aider l'équipe à identifier les paramètres influençant l'apprentissage du modèle. Le reste de la semaine a été consacré à l'entraînement et à l'évaluation des performances du modèle final. Enfin, j'ai analysé et mis en forme les résultats, que j'ai présentés à l'équipe lors de la dernière journée du stage. Cette présentation a permis de discuter la pertinence de l'approche proposée et les perspectives de l'utilisation de l'IA.

Ainsi, ce stage m'a permis de découvrir le métier d'ingénieur en bureau d'études sous tous ses aspects, notamment la collaboration avec d'autres ingénieurs et l'élaboration d'un compte rendu clair et exploitable pour l'équipe. Cette expérience a été enrichissante, tant sur le plan technique que professionnel.

### **3. Mes missions**

#### **3.1. Prise en main de l'environnement de développement**

Durant ce stage, ma principale mission était d'entraîner un modèle YOLOv8 sur un ensemble de données (dataset) spécifique afin qu'il puisse détecter les troncs d'oliviers. Pour y parvenir, j'ai adopté une approche méthodique en me familiarisant d'abord avec Docker, un outil qui m'a permis de créer un environnement de développement isolé et reproductible, évitant ainsi les problèmes liés aux dépendances et aux configurations spécifiques des machines, et garantissant une exécution fiable du code sur diverses plateformes. En parallèle, l'utilisation de notebook Jupyter m'a permis de structurer mon code de manière organisée tout en intégrant des explications détaillées et des résultats directement dans le même document, ce qui a facilité la compréhension, le suivi et la présentation de mon travail.

Pour manipuler les données et implémenter et élaborer le modèle YOLOv8, j'ai utilisé des bibliothèques Python telles que PyTorch et OpenCV, qui m'ont permis de gagner un temps précieux grâce à leurs fonctions optimisées et bien documentées. Ces bibliothèques ont particulièrement facilité le prétraitement des images du dataset ainsi que l'entraînement du modèle, en offrant des outils puissants et adaptés à la gestion des données visuelles et à l'optimisation des performances de modèle.

Ce stage m'a permis d'appliquer et d'approfondir mes connaissances en Linux et en Python, acquises lors de mes cours à l'École des Mines de Saint-Étienne. La maîtrise de Linux m'a été particulièrement utile pour la configuration et la gestion de l'environnement Docker, ainsi que pour l'exécution

efficace de scripts et de commandes complexes. Quant à Python, il a été au cœur de mon travail, aussi bien pour la manipulation des données que pour l'utilisation des bibliothèques de Deep Learning. Ces compétences développées en école m'ont permis de m'adapter rapidement aux exigences techniques du projet et d'y contribuer de manière significative.

#### 3.2. Collecte des données

La collecte de données en vue du réentraînement du modèle YOLO pour la détection des troncs d'oliviers représente une étape cruciale, qui requiert une attention particulière à plusieurs éléments clés. Il est indispensable, en premier lieu, d'établir un protocole d'acquisition rigoureux avant même de choisir le site de collecte des données. Ce dernier doit être soigneusement sélectionné afin de refléter les conditions représentatives des diverses situations auxquelles le modèle pourra être confronté, en veillant notamment à inclure une diversité d'oliviers.

Dans le cadre de mon stage, j'ai tout d'abord collaboré avec l'équipe R&T à l'élaboration d'un protocole d'acquisition de données, initialement testé et validé sur le parking de l'entreprise, où des oliviers étaient implantés. Par la suite, nous avons choisi la parcelle agricole à Bournissac, à Noves, pour y mener des campagnes d'acquisition conformément au protocole validé. Celles-ci ont été réalisées à l'aide de plusieurs capteurs, notamment une caméra RealSense, une caméra ZED, un LiDAR et un radar, tous montés sur un chariot mobile permettant d'ajuster facilement les positions et les angles de prise de vue.



Figure 18 : Parcelle agricole à Bournissac, à Noves



Figure 19 : Chariot mobile d'acquisition de données

### 3.3. Construction du dataset annoté

Pour la détection des troncs d'oliviers, l'annotation des données est une étape essentielle pour garantir un entraînement efficace du modèle. Dans ce cadre, j'ai choisi d'utiliser la plateforme Roboflow. Cette dernière facilite la tâche d'annotation en offrant une interface intuitive permettant de délimiter précisément les contours des troncs grâce à la segmentation d'instance, une technique plus précise que la simple détection par boîtes englobantes. En effet, cette technique permet d'identifier non seulement la présence d'un objet, mais aussi ses contours exacts.

Le processus commence par l'importation des images et vidéo de la dataset. Chaque image est ensuite annotée manuellement. J'ai structuré mon travail d'annotation en définissant 12 classes distinctes : quatre pour les éléments parasites (feuillage, humains, poteaux, tuyaux) pouvant interférer avec la détection, et huit classes détaillées caractérisant les troncs selon leur distance (loin ou proche), leur obstruction (obstrué ou non) et leur ombrage (ombragé ou non). Par exemple, un tronc proche, non obstrué et ombragé est labellisé "proche\_Nobs\_omb", comme illustré dans la figure 20.



Ce travail d'annotation manuelle, bien que long et fastidieux, est absolument essentiel, car il impacte directement la précision et les performances finales du modèle de détection.

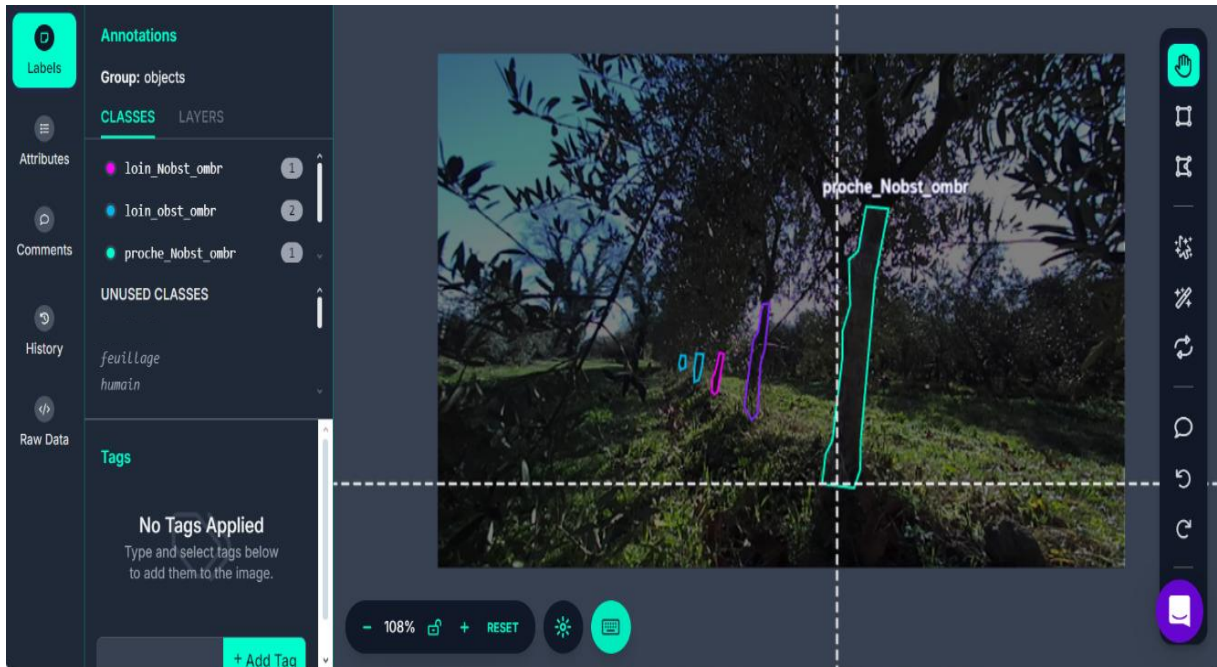


Figure 20 : Annotation manuelle

Après finalisation des annotations, j'ai exploité les fonctionnalités d'augmentation de données de Roboflow (rotations, ajustements de luminosité, ajout de bruit) pour enrichir et diversifier le dataset, améliorant ainsi les capacités de généralisation du modèle. Une attention particulière a été portée sur l'augmentation des variations angulaires, d'exposition et d'éclairage. Ces trois paramètres représentant les principales sources de variabilité dans les conditions réelles d'acquisition, dues aux variations lumineuses solaires, aux ombres portées du feuillage et aux mouvements du robot lors de la capture. Cette approche permet d'accroître significativement la robustesse du modèle, garantissant ainsi une détection plus fiable en conditions réelles. J'ai ensuite généré une version du dataset avec la répartition suivante : 70 % pour l'entraînement, 20 % pour la validation et 10 % pour les tests, assurant un équilibre optimal pour l'apprentissage et l'évaluation du modèle.

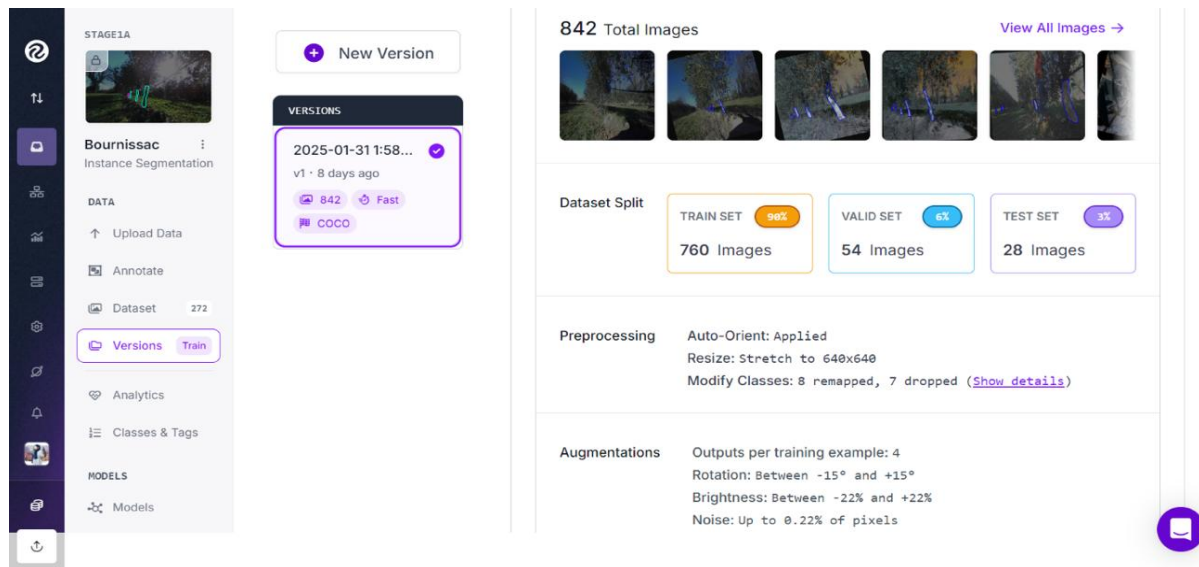


Figure 21 : Génération d'une version du dataset

Enfin, les annotations sont exportées dans un format compatible avec YOLOv8. Afin d'optimiser le flux de travail lors de l'entraînement du modèle, j'ai opté pour un chargement direct du dataset dans le notebook Jupyter. Alternativement, le dataset peut être téléchargé localement sous forme d'un répertoire structuré en trois sous-dossiers : entraînement, validation et test. Chacun de ces dossiers contient deux sous-répertoires distincts : l'un pour les images et l'autre pour les labels. Ces derniers sont stockés sous forme de fichiers texte, précisant l'identifiant des classes ainsi que les coordonnées des contours des objets annotés, conformément aux spécifications techniques requises par l'architecture YOLO.

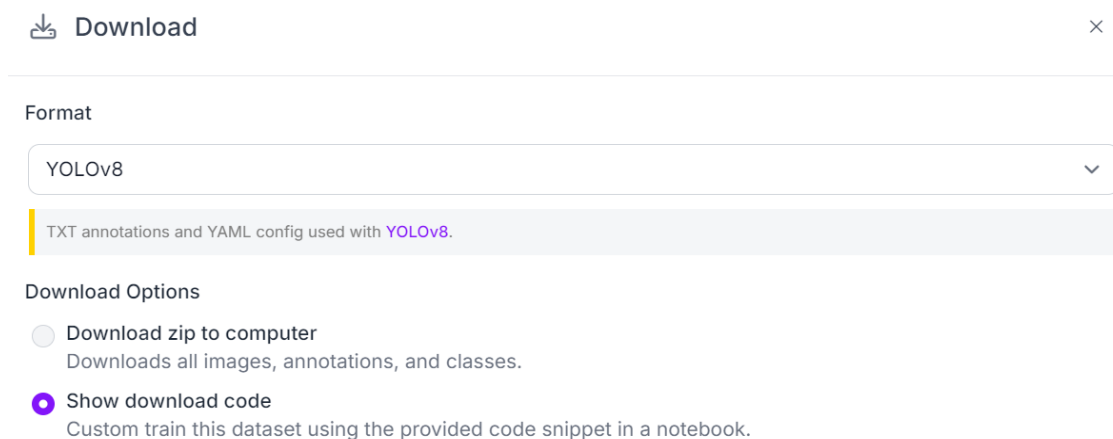


Figure 22 : Options pour exporter le dataset

#### 3.4. Filtrage du dataset

Dans le cadre de la préparation des données pour l'entraînement du modèle YOLOv8, une démarche de filtrage des données a été mise en place afin d'évaluer l'impact d'un paramètre spécifique sur les performances du modèle. À cet effet, j'ai développé un script Python (voir Annexe) structuré autour de deux fonctions principales :

- La fonction « filtre » qui traite individuellement les fichiers d'annotation en appliquant un ensemble de règles prédéfinies (suppression et remplacement de classes, exclusion d'éléments spécifiques). Elle réorganise ensuite les images d'entraînement associées dans une nouvelle arborescence structurée.
- La fonction « filter\_all » : Elle généralise cette approche à l'ensemble du dataset, en parcourant systématiquement les sous-répertoires (test, validation et entraînement). De plus, elle met à jour automatiquement le fichier de configuration « data.yaml » afin de maintenir la cohérence des données après la modification des classes.

Prenons l'exemple de l'impact de la présence de chaussettes sur les troncs : dans un premier temps, le modèle est entraîné sur l'ensemble du dataset. Ensuite, une version filtrée du dataset est générée en excluant toutes les images contenant des troncs avec chaussettes, et un second modèle est entraîné sur ce dataset modifié. La comparaison des précisions des deux modèles permet alors de quantifier l'influence du paramètre étudié sur la capacité de détection du modèle.

#### 3.5. Réentraînement du modèle YOLOv8

Le réentraînement (fine-tuning) du modèle YOLOv8 pour la détection des troncs d'oliviers consiste à adapter une architecture pré-entraînée à des données spécifiques, afin d'optimiser ses paramètres pour cette tâche particulière. Cette approche permet d'améliorer significativement la précision et la robustesse des détections, tout en conservant les performances en temps réel. Cette tâche nécessite la préparation préalable de l'ensemble de données, incluant l'annotation des images des troncs d'arbres, suivie de l'organisation structurée des données en ensembles d'entraînement, de validation et de test, définis dans le fichier de configuration « data.yaml ».

J'ai choisi d'utiliser la version pré-entraînée « yolov8s-seg.pt » du modèle YOLOv8. Ce choix repose sur un compromis optimal entre performance, efficacité et vitesse d'inférence qui sont cruciales pour des applications agricoles en temps réel. Ainsi, l'utilisation de poids pré-entraînés sur de larges



ensembles de données (comme COCO) permet de réduire significativement le besoin en données annotées spécifiques et d'accélérer la convergence du modèle.

Avant de lancer l'entraînement, il est essentiel de configurer l'environnement de travail en installant les bibliothèques requises (voir Annexe). Cela inclut OpenCV pour le traitement d'images, NumPy pour les calculs numériques et Ultralytics pour l'exploitation du modèle YOLO.

Le modèle pré-entraîné « yolov8s-seg.pt » est chargé puis configuré en ajustant les poids du réseau de neurones via un processus d'optimisation, où plusieurs paramètres sont initialement définis, tels que le nombre d'époques (il représente le nombre de fois où l'ensemble des données d'entraînement est entièrement parcouru par le modèle), la taille en pixels des images en entrée et la taille du batch (définit le nombre d'images traitées simultanément avant que le modèle ne mette à jour ses poids). Pendant l'entraînement, le modèle améliore progressivement sa capacité à détecter les troncs d'oliviers en réduisant l'écart entre ses prédictions et les annotations réelles.

Une fois l'apprentissage achevé, il est évalué sur un ensemble d'images de test afin d'analyser sa robustesse et sa capacité de généralisation. Les résultats de performance sont enregistrés dans un fichier JSON « data.json » pour une analyse ultérieure, tandis que le modèle final est sauvegardé sous le nom « best.pt », prêt à être utilisé pour la détection en temps réel des troncs d'oliviers.

### **3.6. Évaluation du modèle**

#### **3.6.1. Métriques d'évaluation**

##### **3.6.1.1. IoU (Intersection over Union)**

Étant donné que mon approche repose sur un modèle de segmentation, l'IoU constitue une métrique essentielle. Elle quantifie le chevauchement entre la région prédite et la région réelle (Une IoU élevée indique une segmentation précise des troncs), selon la formule suivante :

$$\frac{\text{Nombre d'éléments correctement prédits pour une classe}}{\text{Nombre d'éléments appartenant à la classe}}$$

##### **3.6.1.2. Précision (Precision)**

Cette métrique mesure la fiabilité des détections positives. Elle correspond au rapport entre le nombre de vrais positifs et le total des prédictions positives:

$$\frac{VP}{VP + FP}$$

Où :

- VP (Vrais Positifs): détections correctes (par exemple, lorsque le modèle localise correctement un tronc réel)
- FP (Faux Positifs): détections incorrectes (par exemple, lorsque le modèle identifie à tort un tronc sur une feuille ou une ombre).

#### 3.6.1.3. Rappel ("Recall")

Le rappel mesure la capacité du modèle à identifier tous les objets d'intérêt présents dans l'image. Il se calcule comme suit :

$$\frac{VP}{VP + FN} \quad \text{où :}$$

- FN (Faux Négatifs) : objets non détectés (par exemple, lorsqu'un tronc partiellement masqué ou mal éclairé n'est pas identifié).

#### 3.6.1.4. mAP ("mean Average Precision")

La moyenne de la précision moyenne (mAP) permet d'évaluer la performance globale du modèle en combinant la précision des détections et leur exhaustivité (rappel), tout en tenant compte de la précision de localisation (différents seuils d'IoU) :

- La notation mAP50 fait référence à un seuil d'IoU de 50%( IoU = 0.5 : localisation approximative)
- La notion mAP50:95 est utilisée pour des mAP calculées à différents seuils d'IoU, de 50 % à 95 %, avec un pas de 5 %, reflétant différents niveaux d'exigence.
- Les suffixes **B** et **M** correspondent respectivement à la détection d'objets (bounding boxes) et à la segmentation d'instances (masques).

#### 3.6.1.5. Fonction de perte (Loss)

La fonction de perte quantifie l'écart entre les prédictions du modèle et les annotations réelles (vérité terrain). Elle constitue un indicateur fondamental de la qualité de l'apprentissage. En effet, au cours de l'entraînement et de la validation, plusieurs courbes de perte peuvent être tracées afin d'illustrer l'évolution de cette fonction dans le temps. Une diminution de la perte est généralement le signe que

le modèle apprend efficacement. L'analyse conjointe des courbes de perte d'entraînement et de validation permet de détecter un phénomène de surapprentissage (overfitting), typiquement observé lorsque la perte sur les données d'entraînement continue de décroître alors que celle sur les données de validation augmente.

Pour l'évaluation du modèle YOLOv8 élaboré, les pertes sont définies par plusieurs composantes spécifiques :

- **box\_loss** (perte de localisation) qui mesure l'écart entre les *bounding boxes* prédites et les boîtes réelles.
- **seg\_loss** (perte de segmentation) qui évalue la qualité des masques de segmentation (pixels classés comme appartenant à l'objet détecté) produits par le modèle
- **cls\_loss** (perte de classification) qui mesure les erreurs de classification des objets détectés.
- **dfl\_loss** ("*Distribution Focal Loss*") qui mesure l'incertitude dans la prédiction des coordonnées des boîtes en modélisant ces dernières comme des distributions de probabilité. Elle pénalise davantage les prédictions incertaines ou imprécises.

### 3.6.2. Résultats et discussion

La figure 23 représente l'ensemble de courbes d'évaluation obtenues lors de l'entraînement et la validation du modèle YOLOv8 proposé pour la détection des troncs d'oliviers. L'analyse des différentes métriques permet d'évaluer la progression du modèle au fil des époques.

Tout d'abord, les courbes de pertes (train/box\_loss, train/seg\_loss, train/cls\_loss, train/dfl\_loss) montrent une diminution constante, indiquant que le modèle apprend efficacement à ajuster ses prédictions. Cependant, les pertes de validation (val/box\_loss, val/seg\_loss, val/cls\_loss, val/dfl\_loss) indique la nécessité de faire certains ajustements pour assurer une meilleure généralisation du modèle sans sur-apprentissage.

Les courbes associés aux métriques de précision et de rappel (metrics/precision(B), metrics/recall(B), metrics/precision(M), metrics/recall(M)) montre également une progression avec les époques, justifiant une amélioration de la capacité du modèle à détecter correctement les troncs et à minimiser les faux négatifs et faux positifs.

L'évolution du mAP50 et du mAP50-95 montre une tendance à la hausse, ce qui confirme que le modèle devient progressivement plus précis dans la localisation des objets. Cependant, certaines

fluctuations, notamment dans le rappel et le mAP50-95, peuvent indiquer des variations dans l'apprentissage qui pourraient être atténuées par un ajustement des hyperparamètres comme le taux d'apprentissage ou la taille du batch.

En résumé, ces courbes suggèrent que l'entraînement du modèle est efficace et que sa performance s'améliore progressivement, bien que certaines métriques puissent encore être optimisées pour une meilleure robustesse et précision dans la détection des troncs d'olivier.

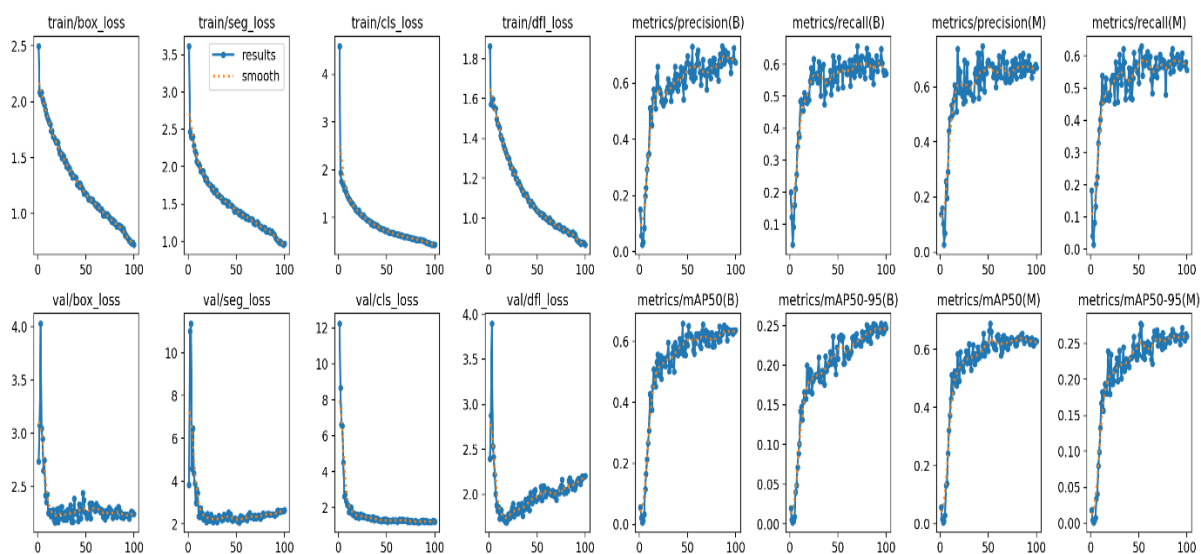


Figure 23 : Courbes d'évaluation de la performance du modèle

Les figures 24 et 25 présentent les résultats obtenus lors du test du modèle YOLOv8 après son réentraînement à l'identification des troncs d'olivier. Les annotations en bleu indiquent les troncs détectés, accompagnés de leurs scores de confiance.



Figure 24 : Résultats de test de détection des troncs d'oliviers avec objets au premier plan



Figure 25 : Test de détection des troncs d'oliviers avec présence d'éléments parasites

Globalement, le modèle parvient à identifier plusieurs troncs présents dans la scène associée à chaque image de test. La figure 24 affiche des scores de confiance compris entre 0.68 et 0.91, tandis que ceux de la figure 25 varient entre 0,61 et 0,85. Les troncs situés au premier plan sont détectés avec une bonne fiabilité, comme en témoigne celui positionné à droite de la figure 24, affichant un score de confiance de 0.91. En revanche, les troncs plus éloignés présentent des scores plus faibles, ce qui peut traduire une sensibilité réduite du modèle face aux objets de petite taille ou partiellement occultés par le feuillage.

Un aspect positif de cette évaluation réside dans la qualité des boîtes englobantes, qui semblent bien ajustées à la forme des troncs, indiquant une bonne capacité de localisation. Toutefois, certaines détections pourraient être affinées, notamment en ajustant le seuil de confiance minimal afin d'améliorer la détection des troncs moins visibles. De plus, la variabilité observée dans les scores de confiance souligne la nécessité de renforcer la cohérence des prédictions.

En conclusion, bien que les résultats obtenus soient prometteurs, des améliorations sont encore possibles pour optimiser la précision et la robustesse du modèle. Cela pourrait passer par l'enrichissement de l'ensemble de données d'entraînement avec des images prises sous différents angles et à différentes distances, ainsi que par l'ajustement de certains hyperparamètres, tels que la résolution des images d'entrée.

#### 4. Retour d'expérience

Au cours de mon stage d'un mois au sein de l'entreprise PELLENC, j'ai pu constater un cadre particulièrement bien organisé pour l'accueil des stagiaires. L'entreprise dispose d'une réelle culture d'intégration, avec des dispositifs établis pour faciliter l'arrivée des nouveaux collaborateurs.

Comme tout nouvel employé, j'ai suivi une formation de sécurité et une visite complète du site. Dès mon arrivée, j'ai bénéficié de l'ensemble des moyens nécessaires pour travailler dans de bonnes conditions : un badge d'accès personnalisé, un poste de travail équipé dans l'open-space aux côtés de mon équipe, un ordinateur préconfiguré, une adresse e-mail professionnelle, ainsi qu'un jeton quotidien offrant une boisson gratuite.

Intégrée au bureau d'études, j'ai été considérée comme une ingénieure à part entière, ce qui a grandement favorisé mon intégration, tant sur le plan professionnel que relationnel. Cette reconnaissance et cette confiance accordées dès le départ ont largement contribué à mon implication et à la qualité de mon expérience au sein de l'entreprise.



Figure 26 : Open-space R&D



Figure 27 : Mon poste de travail

Au-delà des aspects matériels, j'ai particulièrement apprécié la richesse des échanges au sein de l'équipe. En effet, des réunions de suivi régulières ont été organisées tout au long de mon stage, jouant un rôle déterminant dans l'avancement de mes missions. Ces moments d'échange ont constitué pour moi une occasion précieuse de mieux appréhender certains concepts techniques ou méthodologiques

qui, au départ, me semblaient complexes. Lors de ces réunions, j'avais l'opportunité de présenter l'état d'avancement de mon travail, de partager les résultats obtenus et de discuter des difficultés rencontrées. Ces discussions ont toujours été constructives, me permettant de bénéficier des conseils et des orientations de mes collègues, ce qui m'a aidé à surmonter les obstacles et à ajuster ma démarche. Grâce à ces échanges, j'ai pu approfondir ma compréhension des enjeux du projet et progresser de manière significative. Ces réunions ont également renforcé ma capacité à m'exprimer de manière claire et structurée, tout en mettant en évidence l'importance du travail d'équipe et de la communication régulière dans un environnement professionnel.

Ce stage m'a permis d'acquérir un large éventail de compétences. Dès la première semaine, j'ai été impliqué dans des travaux de recherche, ce qui m'a permis d'apprendre à analyser des documents techniques et à en extraire les informations pertinentes pour évaluer des technologies et des modèles. J'ai également renforcé mes compétences en programmation, en découvrant la rigueur requise pour coder dans des environnements complexes. Par ailleurs, le processus d'annotation manuelle, nécessitant précision et minutie, m'a permis de développer une plus grande patience afin de garantir l'exactitude des résultats. J'ai également amélioré ma gestion du temps et de mes priorités, en organisant ma charge de travail de manière optimale tout en restant flexible pour m'adapter aux besoins évolutifs du projet. Cette expérience m'a offert un aperçu précieux du domaine de la recherche et des libertés qu'il offre en termes de conditions de travail. Elle a éveillé en moi un véritable intérêt pour une carrière dans ce secteur, où je pourrais mettre à profit mes compétences techniques et contribuer à des projets innovants et concrets.

Sur le plan relationnel et managérial, j'ai appris à m'intégrer dans l'entreprise et à m'adapter à sa culture. Chez PELLENC, l'usage du tutoiement entre collègues implique une connaissance personnelle des membres de l'équipe, ce qui favorise un environnement de travail informel et collaboratif. Les pauses déjeuner constituaient également des moments privilégiés d'échange, où la convivialité permettait des discussions sur des sujets variés, ce que j'ai particulièrement apprécié. J'ai appris à adopter une conduite professionnelle lors des réunions, en reconnaissant l'importance du travail en équipe, étant régulièrement amené à collaborer avec différents ingénieurs. Cela m'a permis de m'adapter à des personnalités et des expertises diverses, tout en affinant mes compétences interpersonnelles. Enfin, j'ai renforcé mes compétences en communication écrite et orale en rédigeant des rapports techniques et en présentant mes résultats lors de réunions d'équipe. Ces exercices ont été l'occasion de développer ma clarté d'expression et ma confiance dans la transmission de mes idées.

### **III. L'IA et l'agriculture: une révolution entre opportunités et défis**

---

#### **1. Contexte et problématique**

Dans un contexte mondial marqué par des tensions géopolitiques croissantes, des perturbations des chaînes d'approvisionnement et des événements climatiques extrêmes, le secteur agricole est soumis à une pression sans précédent. Ces facteurs fragilisent la résilience des filières agricoles et compromettent les efforts pour éradiquer la faim et l'insécurité alimentaire. La pandémie de COVID-19 a également mis en lumière les vulnérabilités structurelles du secteur agricole, tout en accentuant les inégalités économiques et sociales existantes.

Parallèlement, la croissance démographique mondiale accentue ces enjeux : selon l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), la population mondiale devrait atteindre près de 10 milliards d'individus d'ici 2050, nécessitant une augmentation de 70 % de la production alimentaire pour répondre à la demande croissante.

Dans ce contexte, l'intelligence artificielle (IA) se présente comme un levier technologique majeur pour transformer l'agriculture. Capable d'optimiser les processus de production, de renforcer la gestion des ressources naturelles et d'anticiper les risques climatiques ou sanitaires, l'IA pourrait contribuer à rendre les systèmes agricoles plus durables, efficaces et résilients. Cependant, son adoption soulève également des défis majeurs : fracture technologique, éthique des données, ou encore impact sur l'emploi rural. Dès lors, comment concilier les opportunités offertes par l'IA avec les enjeux socio-économiques et environnementaux qu'elle implique ? En d'autres termes, dans quelle mesure l'IA peut-elle concourir à une agriculture à la fois performante, durable et inclusive ?

#### **2. L'IA au service de l'agriculture**

##### **2.1. Évolutions technologiques du domaine agricole**

L'agriculture a toujours été un secteur en constante évolution, s'adaptant aux besoins des sociétés et aux avancées techniques. Depuis la révolution agricole du XIXe siècle jusqu'à l'ère de la mécanisation au XXe, chaque étape a permis d'améliorer la productivité et d'alléger le travail des agriculteurs.

Aujourd'hui, une nouvelle phase de transformation est en cours, portée par les technologies numériques et l'IA, redéfinissant les pratiques et les perspectives du secteur. En effet, de nombreuses



innovations ont vu le jour dans le domaine agricole. A titre d'exemple, les capteurs connectés permettent de surveiller en temps réel l'état des sols, le niveau d'humidité ou encore les besoins nutritifs des plantes. Parallèlement, les drones, eux, offrent une vue d'ensemble des parcelles et facilitent la détection des maladies ou des zones en stress hydrique. Également, les tracteurs autonomes et les robots agricoles commencent à être déployés, apportant plus de précision et d'efficacité dans les travaux des champs.

Ces évolutions ne se limitent pas à la seule production. Elles impactent également la logistique, la transformation des produits et la commercialisation, avec l'essor de plateformes numériques connectant directement producteurs et consommateurs.

Ainsi, l'agriculture s'engage dans une ère nouvelle, où la technologie agit comme un catalyseur de durabilité, d'efficience et de compétitivité.

## **2.2. Opportunité de l'IA dans l'agriculture**

### **2.2.1. Le rôle de l'IA dans l'optimisation de la production agricole**

L'un des apports les plus significatifs de l'IA dans le domaine agricole réside dans sa capacité à optimiser la production de manière intelligente, durable et personnalisée. Grâce à l'analyse avancée des données et aux capacités prédictives de ses algorithmes, l'IA permet de repenser les pratiques agricoles traditionnelles et d'augmenter significativement les rendements tout en maîtrisant les coûts et en réduisant les impacts environnementaux.

Concrètement, l'IA intervient dès les premières étapes du cycle de production agricole, en optimisant la planification des cultures. En exploitant des données historiques telles que les rendements, les conditions météorologiques, la nature du sol ou encore la rotation des cultures, les systèmes intelligents peuvent générer des recommandations précises sur les variétés à semer, les périodes idéales d'implantation et les parcelles les mieux adaptées. Cette approche permet non seulement une utilisation optimisée des surfaces agricoles, mais aussi une projection plus fiable des résultats, intégrant les contraintes agronomiques et les particularités climatiques locales.

Pendant la phase de croissance des cultures, l'IA joue un rôle essentiel dans le suivi en temps réel des parcelles en utilisant des capteurs connectés qui mesurent l'humidité du sol, la température, la luminosité, ou encore la teneur en nutriments. Ces données sont traitées en continu par des algorithmes capables de détecter des anomalies, comme un stress hydrique ou une carence

nutritionnelle. L'agriculteur peut ainsi intervenir rapidement et de manière ciblée, évitant des pertes potentielles et réduisant le recours excessif aux intrants (engrais, eau, produits phytosanitaires).

En parallèle, l'intégration de l'imagerie satellitaire et des drones avec des modèles prédictifs avancés permet d'établir des estimations précises des rendements agricoles plusieurs semaines avant la récolte. Ces données prospectives offrent une valeur stratégique considérable en permettant d'optimiser la planification logistique, d'ajuster les stratégies commerciales et d'anticiper les besoins en personnel saisonnier. En fournissant une visibilité accrue sur la production à venir, ces technologies favorisent une gestion plus efficiente des ressources tout en renforçant la rentabilité des exploitations.

L'optimisation de la production ne s'arrête pas aux champs. Dans les élevages, l'IA permet également de suivre la santé et la productivité des animaux en temps réel, à travers des capteurs mesurant leur activité, leur alimentation, ou leur comportement. Ces informations permettent d'agir de manière préventive et d'améliorer les performances globales de l'élevage.

En résumé, l'IA offre aux agriculteurs une vision fine, globale et dynamique de leur exploitation, leur permettant d'agir au bon moment, au bon endroit, et de la bonne manière. Elle favorise une agriculture plus efficace, rentable et résiliente, mieux adaptée aux aléas climatiques et aux exigences de durabilité.

### **2.2.2. L'agriculture de précision et la gestion durable des ressources grâce à l'IA**

L'IA joue un rôle fondamental dans le développement de l'agriculture de précision, une approche qui vise à optimiser l'utilisation des ressources agricoles en intervenant de manière ciblée, parcelle par parcelle, voire plante par plante. Grâce à l'IA, cette vision devient de plus en plus concrète, en rendant possible une gestion fine, réactive et durable de l'eau, des engrais, des pesticides et de l'énergie.

L'agriculture de précision repose sur la collecte de données en temps réel via un ensemble de capteurs connectés, de systèmes GPS, de drones, de caméras multispectrales et d'images satellites. Ces données sont ensuite analysées par des algorithmes d'intelligence artificielle, capables d'identifier des patterns, d'anticiper des besoins ou des risques, et de fournir des recommandations personnalisées à l'exploitant.

L'un des axes majeurs d'application de l'IA dans ce domaine est la gestion intelligente de l'irrigation. Grâce à l'analyse combinée des données météorologiques, de l'humidité du sol et des besoins hydriques spécifiques à chaque culture, les systèmes d'irrigation pilotés par l'IA peuvent ajuster les apports en eau de manière précise. Cela permet d'éviter le gaspillage, de préserver les ressources en

eau, particulièrement critiques dans certaines régions, et de maintenir des conditions de croissance optimales. Des études ont montré que ces systèmes peuvent permettre jusqu'à 30 % d'économie d'eau, tout en maintenant, voire en augmentant, les rendements.

La même logique s'applique à l'épandage d'engrais et de produits phytosanitaires. L'IA permet de déterminer les zones qui nécessitent un traitement spécifique, en évitant les applications uniformes sur l'ensemble des champs. En limitant les apports aux seules zones nécessaires, cette stratégie réduit les coûts, limite la pollution des sols et des nappes phréatiques, et diminue la résistance des parasites et des mauvaises herbes aux traitements.

Un autre apport notable concerne la cartographie des sols. En croisant les données issues de prélèvements physiques avec des analyses géospatiales et des données historiques, les systèmes intelligents peuvent produire des cartes détaillées de la fertilité, de l'acidité ou de la texture des sols. Cela permet à l'agriculteur de mieux adapter ses pratiques culturales et de choisir les variétés les mieux adaptées à chaque zone de son exploitation.

De plus, cette approche localisée favorise une réduction de l'empreinte carbone des exploitations agricoles. En évitant les surutilisations d'intrants chimiques et en limitant les passages d'engins agricoles, on réduit les émissions de gaz à effet de serre. À plus grande échelle, l'IA pourrait ainsi jouer un rôle décisif dans la transition vers une agriculture plus durable, compatible avec les objectifs de développement durable (ODD) fixés par les Nations Unies.

En somme, l'IA renforce les fondements de l'agriculture de précision en la rendant plus accessible, plus efficace et plus durable. Elle permet une gestion raisonnée des ressources, alliant performance économique et respect de l'environnement, dans une logique d'anticipation plutôt que de réaction.

#### **2.2.3. Automatisation et robotisation des tâches agricoles grâce à l'IA**

L'un des apports les plus spectaculaires de l'IA dans le secteur agricole réside dans la robotisation et l'automatisation des tâches. Grâce à la combinaison de machines autonomes et d'algorithmes intelligents, de nombreuses opérations agricoles autrefois manuelles ou très exigeantes peuvent aujourd'hui être réalisées de façon automatique, avec une précision et une efficacité accrue. Cette révolution technologique redéfinit profondément l'organisation du travail agricole et ouvre la voie à une productivité renouvelée.

Parmi les applications les plus courantes, on trouve les tracteurs autonomes, capables de semer, de fertiliser ou de labourer un champ sans intervention humaine. Ces véhicules sont équipés de capteurs,

de caméras, de GPS et de logiciels de pilotage automatique alimentés par l'IA, qui leur permettent de suivre des trajets précis, d'éviter les obstacles et de s'adapter aux conditions du terrain en temps réel. L'automatisation de ces tâches répétitives réduit considérablement la pénibilité du travail, tout en augmentant la régularité et la précision des opérations.

Dans le même esprit, les robots de désherbage intelligents représentent une alternative prometteuse à l'utilisation massive d'herbicides. Grâce à la reconnaissance d'image, ces robots peuvent identifier les mauvaises herbes au milieu des cultures, puis les éliminer mécaniquement ou de manière ciblée avec un microdosage de produit. Ce type de solution permet non seulement de réduire la consommation de produits chimiques, mais aussi de préserver la santé des agriculteurs et des écosystèmes environnants.

L'IA est également utilisée dans la récolte automatisée, un domaine en plein essor. Des robots cueilleurs, équipés de bras articulés et de capteurs, sont aujourd'hui capables de récolter certains fruits ou légumes (comme les pommes, fraises, tomates) en reconnaissant leur degré de maturité et en adaptant leur geste pour ne pas les abîmer. Ces machines peuvent travailler jour et nuit, ce qui représente un atout majeur face à la pénurie de main-d'œuvre saisonnière, un problème récurrent dans de nombreux pays.

Dans les élevages, l'automatisation a aussi fait ses preuves. Les robots de traite, par exemple, permettent de traire les vaches de manière autonome tout en collectant des données sur leur état de santé, leur alimentation ou leur rendement lacté. Des systèmes d'alimentation automatisée, associés à des algorithmes d'optimisation, ajustent la ration de chaque animal en fonction de ses besoins spécifiques.

Au-delà du gain de productivité, l'automatisation contribue à réduire les marges d'erreur humaines, à améliorer la traçabilité des opérations, et à optimiser l'utilisation des ressources. Elle permet aussi aux agriculteurs de se libérer de certaines tâches contraignantes pour se concentrer sur des activités à plus forte valeur ajoutée, comme la gestion stratégique de l'exploitation.

Dès lors, l'automatisation rendue possible par l'IA transforme profondément les pratiques agricoles. En améliorant la productivité, en réduisant les charges de travail et en favorisant une agriculture plus précise et durable, elle représente une véritable opportunité d'avenir pour le secteur.

### **3. Les limites et défis de l'IA dans l'agriculture**

#### **3.1. Une fracture numérique au détriment des petits exploitants**

L'introduction de l'IA dans l'agriculture constitue une avancée majeure, mais elle soulève également un enjeu fondamental de fracture numérique. Si l'IA offre des gains de productivité et une gestion plus efficace des ressources, son accès reste inégalement réparti, notamment entre grandes exploitations disposant de moyens importants et petites fermes, souvent situées dans des zones rurales ou isolées, en particulier dans les pays en développement.

Cette fracture s'exprime d'abord par une disparité dans l'accès aux infrastructures numériques de base. De nombreuses zones rurales dans le monde, y compris en Europe ou en Afrique du Nord, souffrent encore d'une couverture Internet insuffisante, d'un réseau mobile instable ou d'un manque d'accès à l'électricité. Pourtant, le déploiement des solutions d'intelligence artificielle est intrinsèquement dépendant d'une connectivité robuste : qu'il s'agisse des capteurs IoT, des plateformes cloud, des systèmes de télédétection ou des outils d'analyse de données, tous ces dispositifs requièrent impérativement une connexion haut débit stable.

Au-delà des infrastructures, l'adoption de l'IA est également freinée par le coût des technologies. L'acquisition de capteurs, de drones, de logiciels intelligents ou de robots agricoles représente un investissement souvent hors de portée pour les petites exploitations ou les agriculteurs indépendants. En l'absence d'aides publiques, de mutualisation ou de subventions ciblées, ces outils restent l'apanage des grandes entreprises agroalimentaires ou de coopératives bien financées. Ce déséquilibre accentue la concentration des moyens de production et pourrait aggraver les inégalités économiques dans le secteur agricole.

Par ailleurs, le manque de compétences numériques constitue un autre obstacle majeur. L'usage de technologies d'intelligence artificielle suppose un certain niveau de formation technique, voire de connaissances en analyse de données. De nombreux agriculteurs, notamment les plus âgés ou les moins alphabétisés, peuvent se retrouver exclus de cette transformation faute de formation adaptée. Ce fossé générationnel et éducatif peut ralentir l'adoption de solutions technologiques pourtant utiles.

Enfin, cette fracture numérique pose des questions plus larges de justice sociale et territoriale. Si seuls certains exploitants peuvent profiter des bénéfices de l'IA (meilleurs rendements, réduction des coûts, anticipation des risques), alors les inégalités structurelles entre agricultures intensives et agricultures familiales risquent de se creuser. Dans les régions du Sud, par exemple, où l'agriculture reste le

principal moyen de subsistance, le retard d'accès à l'innovation pourrait compromettre la sécurité alimentaire et le développement durable à long terme.

En somme, si l'IA représente un levier puissant pour l'avenir de l'agriculture, son potentiel ne pourra être pleinement exploité que si son accès est rendu équitable et inclusif, au bénéfice de tous les acteurs du monde agricole, quels que soient leur taille ou leur localisation.

#### **3.2. Dépendance aux données**

Le potentiel de l'intelligence artificielle en agriculture est étroitement lié à l'accessibilité et à la fiabilité des données exploitées. En effet, les algorithmes d'IA nécessitent des jeux de données volumineux et variés pour fonctionner de manière optimale : paramètres météorologiques, analyses pédologiques, archives de rendements, indicateurs sanitaires des cultures ou des cheptels, imagerie satellitaire, entre autres. Cette exigence en matière de collecte et de traitement de l'information représente l'un des principaux enjeux du secteur agricole dans sa transition numérique, soulignant la nécessité de mettre en place des infrastructures adaptées et des protocoles rigoureux pour garantir la pertinence des modèles prédictifs.

Le premier défi réside dans l'hétérogénéité marquée des données disponibles, qui varie considérablement selon les zones géographiques, les types d'exploitations et les filières agricoles. Une proportion significative de petites et moyennes exploitations ne possède pas l'équipement technologique adéquat (capteurs connectés, stations météorologiques automatisées, drones de surveillance) pour collecter systématiquement les données nécessaires. Par ailleurs, les méthodes de collecte manuelle, encore largement utilisées, introduisent des risques d'erreurs et d'inconsistances dans les jeux de données. Cette situation engendre des bases de données souvent fragmentaires, peu homogènes et parfois biaisées, limitant ainsi la fiabilité et la performance des modèles d'intelligence artificielle qui en dépendent. Cette disparité technologique crée ainsi un déséquilibre dans la capacité des différents acteurs à bénéficier pleinement des avantages de l'agriculture numérique.

Au-delà des enjeux techniques, la question éthique de la gouvernance des données agricoles représente un défi majeur. La problématique centrale réside dans la répartition des droits d'accès et d'exploitation des données générées par les exploitations : doivent-elles relever de l'agriculteur, des fabricants de capteurs ou des éditeurs de solutions logicielles ? Dans la pratique, ces informations sont fréquemment hébergées sur des infrastructures cloud détenues par des acteurs privés internationaux, qui peuvent en exploiter la valeur commerciale sans nécessairement garantir une

transparence absolue sur leurs usages. Cette situation crée une asymétrie préoccupante, mettant en lumière des risques évidents pour la souveraineté numérique, tant au niveau individuel (protection des exploitants) qu'au niveau national (autonomie stratégique des États), particulièrement dans les économies émergentes.

En conclusion, bien que les données constituent le fondement indispensable au développement de l'IA en agriculture, elles en représentent également la principale vulnérabilité. La mise en place d'un cadre de gouvernance rigoureux - garantissant à la fois l'intégrité, l'équité et la sécurité des données - s'avère essentielle pour assurer la pertinence opérationnelle, l'efficacité technologique et l'acceptabilité sociétale des solutions d'IA dans le secteur agricole.

### **3.3. Des risques pour l'emploi et les savoir-faire traditionnels**

L'introduction croissante de l'IA dans les systèmes agricoles transforme profondément les pratiques et les modèles de production. Si cette mutation technologique offre de nombreux avantages — gain de productivité, précision des interventions, réduction des coûts — elle n'est pas sans conséquences sur le marché de l'emploi rural et sur la transmission des savoir-faire agricoles traditionnels.

L'un des principaux risques soulevés par l'automatisation et l'IA réside dans la substitution progressive du travail humain par des machines intelligentes. Les tâches les plus répétitives, pénibles ou à faible valeur ajoutée — comme la récolte, le désherbage, l'irrigation ou la surveillance des cultures — sont aujourd'hui de plus en plus assurées par des robots, des capteurs et des logiciels pilotés par IA. Cette évolution peut conduire, dans certains cas, à une réduction du besoin de main-d'œuvre, notamment saisonnière ou peu qualifiée. Dans les territoires où l'agriculture représente une source d'emploi majeure, cette dynamique risque d'aggraver la précarité et le chômage rural.

Par ailleurs, l'usage massif de technologies intelligentes peut engendrer une transformation profonde des compétences requises dans le secteur agricole. Les agriculteurs sont désormais confrontés à la nécessité de comprendre, manipuler et entretenir des outils numériques complexes, qui relèvent souvent du domaine de l'informatique, de l'électronique ou de la modélisation algorithmique. Or, tous ne disposent pas des ressources ou du temps nécessaire pour se former à ces nouveaux outils. Ce décalage risque d'exclure une partie de la population agricole de la transition technologique, créant ainsi une forme de fracture générationnelle ou sociale.

En parallèle, l'omniprésence de l'IA peut contribuer à une érosion progressive des savoir-faire traditionnels, transmis de génération en génération. Les gestes ancestraux, l'observation fine des cycles naturels, les connaissances empiriques sur les sols ou les cultures peuvent être perçus comme



dépassés face à l'efficacité des algorithmes. Pourtant, ces savoirs jouent un rôle clé dans la résilience et la durabilité des systèmes agricoles, notamment en agriculture familiale ou agroécologique. Leur disparition pourrait appauvrir la diversité des pratiques et la capacité d'adaptation du monde agricole.

Ce constat soulève également des enjeux identitaires et culturels : l'agriculture ne se réduit pas à un ensemble de tâches techniques à automatiser. Elle constitue un mode de vie, un héritage culturel, un lien étroit avec la nature et les territoires. En transformant la relation entre l'humain, la terre et les cycles biologiques, l'IA redéfinit le sens même du métier d'agriculteur.

En somme, le développement de l'IA en agriculture ne doit pas se faire au détriment du tissu humain, social et culturel qui fonde l'agriculture depuis des millénaires. Il s'agit de bâtir une modernisation inclusive et respectueuse des équilibres humains, en conjuguant innovation technologique et préservation des savoirs vivants.

#### **4. Vision future de l'Agriculture Intelligente**

##### **4.1. Vers une collaboration homme-machine pour une agriculture plus durable**

L'avenir de l'agriculture intelligente ne se joue pas dans une opposition entre l'être humain et la machine, mais dans la capacité à créer une collaboration équilibrée, éthique et durable entre les deux. L'IA, loin de remplacer totalement le travail humain, pourrait devenir un partenaire stratégique pour les agriculteurs, en les accompagnant dans leurs décisions, en réduisant leur charge de travail, et en les aidant à cultiver plus efficacement, avec un impact environnemental limité.

Cette vision repose sur un principe fondamental : l'IA comme outil d'aide à la décision, et non comme substitut à l'intelligence humaine ou au savoir agricole. Les systèmes intelligents peuvent analyser des données complexes à grande échelle (climat, état des sols, maladies émergentes, besoins nutritionnels des plantes...), mais c'est encore à l'humain de juger, d'interpréter et d'agir. Cette complémentarité permet de combiner la rigueur algorithmique et l'intuition humaine, souvent ancrée dans l'expérience du terrain. L'IA devient alors une forme d'extension de la perception humaine, capable d'alerter, de prévoir, d'optimiser, tout en laissant à l'agriculteur le contrôle final.

Dans cette logique de collaboration, l'IA peut aussi aider à renforcer les principes d'une agriculture plus durable. En fournissant des données précises et localisées, elle permet une gestion plus fine des ressources naturelles : eau, engrais, pesticides, énergie. Elle évite les gaspillages, limite les interventions inutiles et contribue à réduire l'empreinte écologique des exploitations agricoles. L'IA peut également être mise au service de modèles agroécologiques ou biologiques, en soutenant une

observation continue de la biodiversité, des rotations de cultures, ou encore des interactions entre plantes et insectes.

Mais cette collaboration homme-machine suppose un changement culturel et structurel. Elle nécessite de repenser la formation agricole, en intégrant les compétences numériques et la compréhension des systèmes intelligents dans les cursus. Elle suppose également un effort d'accessibilité technologique pour que les outils d'IA soient adaptés aux réalités diverses des exploitations, en particulier pour les petites et moyennes structures. La technologie ne doit pas être un facteur d'exclusion, mais au contraire, un levier d'inclusion pour une agriculture plus équitable.

En somme, l'agriculture du futur ne sera pas seulement technologique : elle sera humaine et augmentée. L'IA ne remplace pas la main, ni le regard, ni le bon sens du paysan. Mais elle peut, si elle est bien pensée, accompagner une transition vers un modèle plus sobre, plus intelligent, et plus durable, où l'humain garde le cap, et la machine éclaire le chemin.

#### **4.2. Encadrer l'IA pour une agriculture éthique, équitable et régulée**

Alors que l'IA s'impose progressivement dans le monde agricole, se pose une question essentielle : comment encadrer son développement pour qu'il serve réellement l'intérêt général ? Car si les technologies intelligentes promettent des avancées considérables en matière de productivité, de durabilité et d'anticipation des risques, elles peuvent aussi, en l'absence de régulation, accentuer les inégalités, concentrer le pouvoir entre les mains de quelques acteurs technologiques, et menacer la souveraineté alimentaire des États.

L'un des premiers enjeux majeurs est celui de la gouvernance des données agricoles. L'IA repose sur des quantités massives d'informations — sur les sols, les cultures, le climat, les rendements, les pratiques agricoles — souvent collectées à l'insu des agriculteurs, via capteurs, satellites ou plateformes numériques. Or, ces données sont une ressource stratégique : elles permettent d'alimenter les algorithmes d'optimisation, mais aussi de générer de la valeur économique. Il devient donc impératif de garantir la transparence dans la collecte, l'usage et la propriété des données, afin d'éviter leur captation exclusive par de grandes firmes agro-tech ou plateformes de services. La création de communs numériques agricoles ou de coopératives de données pourrait permettre aux agriculteurs de garder la maîtrise de leurs informations, tout en bénéficiant collectivement des avancées technologiques.

Un autre défi majeur concerne la concentration du pouvoir technologique. À l'heure actuelle, le marché des solutions d'IA en agriculture est largement dominé par quelques géants du numérique qui

proposent des outils intégrés mêlant semences, intrants, capteurs, services de conseil et plateformes d'analyse. Cette logique d'intégration verticale peut enfermer les agriculteurs dans des systèmes fermés, coûteux et dépendants, les privant de liberté de choix et de marges de manœuvre. Il est donc nécessaire de favoriser la pluralité des acteurs (startups locales, institutions publiques, coopératives), l'interopérabilité des technologies, et de soutenir les innovations ouvertes, accessibles et modulables.

Sur le plan éthique, l'introduction de l'IA dans l'agriculture soulève également des interrogations sur les valeurs que nous voulons y inscrire. Faut-il prioriser l'efficacité à tout prix, ou veiller à préserver la diversité des modèles agricoles, les équilibres sociaux, et les liens humains ? La technologie ne doit pas être pensée comme neutre : elle est porteuse d'une vision du monde. Il est donc crucial d'impliquer dans sa conception et son déploiement l'ensemble des parties prenantes : agriculteurs, chercheurs, citoyens, décideurs politiques. Des comités éthiques agri-tech, ou des processus de démocratie technologique, pourraient voir le jour pour éclairer les choix de société liés à ces innovations.

Enfin, un cadre de régulation national et international devra accompagner cette transformation. L'Union européenne, par exemple, s'engage déjà sur une IA de confiance, fondée sur le respect des droits fondamentaux et la non-discrimination. Ce type de cadre doit être adapté aux spécificités du monde agricole, en garantissant l'équité d'accès aux technologies, la sécurité des systèmes automatisés, et la protection des exploitants face aux dérives potentielles.

En somme, encadrer l'IA dans l'agriculture, c'est refuser de subir l'innovation pour mieux la réorienter vers des objectifs sociaux, environnementaux et démocratiques. C'est faire le choix d'une technologie au service de l'humain, et non l'inverse.

### **4.3. Scénarios pour l'agriculture de demain**

Imaginer l'agriculture de demain, c'est réfléchir aux trajectoires possibles qu'elle pourrait emprunter sous l'influence croissante de l'IA, du changement climatique, de la transition écologique, et des bouleversements géopolitiques. Loin d'un avenir figé ou linéaire, plusieurs scénarios contrastés se dessinent, chacun porteur de visions du monde, de priorités économiques, sociales et environnementales différentes. Cette multiplicité de futurs potentiels invite à une réflexion critique sur les choix à faire aujourd'hui pour façonner une agriculture à la fois performante, juste et résiliente.

Le premier scénario, que l'on pourrait qualifier de technologique et productiviste, s'appuie sur une généralisation massive des outils d'IA et de robotique pour maximiser les rendements. L'exploitation

agricole devient une “usine verte” hyperconnectée, optimisée en temps réel, où la machine joue un rôle central dans la prise de décision et l'exécution. Ce modèle peut répondre efficacement aux défis de la sécurité alimentaire mondiale, mais il soulève des questions cruciales : quelle place reste-t-il pour l'humain, la biodiversité, la souveraineté alimentaire ? Et quels risques de dépendance accrue vis-à-vis des grandes entreprises technologiques ?

Un deuxième scénario cherche à réconcilier haute technologie et durabilité, dans une logique de transition équilibrée. Il s'appuie sur l'IA pour améliorer l'efficacité des pratiques agricoles tout en intégrant des impératifs environnementaux, sociaux et éthiques. L'innovation y est pensée comme un levier de transformation vers des modèles hybrides : fermes verticales en milieu urbain, agriculture circulaire, intégration paysagère, gestion adaptative des ressources. Ce scénario repose toutefois sur une volonté politique forte, des choix collectifs concertés et un effort de formation pour accompagner les agriculteurs dans ce changement.

Enfin, un scénario plus critique envisage une fragmentation du monde agricole, où les inégalités d'accès à la technologie creusent un fossé entre différents modèles. D'un côté, des exploitations à la pointe de l'IA, hautement performantes mais concentrées et déshumanisées ; de l'autre, des agricultures exclues du numérique, fragilisées, voire marginalisées. Ce scénario souligne l'urgence d'un encadrement démocratique de l'innovation agricole, pour éviter une agriculture à deux vitesses.

Face à ces scénarios, aucun futur n'est écrit d'avance. L'agriculture de demain dépendra des valeurs que la société choisira de mettre au cœur de son développement technologique : performance ou résilience, centralisation ou autonomie, profit ou équité. L'IA, en tant que catalyseur de ces transformations, pose une question clé : serons-nous capables de la mettre au service d'une vision collective et responsable du monde agricole, ou la laisserons-nous façonner l'avenir à notre place ?

## Conclusion

---

Ce stage réalisé au sein de l'équipe R&T à PELLENC a été bien plus qu'une simple expérience professionnelle, il représente une véritable étape dans mon parcours personnel et académique. En effet, en seulement quelques semaines, j'ai eu l'opportunité de découvrir de l'intérieur le fonctionnement d'une entreprise innovante, et de contribuer à des projets concrets, en lien direct avec le domaine de l'intelligence artificielle.

Ce qui m'a le plus marquée durant ce stage, c'est la richesse des apprentissages, à la fois techniques et relationnels. J'ai pu me confronter à des problématiques réelles, apprendre à trouver des solutions de manière autonome, tout en bénéficiant des conseils et de l'expertise des ingénieurs qui m'ont encadrée. J'ai ainsi pu développer ma capacité d'analyse, ma rigueur, et surtout ma curiosité. J'ai également appris à mieux m'organiser, à gérer mon temps et à travailler en équipe dans un environnement professionnel exigeant.

Cette immersion m'a également permis de mieux me connaître. J'ai pris confiance en mes capacités, et j'ai réalisé à quel point le domaine de l'intelligence artificielle embarquée me passionne. Voir concrètement comment l'IA peut améliorer les performances des machines agricoles, optimiser les processus, et contribuer à une agriculture plus précise et plus durable a été très inspirant. Ce stage a donc joué un rôle clé dans la construction de mon projet professionnel, il m'a confirmé que je souhaite devenir ingénieure spécialisée dans ce domaine. Ainsi, pour mon prochain stage de deuxième année, je souhaite m'orienter plus spécifiquement vers les systèmes embarqués, afin d'approfondir mes compétences techniques et de continuer à m'investir dans des projets concrets, innovants et porteurs de sens.

Même si ce mois de janvier est passé très vite, il restera une expérience forte et fondatrice. Je repars de cette aventure avec de nouvelles compétences, une vision plus claire de mon avenir, et une motivation renforcée pour poursuivre mes études avec ambition et détermination.

# Glossaire

---

**Management directif :** C'est un style où le manager s'intéresse aux processus, aux résultats. Il donne des consignes, suit la progression vers des objectifs précis, définit des directives régulièrement. Il informe plus qu'il ne communique et prend seul les décisions.

**La méthode agile :** Un ensemble de pratiques de gestion de projet qui consiste à développer un produit par petites étapes, en ajustant régulièrement les priorités et en favorisant la communication continue entre les équipes et les parties prenantes.

**POC (Proof of Concept) :** C'est une démonstration ou un prototype qui permet de valider la faisabilité d'une idée, d'une technologie ou d'une solution avant de la développer à grande échelle.

**Dataset :** Un ensemble de données structurées utilisées pour entraîner, tester ou valider des modèles d'IA.

**L'apprentissage automatique (Machine Learning) :** C'est une branche de l'intelligence artificielle qui permet à une machine d'apprendre à partir de données pour faire des prédictions ou prendre des décisions sans être explicitement programmée pour chaque tâche.

**La vision par ordinateur (Computer Vision) :** C'est un domaine spécifique de l'IA qui utilise le ML pour permettre aux machines d'interpréter et d'analyser des images ou des vidéos de la même manière que le font les êtres humains.

**L'apprentissage profond (Deep Learning) :** C'est une branche du machine learning qui utilise des réseaux de neurones artificiels très, inspirés du cerveau humain, pour apprendre à partir de grandes quantités de données.

**Couche de convolution :** Sert à extraire automatiquement des caractéristiques importantes d'une image, en appliquant un filtre (matrice carré). Ce filtre balaye l'image et capte des motifs visuels en effectuant des opérations mathématiques simples. Le résultat est une carte de caractéristiques qui met en valeur certaines parties de l'image.

**Réseaux de neurones convolutifs :** Réseaux de neurones formés de couches de convolution afin de traiter les données fournies en entrée.

**Pooling** : Consiste à réduire la taille des cartes de caractéristiques en conservant l'information la plus significative, généralement en prenant la valeur maximale dans de petites zones (max pooling).

**Les boîtes englobantes (bounding boxes)** : Ce sont des rectangles utilisés en vision par ordinateur pour délimiter et localiser un objet dans une image.

**Docker** : C'est une plateforme qui permet de conteneuriser des applications, c'est-à-dire de les emballer avec tout ce dont elles ont besoin pour fonctionner, afin qu'elles puissent être exécutées de manière cohérente sur n'importe quel système.

**La segmentation d'instance** : Elle permet de distinguer et délimiter chaque objet individuel dans une image, même s'ils appartiennent à la même catégorie (par exemple, séparer chaque personne dans une foule).

**L'augmentation de données** : C'est une technique qui consiste à modifier les données existantes (par exemple, en les faisant pivoter, les redimensionnant, ou en ajoutant du bruit) pour créer de nouvelles variations et ainsi améliorer la robustesse du modèle d'apprentissage automatique.

**Epoch** : Représente le nombre de fois où l'ensemble des données d'entraînement est entièrement parcouru par le modèle.

**Batch** : Définit le nombre d'images traitées simultanément avant que le modèle ne mette à jour ses poids.

**Métrique d'évaluation** : C'est une mesure utilisée pour quantifier la performance d'un modèle d'apprentissage automatique. Elle permet de déterminer à quel point le modèle est précis, efficace ou adapté à la tâche, en fonction de critères spécifiques.



# Bibliographie

---

## I. Présentation de l'entreprise d'accueil

### 1. Présentation de l'entreprise d'accueil

[1] Site officiel de PELLENC : <https://www.pellenc.com>

[2] Matériel Agricole.Info. (2021). Parts du marché des machines à vendanger:  
<https://www.materielagricole.info/indicateurs-marches/article/731484/machines-a-vendanger-automotrices-new-holland-reste-en-tete-du-podium-de-2021>

### 2. Organisation du siège social à Pertuis

#### 2.2. Mode managérial

[3] Management et ingénieur, cycle ISMIN – Management durable : l'essentiel.

F. Jaujard, T. Ricordeau – L'essentiel du management, Tome 1 (2024)

## II. Missions et réalisations

### 1. Fondements théoriques

[4] Margaret A. Boden. L'intelligence artificielle. EDP Sciences, 2021. ISBN: 9782759825806

[5] Introduction à l'apprentissage automatique - éducol

<https://eduscol.education.fr/sti/sites/eduscol.education.fr/sti/files/ressources/pedagogiques/14512/14512-introduction-lapprentissage-automatique-ensps.pdf>

[6] Types de techniques d'apprentissage de l'IA utilisées dans la vision par ordinateur

<https://www.ultralytics.com/fr/blog/types-of-ai-learning-techniques-used-in-computer-vision>

[7] Convolutional Neural Network : Tout ce qu'il y a à savoir

Article détaillant les CNN, leur architecture et leurs applications.

<https://datascientest.com/convolutional-neural-network>

[8] Ultralytics YOLOv5 Architecture - Ultralytics YOLO Docs

Description approfondie de l'architecture de YOLOv5, des stratégies d'augmentation des données et des méthodologies d'entraînement.

[https://docs.ultralytics.com/fr/yolov5/tutorials/architecture\\_description/Home](https://docs.ultralytics.com/fr/yolov5/tutorials/architecture_description/Home)

### **III. L'IA et l'agriculture : une révolution entre opportunités et défis**

#### **1. Contexte et problématique**

[9] Académie d'Agriculture de France. (2025). L'intelligence artificielle (IA) est-elle devenue indispensable en agriculture ? [PDF]. [https://www.academie-agriculture.fr/sites/default/files/publications/encyclopedie/intelligence\\_artificielle\\_agriculture.pdf](https://www.academie-agriculture.fr/sites/default/files/publications/encyclopedie/intelligence_artificielle_agriculture.pdf)

#### **2. L'IA au service de l'agriculture**

[10] Inria & INRAE. (2022). *Livre blanc : Agriculture et numérique*. [https://www.inria.fr/sites/default/files/2022-02/livre-blanc-agriculture-numerique-2022\\_INRIA\\_BD.pdf](https://www.inria.fr/sites/default/files/2022-02/livre-blanc-agriculture-numerique-2022_INRIA_BD.pdf)

[11] ministère de l'Agriculture et de la Souveraineté alimentaire. (2022). *Feuille de route Agriculture et numérique*. <https://agriculture.gouv.fr/telecharger/129515>

#### **3. Les limites et défis de l'IA dans l'agriculture**

[12] Organisation des Nations Unies pour l'alimentation et l'agriculture (2020). L'agriculture numérique pour les petits exploitants. <https://www.fao.org/3/ca4887en/ca4887en.pdf>

# Annexes

---

## Annexe 1 : Script filtrage du dataset

```
import os
import shutil
import yaml
import copy
import json

!pip install roboflow

from roboflow import Roboflow
rf = Roboflow(api_key="5jXjahU4H0jPFMAIbync")
project = rf.workspace("stage-ykecb").project("tronc-p3")
version = project.version(7)
dataset = version.download("yolov8")

def filtre(cc: list, bg: list, pm: list, source: str, destination: str, isTest=False):
    #cc est une liste de liste indiquant les classes que l'on veut remplacer ainsi que leur remplaçant
    #bg est la liste des classes qui deviennent nulles, 'background'
    #pm est la liste des classes dont on veut voir l'effet sur l'entraînement
    #source est le repertoire source
    #destination est le repertoire de destination
    #isTest est un booléen égal à True uniquement si on le précise, utile pour ne pas supprimer d'images de test contenant les classes incluses dans pm

    print("source name = ", source)
    label_dir = source + "/labels"
    image_dir = source + "/images"

    for filename in os.listdir(label_dir): # boucle sur les fichiers de source/labels
        rm_file_flag = False # booléen indiquant si on va supprimer le fichier
        f = os.path.join(label_dir, filename)
        nom_fich = filename.split('/')[-1]
        file_root_list = nom_fich.split('.')
        file_root = ""
        for i in range(len(file_root_list)-1):
            if i == 0:
                file_root = file_root_list[i]
            else:
                file_root = file_root + '.' + file_root_list[i] #pour avoir le nom du fichier sans ".txt", utile pour la copie de l'image

        if os.path.isfile(f) and nom_fich.endswith('.txt'):
            nouv_fichier = destination + '/labels/' + nom_fich
            file = open(f, 'r')
            contenu = file.readlines() #liste des lignes du fichier d'origine
            file.close()
            new_file = open(nouv_fichier, 'w') #création et ouverture en écriture du nouveau fichier

            for ligne in contenu: # boucle sur les lignes du fichier d'origine
                ind = int(ligne[0]) # classe de la ligne
                new_ligne = ''
                if ind not in bg:
                    for c in cc:
                        if ind == c[0]:
                            new_ligne = str(c[1]) + ligne[1:] # nouvelle ligne avec le changement de classe
                            new_file.write(new_ligne) # écriture de la ligne dans le nouveau fichier
                            new_file.flush()
                            break
                        #print(new_ligne)
                    ligne = new_ligne # remplacement de l'ancienne ligne par la nouvelle
```

```

        if ind in pm and not isTest:
            new_file.close()
            os.remove(nouv_fichier) # suppression du fichier car il y a le parametre que l'on veut enlever et ce n'est pas le set de test
            rm_file_flag = True
            break

    if not rm_file_flag: # si on n'a pas supprimer le fichier
        image_file = file_root + ".jpg"
        new_image_path = os.path.join(destination, 'images', image_file)
        image_path = os.path.join(image_dir, image_file)
        shutil.copyfile(image_path, new_image_path) #copie de l'image dans destination/images
    new_file.close()

```

```

def filter_all(cc: list, bg: list, pm: list, source:str, destination: str):
    test_source = os.path.join(source, 'test')
    valid_source = os.path.join(source, 'valid')
    train_source = os.path.join(source, 'train')
    test_destination = os.path.join(destination, 'test')
    valid_destination = os.path.join(destination, 'valid')
    train_destination = os.path.join(destination, 'train') #creation des repertoires de test,valid, et train
    filtre2(cc,bg, pm , valid_source, valid_destination) # applications de la fonction précédente
    filtre2(cc,bg, pm , train_source, train_destination)
    filtre2(cc,bg, pm , test_source, test_destination, isTest=True)
    source_yaml = source + "/data.yaml"
    destination_yaml = destination + "/data.yaml"
    with open(source_yaml, 'r') as file: #ouverture du fichier yaml et fermeture 1 ligne apres
        file = yaml.safe_load(file)

    file_copy = copy.copy(file) #copie en valeurs sinon modification du fichier d'origine
    noms = copy.copy(file["names"]) # copie de la liste des classes
    print(noms)
    for c in cc:
        noms[c[0]] = noms[c[1]] #changement de la copie de la liste des classes

    noms = list(set(noms)) #enleve les redondances dans la list noms
    new_nc = len(noms)
    file_copy["names"] = noms
    file_copy["nc"] = new_nc
    if "roboflow" in file_copy:
        del file_copy["roboflow"]
    with open(destination_yaml, 'w') as outfile:
        yaml.dump(file_copy, outfile)

```

## Annexe 2 : Script de réentraînement du modèle YOLOv8

```
!pip uninstall numpy -y
!pip uninstall numpy -y
!pip install numpy==1.26.4
!pip install --upgrade pandas==2.0.0
!pip uninstall opencv
!pip install --upgrade opencv-python==4.8.0.74

import os
import cv2
import numpy as np
from ultralytics import YOLO
import threading as th

rep_travail = '/workspace/stage1A/imeane_test/test_entrainement'
os.chdir(rep_travail)

def train(dataset, length_batch, file_name):
    model = YOLO('yolov8s-seg.pt') # Load a pretrained model
    model.train(data=dataset, batch=length_batch, epochs=100, imgsz=640, name=file_name) # Train the model
    #batch c'est le nbr d'images par batch

def default_training():
    dataset = f"{rep_travail}/dataset/data.yaml"
    length_batch = 16 #default 16
    train(dataset,length_batch,'test')

import json

with open(f'{rep_travail}/dataset/data.json', 'r') as file :
    data = json.load(file)

print(data)

length_batch = 16

for model in data :
    train(f'{rep_travail}/dataset/data.yaml', length_batch, model['name'])
    model['chemin_metrics'] = f"/workspace/stage1A/imeane_test/test_entrainement/runs/segment/{model['name']}"
    model['chemin_poids'] = f"/workspace/stage1A/imeane_test/test_entrainement/runs/segment/{model['name']}/weights/best.pt" #Si on prends best.pt et non last.pt

with open(f'{rep_travail}/dataset/data.json', 'w') as file :
    json.dump(data,file)
```