

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

ÉCOLE NATIONALE POLYTECHNIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



Département Génie Industriel

Mémoire de projet de fin d'études

En vue de l'obtention du diplôme d'Ingénieur d'État en Génie Industriel
Option : Data Science and Intelligence Artificielle

Systeme d'Inspection Qualité pour les Sachets de Détergent par
Vision Industrielle et Intelligence Artificielle

LAKACHE Ahmed

Sous la direction de **Mr. ZOUAGHI Iskander ENP & Mme BARKAT
Nouha HENKEL**

Présenté et soutenu publiquement le (11/07/2024)

Composition du jury :

Président :	Mr. ADNANE Mourad	Professeur	ENP
Promoteur :	Mr. ZOUAGHI Iskander	MCA	ENP
Promotrice :	Mme BARKAT Nouha	Ingénieur	Henkel
Examinatrice :	Mme. BELDJOUDI Samia	MCA	ENP

ENP 2024

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

ÉCOLE NATIONALE POLYTECHNIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



Département Génie Industriel

Mémoire de projet de fin d'études

En vue de l'obtention du diplôme d'Ingénieur d'État en Génie Industriel
Option : Data Science and Intelligence Artificielle

Systeme d'Inspection Qualité pour les Sachets de Détergent par
Vision Industrielle et Intelligence Artificielle

LAKACHE Ahmed

Sous la direction de ZOUAGHI Iskander ENP & BERKAT Nouha
HENKEL

Présenté et soutenu publiquement le (11/07/2024)

Composition du jury :

Président :	Mr. ADNANE Mourad	Professeur	ENP
Promoteur :	Mr. ZOUAGHI Iskander	MCA	ENP
Promotrice :	Mme BARKAT Nouha	Ingénieur	Henkel
Examinatrice :	Mme. BELDJOUDI Samia	MCA	ENP

ENP 2024

ملخص

يهدف هذا المشروع إلى حل التحديات المتعلقة بفحص الجودة في عملية إنتاج المنظفات لدى شركة هنكل. الهدف الرئيسي هو تطوير نظام متكامل لمراقبة الجودة باستخدام تقنيات متقدمة في التعلم الآلي والرؤية الصناعية للكشف عن العيوب في الوقت الفعلي. من خلال تطبيق نماذج رؤية حاسوبية متطورة، يسعى هذا المشروع إلى تحسين كفاءة الإنتاج، خفض التكاليف المرتبطة بعدم الجودة، وضمان رضا العملاء بشكل أفضل. بالإضافة إلى ذلك، يمكن أن يُستخدم النظام لاحقاً لجمع البيانات في الوقت الفعلي لتطوير نموذج صيانة تنبؤية، مما قد يساهم في تحقيق سياسة الصفر خسائر لشركة هنكل.

الكلمات المفتاحية : فحص الجودة - إنتاج المنظفات - نظام مراقبة الجودة - التعلم الآلي - الرؤية الصناعية - الكشف عن العيوب - كفاءة الإنتاج - رضا العملاء - جمع البيانات في الوقت الفعلي

Abstract

This project aims to tackle the challenges of quality inspection in Henkel's detergent production process. The primary goal is to develop an integrated quality control system using advanced machine learning and computer vision techniques to detect defects in real-time. By leveraging state-of-the-art computer vision models, the project seeks to enhance production efficiency, reduce non-conformity costs, and ensure higher customer satisfaction. Additionally, the system may later be used to collect real-time data to assist in developing a predictive maintenance model, thereby potentially supporting Henkel's zero-loss policy.

Keywords : quality inspection - detergent production - Henkel - quality control - machine learning - computer vision - defect detection - real-time - computer vision models - production efficiency - non-conformity costs - customer satisfaction - real-time data collection - predictive maintenance - zero-loss policy.

Résumé

Ce projet vise à résoudre les défis liés à l'inspection de la qualité dans la production de détergents chez Henkel. L'objectif principal est de développer un système intégré de contrôle qualité utilisant des techniques avancées d'apprentissage automatique et de vision industrielle pour détecter les défauts en temps réel. Grâce à l'application de modèles de vision par ordinateur, ce projet cherche à améliorer l'efficacité de la production, à réduire les coûts liés aux non-conformités, et à garantir une meilleure satisfaction des clients. De plus, le système pourrait ultérieurement collecter des données en temps réel pour développer un modèle de maintenance prédictive, contribuant ainsi potentiellement à la politique de zéro perte de Henkel.

Mots clés : inspection de la qualité - production de détergents - Henkel - contrôle qualité - apprentissage automatique - vision industrielle - détection des défauts - temps réel - modèles de vision par ordinateur - efficacité de la production - réduction des coûts - non-conformités - satisfaction des clients - collecte de données - maintenance prédictive

Dédicace

*Je dédie ce travail aux deux personnes les plus chères à mon cœur.
À ma mère, pour son amour inconditionnel, sa patience infinie et ses
encouragements constants. Ta sagesse et ta bienveillance ont été ma source
de motivation.*

*À mon père, le pilier solide de notre famille, pour sa force, son soutien
indéfectible et ses conseils avisés. Ta confiance en moi m'a donné le courage
de surmonter tous les obstacles.*

*À mes frères *Khaled, Yahia*, et ma sœur *Khadija*, pour leur soutien
inébranlable et leur amour fraternel.*

*À *Monsaf* (special thanks to him), *Younes, Amine et Issam*, pour leur aide
exceptionnelle dans la réalisation de ce projet. Votre collaboration et votre
amitié ont été inestimables.*

*Bien sûr, une dédicace spéciale aux gens du *Mondât s6*, pour tous les
moments partagés et les souvenirs créés. Le sixième pilier est enfin tombé.*

*Et enfin, je n'oublie pas les personnes qui m'ont accompagné pendant ces six
années d'études, en particulier *Hammid, Anis, Zaki, Abdelhak* et les gens du
M54. Votre camaraderie et votre soutien ont rendu ce voyage académique
inoubliable.*

Merci.

- Ahmed

Remerciements

C'est avec une profonde gratitude que je tiens à exprimer mes remerciements à tous ceux qui ont contribué à la réalisation de ce projet de fin d'études.

Mes premiers remerciements vont à Dieu le tout-puissant, source de force, de volonté et de patience qui m'ont permis de mener à bien ce travail.

Je tiens à exprimer ma gratitude à **Mr. Iskander ZOUAGHI**, mon professeur et encadreur, pour son encadrement et sa patience avec moi. Je le remercie non seulement en tant qu'encadreur, mais également en tant que professeur qui figure parmi les enseignants les plus influents dans mon parcours académique. Son soutien et ses conseils avisés ont été essentiels à la réalisation de ce projet.

Ma gratitude s'étend également à **Mme. Nouha BARKAT**, ma superviseure chez Henkel. Son expertise, sa disponibilité et son soutien indéfectible ont été des piliers essentiels durant mon stage. Son accompagnement bienveillant a transformé cette expérience en une opportunité d'apprentissage exceptionnelle, me permettant de surmonter chaque défi avec confiance.

Je remercie chaleureusement le groupe Henkel pour cette opportunité enrichissante. Une mention spéciale à **Mr. Fadlalah Karoune** et à toute la division qualité pour leur collaboration précieuse. Leur engagement envers l'excellence et leur expertise ont considérablement enrichi mon expérience et inspiré mon travail.

J'adresse mes remerciements anticipés aux membres du jury qui évalueront ce travail. Leur temps, leur expertise et leurs critiques constructives sont grandement appréciés et contribueront sans nul doute à affiner la qualité de ce projet.

Je tiens également à exprimer ma reconnaissance envers le corps enseignant du département de génie industriel de l'École Nationale Polytechnique. Leur dévouement et la qualité de leur enseignement ont été des éléments fondamentaux dans mon développement académique et professionnel.

Enfin, je souhaite remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce projet. Votre soutien, vos encouragements et votre aide ont été inestimables tout au long de ce parcours.

Ce projet représente non seulement l'aboutissement de mes études, mais aussi le début d'une nouvelle étape dans ma carrière professionnelle. Je suis reconnaissant pour toutes les leçons apprises et les expériences vécues, qui seront des atouts précieux pour mes futurs défis.

Table des matières

Liste des figures

Liste des tableaux

Introduction générale	13
1 Contexte Industriel et Problématique	15
1.1 Présentation de l'entreprise Henkel et de son processus de production	15
1.1.1 Présentation de l'entreprise HENKEL	15
1.1.2 Présentation de HENKEL Algérie	17
1.2 Chaîne d'Approvisionnement et Processus de Production de Poudre Détergente	19
1.3 Processus de fabrication	22
1.4 Identification et analyse de la problématique	27
1.5 Analyse des Défauts	27
1.5.1 Analyse des Défauts chez Henkel	32
1.5.2 Analyse des coûts de retour	32
1.6 Problématique	32
1.6.1 Objectifs spécifiques du projet	32
1.7 Conclusion	33
2 Fondements Théoriques et Revue de Littérature	34
2.1 Défauts de production des pouches (sachets)	34
2.2 Mise au point sur les problèmes de soudure et d'étiquetage	35
2.2.1 Défauts de soudure	35
2.2.2 Défauts d'étiquetage : types, causes et conséquences	36
2.2.3 Techniques existantes pour traiter les problèmes de soudure	37

2.2.4	Techniques existantes pour traiter les problèmes d'étiquetage	38
2.2.5	Comparaison des différentes approches	39
2.3	vision industrielle	40
2.3.1	Historique et évolution	41
2.3.2	La Vision par Ordinateur dans l'Industrie 4.0	42
2.3.3	Avantages de la Vision par Ordinateur dans l'Industrie 4.0	43
2.4	Fondamentaux de l'Apprentissage Automatique	43
2.4.1	Catégorisation de l'Apprentissage Automatique	44
2.4.1.1	Apprentissage supervisé	44
2.4.1.2	Apprentissage semi-supervisé :	45
2.4.1.3	Apprentissage non supervisé :	45
2.4.1.4	Apprentissage par Renforcement :	48
2.4.2	Processus d'Apprentissage et d'Évaluation des Modèles	49
2.5	Apprentissage Profond Appliqué à la Machine Vision	58
2.6	Apprentissage par transfert (Transfer Learning)	61
2.6.1	Fine-tuning pour l'adaptation des modèles :	63
2.6.2	Modèles pré-entraînés populaires :	65
2.6.3	Évaluation des modèles pré-entraînés pour la détection d'objet	66
2.6.4	l'apprentissage automatique VS l'apprentissage par transfert	69
2.7	Technologies des Caméras Industrielles	70
3	Conception et Développement du Système de Contrôle Qualité	72
3.1	Partie1 :Inspection de l'impression des Dates	72
3.1.1	Architecture du système	72
3.1.2	Développement des Modèles D'inspection d'impression	75
3.1.3	Choix des Algorithmes et Évaluation des Modèles	77
3.1.4	Evaluation du Modèle	80
3.1.5	Reconnaissance Optique de Caractères (OCR)	83
3.2	Partie2 :Inspection du Scellage	85
3.2.1	Architecture du Système	85
3.2.2	Collecte et Préparation des Données :	87

3.2.3	Développement des Modèles d'inspection de scellage	90
3.3	Conclusion	100
4	Analyse de Rentabilité pour le Projet de Système d'Inspection de Qualité	102
4.1	Investissement Initial	102
4.2	Économies Réalisées Grâce à la Réduction des Défauts	103
4.2.1	Détails des Coûts de Retour	103
4.2.2	Calcul des Coûts de Retour Totaux	103
4.3	Calcul du Temps de Récupération de l'Investissement	104
4.4	Conclusion	104
	Bibliographie	106

Liste des figures

1.1	HENKEL LOGO	15
1.2	Organisation de l'entreprise Henkel dans le monde (Henkel global)	15
1.3	Objectives de Henkel	16
1.4	Organigramme de Henkel Chelghoum Laid	18
1.5	durabilité	21
1.6	Processus de fabrication du détergent en poudre	22
1.7	Cartographie de la chaîne de valeur (VSM) du processus de fabrication de détergent en poudre chez Henkel (une autres figure plus claire est ajoutée au annexe)	26
1.8	L'impact des défaut sur la qualité [1]	28
1.9	L'impact des défaut sur la performances der l'entreprise [1]	29
1.10	L'Impact des Défauts et des Retours sur la Satisfaction des Clients - Coût Tangible de la Qualité : Assurer la Satisfaction des Clients [1]	30
1.11	figure représenatif des défauts de seullage	30
1.12	Figures représentatif des défauts visuelle	31
1.13	défautes d'impression (étiquetages)	31
2.1	Étapes de la machine vision	41
2.2	integration de la machine vision dans l industry 4.0 [2]	42
2.3	utilisation de la machine vision dans l industry 4.0 [2]	43
2.4	apprentissage supervisé [3]	44
2.5	Catégorie de l'apprentissage supervisé [4]	45
2.6	Apprentissage non supervisé [5]	46
2.7	clustring [6]	46
2.8	Règles d'associations [7]	47
2.9	Apprentissage par renforcement [8]	49

2.10	Matrice de confusion [9]	55
2.11	Schéma représentant l'analogie entre le neurone biologique et le neurone artificielles [10]	58
2.12	Réseaux de Neurones Convolutifs (CNN) [11]	59
2.13	Couche Convulsive [12]	60
2.14	Fonction d'Activation (ReLU) [13]	60
2.15	Couche de Pooling [14]	61
2.16	Schéma d'une couche entièrement connectée avec 6 classes [15]	61
2.17	Apprentissage par transfert [16]	62
2.18	Explication de fine-tuning [17]	64
3.1	Position de la caméra dans la ligne volpack	74
3.2	Exemple d'annotation d'images avec Roboflow. Cette figure montre un exemple de l'interface Roboflow utilisée pour annoter les images avec des boîtes englobantes autour des zones d'intérêt.	76
3.3	Comparaison des méthodes classiques de traitement d'image pour la détection de la zone d'impression	77
3.4	Courbes de perte d'entraînement et de validation	80
3.5	Courbes de performance pour la validation et l'entraînement	80
3.6	Matrice de confusion	81
3.7	Courbes F1-Confidence pour box et mask	81
3.8	Courbes Precision-Confidence pour box et mask	82
3.9	Courbes Precision-Recall pour box et mask	82
3.10	Courbes Recall-Confidence pour box et mask,	82
3.11	Prédictions du Modèle YOLOV8 montrant une segmentation précise des zones d'impression sur différentes images	83
3.12	Résultat de segmentation de modèle YoloV8	83
3.13	Application de l'OCR sur la zone segmenté	84
3.14	Position de la caméra par rapport a la chaîne de conditionnement	86
3.15	Interface	88
3.16	annotation avec roboflow	90
3.17	Graphiques des métriques d'évaluation des modèles	91
3.18	Graphiques des métriques d'évaluation des modèles	92

3.19	Détection d'un scellage incorrect avec une précision de 0.94	94
3.20	Détection d'un scellage correct avec une précision de 0.88	94
3.21	annotation avec roboflow	96
3.22	Graphiques des resultat	98
3.23	Détection d'un scellage incorrect	99
3.24	Détection d'un scellage correct	99
4.1	Value stream mapping	113

Liste des tableaux

1.1	Les dates clés de l'évolution du groupe HENKEL	16
2.1	Comparaison des différentes techniques de traitement des problèmes de soudure	39
2.2	Comparaison des différentes techniques de traitement des problèmes d'étiquetage	40
2.3	Comparaison entre l'apprentissage supervisé et non supervisé	48
2.4	Tableau comparatif des types de transfert learning	63
2.5	Modèles pré-entraînés et leurs utilisations	66
2.6	Comparaison entre Machine Learning et Transfer Learning - Caractéristiques	69
2.7	Classification et Applications des Caméras Industrielles	71
3.1	Caractéristiques de la caméra MER2-503-36U3M/C	73
3.2	Vitesse de production et intervalle de capture d'images pour différentes gammes de sachets	75
3.3	Répartition des données pour l'entraînement, la validation et les tests.	77
3.4	YOLOv8m Segmentation	78
3.5	Avantages et inconvénients de la caméra thermique FLIR A35	87
3.6	Comparaison des performances entre YOLO v9 et YOLO v8	92
3.7	Évolution des pertes du modèle YOLOv8 Large (YOLOv8-l) pour la détection de la qualité du scellage	97
3.8	Performances des métriques du modèle YOLOv8 Large (YOLOv8-l) pour la détection de la qualité du scellage	98
3.9	Concept de chaque approche	100
3.10	Caractéristiques des approches	100
4.1	Détails des Coûts de Retour	103
4.2	Répartition des Défauts par Taille de Produit	103
4.3	Coûts de Retour Totaux	104

Liste des sigles et acronymes

- **AI** - Artificial Intelligence
- **CNN** - Convolutional Neural Network
- **CPU** - Central Processing Unit
- **DNN** - Deep Neural Network
- **DFL** - Distribution Focal Loss
- **FPS** - Frames Per Second
- **GPU** - Graphics Processing Unit
- **IoU** - Intersection over Union
- **mAP** - Mean Average Precision
- **ML** - Machine Learning
- **OCR** - Optical Character Recognition
- **RCNN** - Region-based Convolutional Neural Network
- **ROI** - Region Of Interest
- **YOLO** - You Only Look Once

Introduction générale

Dans un contexte industriel en constante évolution, la qualité des produits est devenue un facteur clé de compétitivité pour les entreprises manufacturières. L'inspection de la qualité dans les processus de production est cruciale pour garantir la conformité des produits aux normes et minimiser les défauts. L'évolution rapide des technologies de vision industrielle et d'apprentissage automatique offre de nouvelles opportunités pour améliorer ces processus, en particulier dans l'industrie manufacturière.

Ce mémoire se concentre sur l'application de techniques d'apprentissage automatique et de vision industrielle pour l'inspection de la qualité de l'impression des dates et du scellage thermique des sachets de détergent au sein de l'entreprise Henkel. L'objectif principal est de développer un système intégré de contrôle qualité qui permet de détecter et de corriger les défauts en temps réel, améliorant ainsi l'efficacité de la production et réduisant les coûts liés aux non-conformités. Ce système vise à automatiser le processus d'inspection, réduisant ainsi les erreurs humaines et augmentant la précision et la rapidité des contrôles.

Dans la première partie de ce mémoire, nous examinerons le contexte industriel de l'entreprise Henkel et la problématique de la qualité dans le processus de production de détergent en poudre. Nous décrirons le processus de production, les défis rencontrés en matière de contrôle qualité, et les impacts des défauts sur la performance de l'entreprise.

Ensuite, nous présenterons les concepts et les approches clés en vision industrielle et en apprentissage automatique. Nous aborderons les différentes techniques d'apprentissage telles que l'apprentissage supervisé, non supervisé, et l'apprentissage profond, avec un focus particulier sur les réseaux de neurones convolutifs (CNN) et l'apprentissage par transfert. Nous expliquerons également les principes de la vision industrielle appliquée à l'inspection de qualité.

Dans la troisième partie, nous détaillerons la conception et le développement du système de contrôle qualité proposé. Cette section comprendra la description de l'architecture du système, les équipements utilisés, et les méthodologies employées pour la collecte et le prétraitement des données. Nous discuterons également du développement et de l'évaluation des modèles d'apprentissage automatique pour l'inspection de l'impression des dates et du scellage thermique.

Une partie cruciale de ce mémoire portera également sur l'analyse de rentabilité du projet. Nous montrerons comment l'implémentation du système d'inspection de qualité permet non seulement d'améliorer la détection des défauts et de réduire les coûts de retour des produits défectueux, mais aussi de garantir une récupération rapide de l'investissement initial. L'analyse détaillera les économies réalisées grâce à la réduction des défauts, les coûts de retour évités, et le temps de récupération de l'investissement, démontrant ainsi la viabilité financière et les bénéfices économiques du projet pour l'entreprise Henkel.

Enfin, nous concluons ce mémoire en résumant les principaux points abordés et en soulignant

l'importance de l'application des techniques d'apprentissage automatique et de vision industrielle dans l'industrie manufacturière. Nous présenterons également les perspectives d'amélioration et les futurs axes de recherche pour optimiser davantage les processus de contrôle qualité.

L'objectif de ce mémoire est de fournir une base solide pour la mise en œuvre d'un système de contrôle qualité automatisé chez Henkel, en utilisant des techniques avancées d'apprentissage automatique et de vision industrielle. Ce système vise à améliorer la détection des défauts, à réduire les coûts de production, et à garantir une meilleure satisfaction des clients. Notre approche innovante combine des techniques de pointe en traitement d'image et en apprentissage profond, offrant une solution adaptée aux spécificités de la production de détergents en poudre.

Ce projet s'inscrit dans la dynamique plus large de l'Industrie 4.0, illustrant comment l'intégration de technologies intelligentes peut transformer les processus de production traditionnels et ouvrir la voie à des usines plus efficaces et plus flexibles. En développant ce système de contrôle qualité avancé, nous espérons non seulement améliorer les performances de Henkel, mais aussi contribuer à l'avancement des connaissances dans le domaine de l'inspection automatisée et de l'application de l'intelligence artificielle dans l'industrie manufacturière.

Chapitre 1

Contexte Industriel et Problématique

1.1 Présentation de l'entreprise Henkel et de son processus de production

Historique, secteur d'activité et importance stratégique de l'optimisation de la production et du contrôle qualité

1.1.1 Présentation de l'entreprise HENKEL



FIGURE 1.1 – HENKEL LOGO

HENKEL est une entreprise multinationale spécialisée dans les produits de consommation et les adhésifs. Fondée en 1876, elle a depuis évolué pour devenir un acteur majeur dans plusieurs secteurs industriels.

Le groupe HENKEL est présent dans 125 pays répartis sur les cinq continents. Il emploie plus de 50 000 personnes avec une implantation industrielle dans 75 pays. Le groupe HENKEL a fait ses preuves sur différents marchés.

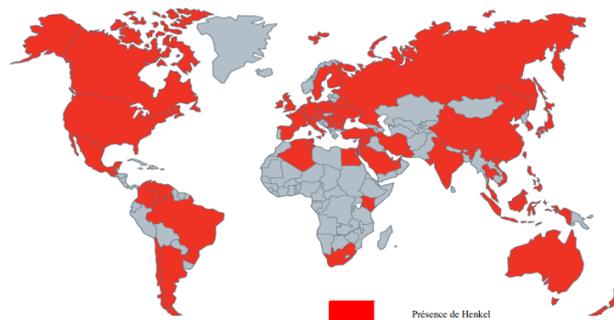


FIGURE 1.2 – Organisation de l'entreprise Henkel dans le monde (Henkel global)

HENKEL se veut un leader mondial des marques et des technologies, mettant en avant des

valeurs centrées sur les clients, la performance financière, le développement durable, et une base familiale solide.



FIGURE 1.3 – Objectives de Henkel

Le tableau 1.1 représente les dates clef de l'évolution du group henkel

TABLE 1.1 – Les dates clés de l'évolution du groupe HENKEL

Année	Événement
1876	Création de la société par Fritz HENKEL à Achen en Allemagne
1907	Lancement de la première marque de détergents (PERSIL)
1913	Création de la première filiale à l'étranger (HENKEL & Cie AG) en Suisse
1923	Démarrage de la production des colles et adhésifs (PATEX)
1924	Lancement de la marque d'hygiène industrielle (TSUNAMI)
1946	Début de la production des produits cosmétiques
1974	Demande de brevet pour SASIL, un substitut de phosphate
1982	Lancement du premier détergent sans phosphate (DIXAN)
1990	Alliance avec une entreprise chinoise ; Ré-acquisition de l'usine à Genthin
1991	Alliance HENKEL-ECOLAB ; Création de HENKEL-ASIA PACIFIC Ltd
1993	Participation dans TONSO à St Petersburg
1996	Ouverture du capital ; Fondation de SCHWARZKOPF & HENKEL Cosmetics
1997-98	Acquisitions de LOCTITE Corps et DEP Corps aux USA
1999	Création de COGNIS
2000	Alliances avec ENAD, MULTICORE, LION, et DEXTER
2001	Cession de COGNIS
2002	Lancement de "Henkel A Brand like a Friend"
2003	Acquisition de DIAL Corporation
2004	Rachat des parts de l'ENAD en Algérie

Les différents secteurs d'activité du groupe HENKEL :

Le groupe HENKEL est présent dans trois secteurs d'activités :

- **Détergents et produits d'entretien :**

Le groupe et les marques de produits HENKEL les plus connus et les plus appréciés, dont beaucoup sont distribués dans le monde, comprennent : le linge, la vaisselle,

l'entretien de la maison. Le portefeuille produit du groupe HENKEL comprend : détergents universels, lessives spéciales, adoucissants et détachants, produits pour lave-vaisselle, produits d'entretien, produits à récurer, produits de soin pour sols et moquettes, nettoyeurs bain et toilettes, nettoyeurs vitres, nettoyeurs cuisines et autres produits d'entretien spécialisés. Marques Dial, Purex, Renuzit, Armour.

- **Cosmétiques :**

Les nombreux produits de marque développés, fabriqués et commercialisés dans le monde, dans le secteur cosmétiques, connus sur le marché sous la marque Schwarzcopf & HENKEL, se décomposent en un certain nombre de gammes de produits. Des articles de marque pour professionnels du secteur de la coiffure, c'est à dire Schwarzcopf professionnel, complètent leur gamme. Cette gamme de produit HENKEL comprend : shampoings et après-shampoings, colorations, produits de styling et de permanente, savons, produits pour le bain et la douche, déodorants, crèmes pour la peau, produits de soin de la peau, produits d'hygiène buccodentaire, parfums et fragrances, produits pour salons de coiffure. Marques Dial : Coast, Tone, Pure & Natural. buccodentaire, parfums et fragrances.

- **Colles adhésifs et produits d'étanchéité :**

L'activité Colles et Adhésifs de HENKEL constitue une seule division responsable de ses activités dans le monde entier : adhésifs pour professionnels et adhésifs grand public. Dans ce secteur, le groupe offre une large gamme de marque de produits pour la maison, le bureau et pour le marché du bricolage domestique. Pour les professionnels, elle fournit des adhésifs et des produits d'étanchéité pour la construction, la pose de sols, la toiture, la rénovation et le mobilier pour quelques uns.

Le chiffre d'affaires de HENKEL pour les secteurs de détergents et d'entretien sont de 35% de l'ensemble des ventes, viennent ensuite le secteur des cosmétiques et des colles et mastiques industriels qui représente chacun approximativement 25% du chiffre d'affaires du groupe HENKEL. Quant au secteur des colles adhésifs et étanchéité, il représente 15% du chiffre d'affaires global.

1.1.2 Présentation de HENKEL Algérie

La société HENKEL Algérie est une société par actions (SPA) avec trois unités de production situées à Réghaia (Centre), Chelghoum El Laid (Est), et Ain Temouchent (Ouest). Son siège est situé à : 22 rue Ahmed OUAKED, Bois des cars 3, Delly Ibrahim, Alger. HENKEL Algérie emploie près de 1200 salariés.

Organigramme de Henkel Chelghoum Laid :

Dans l'organigramme de Henkel Chelghoum Laid, la structure hiérarchique met en lumière les principaux acteurs qui contribuent au bon fonctionnement du site. Le directeur de l'usine supervise l'ensemble des opérations, tandis que les différents départements, tels que la qualité, la sécurité interne et les services généraux, l'ingénierie et les projets, travaillent en synergie pour assurer l'efficacité des processus.

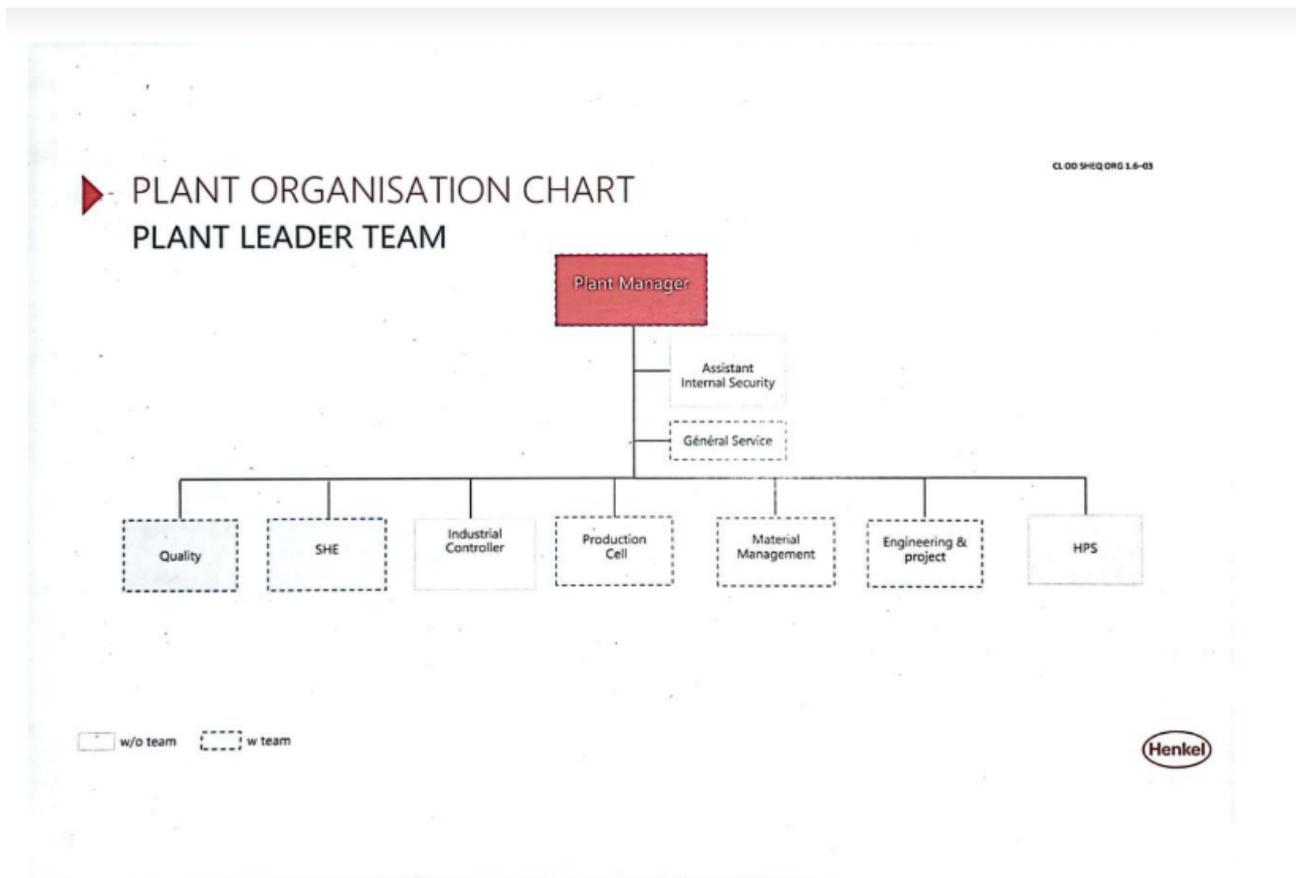


FIGURE 1.4 – Organigramme de Henkel Chelghoum Laid

HENKEL – ENAD Algérie et Passage à HENKEL Algérie

Le partenariat de l'ENAD avec l'un des plus grands producteurs mondiaux de détergents, HENKEL, a été un moyen de mise à niveau afin de répondre à une stratégie gouvernementale algérienne encourageant les investissements étrangers (financement, know-how, mise à niveau). Cela visait également à éviter la perte des parts de marché de l'ENAD face à la concurrence qui est apparue suite à l'ouverture de l'économie algérienne.

Après l'appel d'offre international lancé par ENAD, les trois leaders mondiaux du secteur des détergents, à savoir, PROCTER et GAMBLE, UNILEVER et HENKEL, ont soumissionné. Le choix définitif s'est porté sur le groupe allemand HENKEL qui a présenté la meilleure offre.

L'accord de joint-venture permet à un partenaire étranger de développer un produit qu'il cherche à diffuser dans un pays au sein duquel l'autre partenaire (local) dispose d'un accès aisé. Cette démarche évite donc à l'investisseur étranger les phases les plus difficiles et délicates d'implantation et de lancement.

L'accord de partenariat entre HENKEL et ENAD a mis l'accent sur les points suivants :

- Réalisation d'un programme ambitieux : plus de 300 milliards de dinars ;
- Réalisation d'un programme d'investissement de 185 millions de dollars US dans un délai de trois à cinq ans ;
- Maintien de l'ensemble du personnel des trois unités de Réghaia, Ain Temouchent et Chelghoum El Laid ;
- Assistance technique gratuite fournie par les techniciens de HENKEL pour permettre la mise à niveau et la réhabilitation du complexe de détergents de Chelghoum El Laid ;
- Exploitation de la marque ISIS, une des marques leaders de ENAD, suite au partenariat

HENKEL ENAD Algérie (HEA).

La société par action (SPA) a été créée le 23 mai 2000 avec un capital de 1.760 milliards de dinars réparti en 60% pour HENKEL et 40% pour ENAD. HEA est une société de droit algérien. À noter que la reprise du complexe de détergents de Chelghoum El Laid a été concrétisée le 29 mai 2002 sur décision du conseil de participation de l'État (CPE) du 13 février 2002.

Après quatre années et demi de présence sur le marché algérien, HENKEL a montré son attachement à l'Algérie en devenant 100% détenteur du capital de HEA après avoir racheté les 40% des actions. La cession des parts de ENAD à HENKEL a été concrétisée par la signature d'un acte notarié en date du 8 septembre 2005 entre le président de la SGP Gephac (M. Chérif BOUNAB) et le président du groupe HENKEL France. Les 40% des actions de ENAD représentaient 880 millions de dinars.

C'est au terme de cette opération que ENAD a achevé la privatisation totale de ses trois installations industrielles. HENKEL a investi près de 16 millions de dollars dans la mise à niveau des complexes et usines :

- 15 millions de dollars dans la rénovation des installations de Réghaia et de Ain Temouchent ;
- 500 000 euros dans la mise au niveau du complexe de Chelghoum El Laid.

Les moyens modernes et le système d'information HENKEL Algérie a fait appel à des solutions informatiques afin de réduire le temps nécessaire pour l'accomplissement de ses diverses tâches avec accroissement de l'efficacité notamment :

Internet : Un réseau Internet est installé au niveau des trois usines ainsi que la direction générale.

Intranet : Logiciel de messagerie interne (OUTLOOK), et est le cœur de la communication au sein de l'entreprise grâce au courrier électronique, transfert et partage des fichiers.

Le SAP (Système, Application, Produit) : SAP est un ERP (Entreprise Resource Planning), système d'information et de gestion qui permet, à partir des prévisions de la demande, de disposer du planning de la production, la mise à jour des stocks, le réapprovisionnement des matières premières et des emballages, la livraison des entrepôts et toute la préparation des transports (documents, comptabilité, factures). Son objectif est de réduire les coûts, d'accélérer la diffusion de l'information et d'améliorer les décisions prises. Ce programme est doté d'une particularité, il s'agit du déclenchement automatique d'une simple série d'opérations dès l'introduction d'une donnée, ainsi l'enregistrement d'une simple commande permet l'établissement immédiat d'un programme d'enlèvement, d'un bon de sortie et bien d'autres opérations.

HENKEL Algérie a bénéficié du concours précieux de la filiale française du groupe qui utilise ce système depuis plusieurs années, laquelle a joué un rôle important dans la formation du personnel de HENKEL Algérie à l'utilisation de ce système.

1.2 Chaîne d'Approvisionnement et Processus de Production de Poudre Détergente

Henkel, un leader mondial dans les secteurs des biens de consommation et industriels, a établi un réseau de chaîne d'approvisionnement sophistiqué et étendu qui connecte les fournisseurs,

les installations de production, les centres de distribution et les clients à travers le monde. La stratégie de gestion de la chaîne d'approvisionnement de l'entreprise se concentre sur l'optimisation de l'efficacité, la minimisation des coûts et la garantie de la livraison en temps voulu de produits de qualité supérieure aux clients dans divers marchés.

Gestion des Fournisseurs :

Henkel maintient un réseau complet de plus de 6 000 fournisseurs dans le monde, fournissant une variété de matières premières, de matériaux d'emballage et d'autres intrants essentiels à la production. Ces fournisseurs sont stratégiquement positionnés dans différentes régions pour assurer un approvisionnement stable et adaptable en matériaux. Henkel emploie des processus rigoureux de sélection et d'évaluation des fournisseurs pour évaluer la qualité, la fiabilité et la durabilité de ses fournisseurs. L'entreprise effectue des audits périodiques et des évaluations de performance pour maintenir des normes élevées et cultiver des partenariats durables avec des fournisseurs clés [18][19].

Fabrication :

Avec plus de 180 installations de fabrication réparties dans 57 pays, y compris une usine de production ultramoderne en Algérie, Henkel possède une présence de fabrication mondiale robuste. Ces installations sont équipées de lignes de production et de technologies de pointe, facilitant la production efficace et cohérente d'une gamme diversifiée de produits de poudre détergente. L'usine d'Algérie, notamment, dispose d'une capacité de production dépassant les 200 000 tonnes par an et fournit de l'emploi à environ 500 individus [20].

Les processus de fabrication de Henkel sont régis par des procédures strictes de contrôle et d'assurance qualité pour garantir que les produits respectent les spécifications et normes nécessaires. L'entreprise investit dans des initiatives d'amélioration continue, telles que Lean Six Sigma, pour rationaliser les processus de production, minimiser les déchets et améliorer la productivité [18][19].

Distribution :

Henkel a établi un réseau de distribution complet comprenant plus de 200 entrepôts et centres de distribution dans le monde entier. L'entreprise collabore avec des partenaires logistiques de premier plan pour superviser le transport et la livraison des produits aux clients. Henkel utilise des outils avancés de planification et de programmation, tels que SAP APO (Advanced Planning and Optimization), pour optimiser les niveaux de stock, réduire les délais de livraison et garantir la livraison ponctuelle des produits. Le réseau de distribution de l'entreprise est conçu pour être flexible et réactif aux conditions de marché en évolution et aux exigences des clients. Henkel surveille en permanence les indicateurs clés de performance (KPI) tels que les taux de livraison à temps, la rotation des stocks et les coûts de transport pour identifier les domaines à améliorer et maintenir des niveaux de service élevés [18][19].

Service Client :

Henkel accorde une grande importance à la satisfaction des clients et dispose d'équipes de service client dédiées pour répondre aux demandes, plaintes et commentaires. L'entreprise fournit plusieurs canaux pour que les clients puissent la contacter, y compris par téléphone, par email et via des portails en ligne. Les représentants du service client de Henkel sont formés pour offrir une résolution rapide et efficace des problèmes et préoccupations des clients.

L'entreprise effectue également des enquêtes régulières de satisfaction client et recueille des commentaires pour améliorer continuellement ses produits et services. L'approche centrée sur le client de Henkel a permis à l'entreprise de forger des relations solides avec ses principaux clients et de maintenir une base de clients fidèle [21][22].

Durabilité :

La durabilité est une priorité essentielle pour la gestion de la chaîne d'approvisionnement de Henkel. L'entreprise s'est fixée des objectifs ambitieux pour réduire son empreinte environnementale et promouvoir des pratiques d'approvisionnement responsable. Henkel travaille en étroite collaboration avec ses fournisseurs pour assurer la conformité à sa politique d'approvisionnement durable et effectue des évaluations régulières pour surveiller la performance des fournisseurs sur des critères sociaux et environnementaux.

Henkel a également mis en œuvre diverses initiatives pour réduire les déchets, optimiser la consommation des ressources et minimiser l'empreinte carbone de ses opérations de fabrication et de distribution. Par exemple, l'entreprise a investi dans des sources d'énergie renouvelable, telles que l'énergie solaire et éolienne, pour réduire sa dépendance aux combustibles fossiles

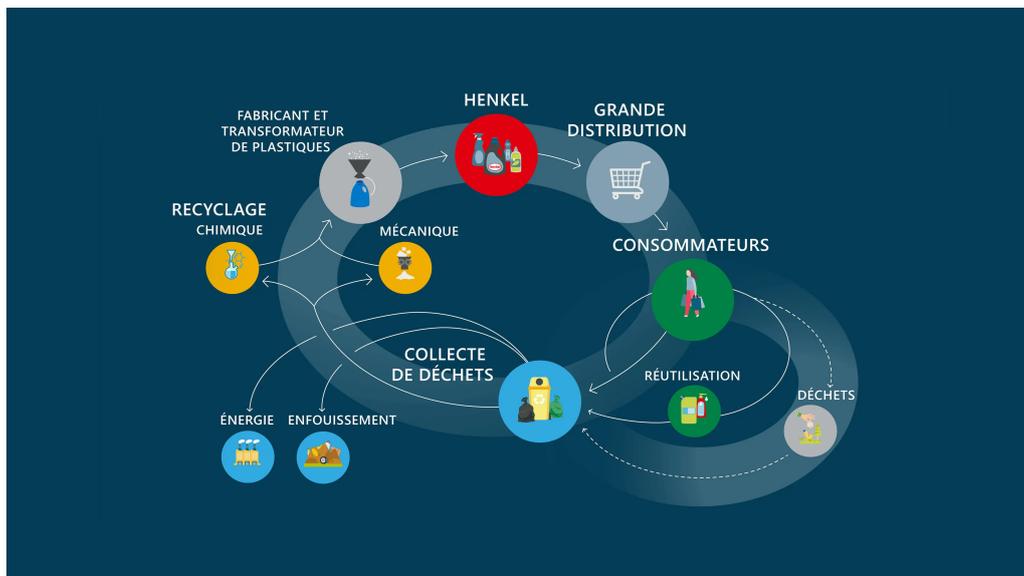


FIGURE 1.5 – durabilité

Intégration et Collaboration :

Le succès de la chaîne d'approvisionnement de Henkel repose sur une intégration et une collaboration robustes entre les diverses fonctions et parties prenantes. L'entreprise utilise des technologies avancées et des outils numériques pour faciliter le partage d'informations et la coordination à travers la chaîne d'approvisionnement. Par exemple, Henkel a mis en place une plateforme cloud de collaboration avec les fournisseurs pour permettre une communication en temps réel et une résolution conjointe des problèmes [23].

Henkel travaille également en étroite collaboration avec ses clients pour comprendre leurs besoins et développer des solutions personnalisées. L'entreprise a établi des initiatives de création de valeur conjointe avec des clients clés pour stimuler l'innovation, optimiser les processus de la chaîne d'approvisionnement et créer une valeur partagée [22].

1.3 Processus de fabrication

La fabrication de poudre détergente chez Henkel est un processus sophistiqué alliant précision et efficacité. De la gestion des matières premières à la distribution finale, chaque étape est optimisée pour garantir un produit de haute qualité répondant aux exigences du marché. Prochainement, nous allons voir le processus de fabrication en détail.

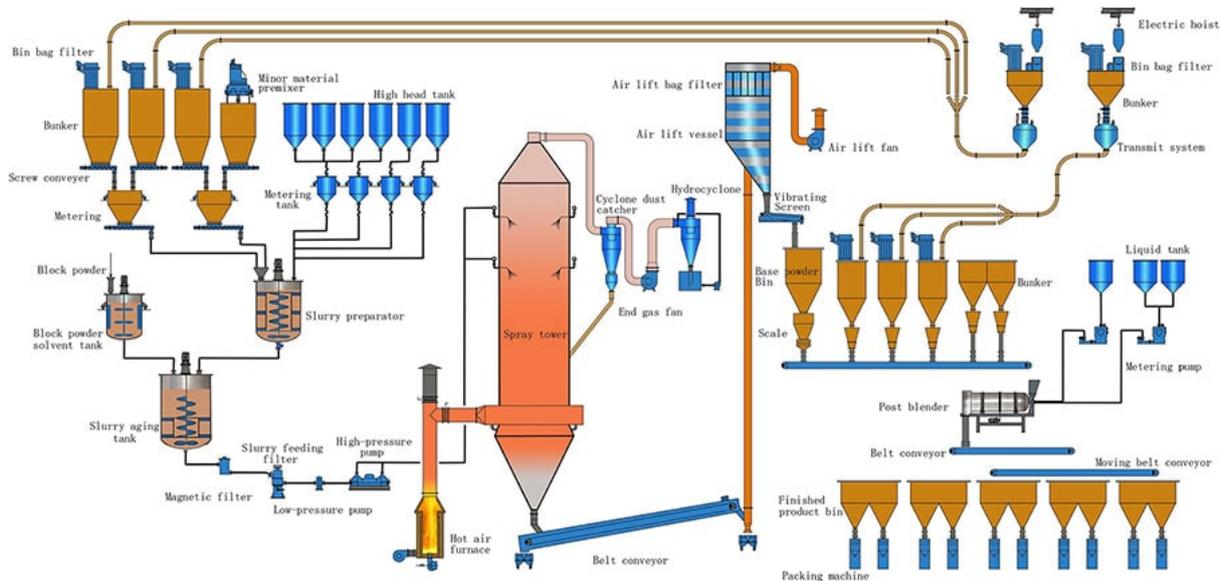


FIGURE 1.6 – Processus de fabrication du détergent en poudre

Stockage et manipulation des matières premières

La production de poudre détergente débute par une étape cruciale : la manipulation et le stockage appropriés des matières premières. Ces ingrédients essentiels se divisent en deux catégories distinctes : les solides et les liquides. Les matières premières solides, telles que le carbonate de soude ou le STPP, arrivent généralement sous forme de poudre, soigneusement emballée dans des sacs tissés. En revanche, les matières premières liquides, comme le LABSA ou la lessive de soude, nécessitent un stockage dans des réservoirs spécialisés. Pour préserver la qualité et l'intégrité de ces composants vitaux, des conditions de stockage spécifiques sont mises en place. Les matières solides trouvent leur place dans l'Unité 900, tandis que les liquides sont sécurisés dans l'Unité 800. Cette séparation et ces conditions de stockage minutieusement contrôlées constituent la base d'un processus de production efficace et de haute qualité.

Dosage des matières premières

Après le stockage, vient l'étape cruciale du dosage précis des matières premières. Ce processus, à la fois complexe et délicat, est entièrement automatisé grâce à un système API (Automate Programmable Industriel) de pointe, tout en restant sous la vigilance attentive d'un opérateur qualifié. Le cœur de ce système réside dans sa capacité à mesurer et doser avec une précision remarquable chaque composant, en suivant scrupuleusement des formules prédéfinies stockées dans sa mémoire. Cette automatisation joue un rôle clé dans la maintenance d'une cohérence et d'une précision constantes entre les différents lots de production, un aspect fondamental pour le contrôle de la qualité et la standardisation du produit final. L'utilisation de ce système automatisé apporte des avantages considérables : elle

permet une précision nettement accrue dans le dosage, réduit significativement les risques d'erreurs humaines, et assure une meilleure reproductibilité entre les lots. Ainsi, chaque sachet de détergent produit est le fruit d'un processus hautement contrôlé, garantissant aux consommateurs un produit d'une qualité constante et fiable.

Préparation et maturation de la pâte

Une fois les matières premières précisément dosées, elles entament un voyage fascinant pour devenir une pâte homogène, cœur du futur détergent. Ce processus se déroule en deux étapes principales, chacune jouant un rôle crucial dans la création du produit final. La première étape, le mélange initial, se déroule dans l'imposant mélangeur Crutcher. Pendant environ une heure, ce géant mécanique agite et mélange vigoureusement les matières avec de l'eau soigneusement traitée, créant ainsi la base de notre détergent. Mais le voyage de notre pâte ne s'arrête pas là. Elle est ensuite transférée dans un réservoir de maturation pour la deuxième étape cruciale : une période de "vieillesse" d'environ 30 minutes. Ce temps de repos, loin d'être oisif, permet à la pâte d'atteindre la consistance parfaite et de développer les propriétés spécifiques qui feront l'efficacité de notre détergent. C'est dans ce calme apparent que la magie opère, transformant un simple mélange en une pâte aux propriétés nettoyantes puissantes.

Phase de séchage par atomisation

La phase de séchage par atomisation marque un tournant décisif dans notre processus de fabrication, transformant la pâte liquide en la poudre sèche que nous connaissons tous. Ce processus, véritable prouesse technologique, se déroule en trois actes captivants. Tout d'abord, la pâte maturée est pompée vers les hauteurs de la tour de séchage, tel un ascenseur vers les cieux. Là-haut, des buses la pulvérisent en une myriade de fines gouttelettes, créant un brouillard de détergent en devenir. C'est alors qu'intervient la chaleur : un four génère un courant d'air chaud qui enveloppe ces gouttelettes. Dans leur chute maîtrisée, elles sont déshydratées, se transformant comme par magie en granulés de poudre. Enfin, ces granulés, désormais secs, sont recueillis au pied de la tour, tels des trésors à la fin d'une chasse. Ils constituent alors un produit semi-fini, prêt pour les étapes finales de personnalisation. Cette étape de séchage est bien plus qu'une simple transformation physique : elle facilite la manipulation et l'emballage du produit, tout en ouvrant la porte à une personnalisation ultérieure selon les besoins spécifiques de chaque gamme de détergent. C'est ici que la science rencontre l'art de la fabrication du détergent, créant un produit à la fois efficace et adapté aux exigences variées des consommateurs.

Processus de filtration

Après avoir traversé la tour de séchage, les granulés de détergent entament une nouvelle étape cruciale : le processus de filtration. Cette phase, bien que moins spectaculaire que l'atomisation, joue un rôle déterminant dans la qualité finale du produit. Son objectif principal est d'une précision chirurgicale : éliminer les résidus de poussière fine et les particules surdimensionnées qui auraient pu se former durant le séchage. L'importance de cette étape ne saurait être sous-estimée, car elle garantit que la poudre de base répond scrupuleusement aux normes de qualité les plus élevées et reste exempte de toute impureté. Le processus de filtration produit deux résultats distincts mais tout aussi importants. D'une part, l'air utilisé durant le séchage, maintenant filtré et refroidi, est évacué en toute sécurité, respectant ainsi les normes environnementales. D'autre part, la poudre de base, désormais parfaitement filtrée, est prête à entamer la prochaine étape de sa transformation. C'est grâce à cette attention méticuleuse aux détails que chaque grain de notre détergent est porteur de qualité et d'efficacité.

Dosage et mélange post-tour

La phase post-tour marque l'étape finale de personnalisation du détergent, où le produit acquiert ses caractéristiques uniques. C'est ici que la poudre de base se transforme en un produit fini aux propriétés spécifiques. Des additifs soigneusement sélectionnés, tels que des enzymes puissantes, des parfums envoûtants et des mouchetures colorées attrayantes, sont introduits dans le mélange. Ces éléments délicats, qui ne supporteraient pas le processus de séchage par atomisation, sont incorporés à ce stade précis pour préserver leur intégrité et leur efficacité. Le processus de mélange est une danse minutieuse, où chaque additif trouve sa place dans la poudre de base. Cette étape cruciale offre une flexibilité remarquable, permettant la création d'une vaste gamme de formules adaptées aux différents besoins et préférences du marché. Ainsi, chaque sachet de détergent devient un produit unique, conçu pour répondre aux attentes spécifiques des consommateurs.

Emballage et étiquetage

Une fois le détergent parfaitement formulé, il entame son voyage vers les consommateurs par l'étape d'emballage et d'étiquetage. Des machines d'emballage, alliant précision manuelle et efficacité automatisée, s'activent pour conditionner le produit dans une variété de formats : des sachets pratiques aux seaux robustes, en passant par des sacs tissés résistants. Chaque emballage devient alors un vecteur d'information, soigneusement étiqueté avec les données essentielles du produit. Cette étape ne se contente pas de protéger le détergent ; elle le prépare pour son grand départ vers les étagères des détaillants et, in fine, les foyers des utilisateurs finaux.

Stockage

Le stockage des détergents s'effectue dans un entrepôt sécurisé, conçu pour assurer une qualité optimale. Les produits emballés sont systématiquement organisés sur des palettes, chacune étant munie d'une identification unique dans un système de gestion des stocks avancé. L'environnement de l'entrepôt est rigoureusement contrôlé, garantissant des conditions idéales pour préserver l'intégrité et l'efficacité de chaque paquet de détergent.

Distribution

La phase finale du périple de notre détergent commence dans une zone de stockage marketing spécialisée. Sous la houlette experte du département marketing, cette étape orchestre la distribution des produits. Elle englobe non seulement la vente stratégique, mais aussi la logistique complexe du transport vers les clients. Avant chaque expédition, des contrôles qualité rigoureux sont effectués, garantissant que seuls les produits parfaits quittent l'entrepôt. Parallèlement, une documentation exhaustive est préparée, assurant une traçabilité complète et une conformité aux normes en vigueur.

Gestion des retours

La gestion des retours constitue un pilier essentiel du processus de production, jouant un rôle crucial dans la satisfaction client et le maintien de standards de qualité élevés. Face à une réclamation, Henkel déploie un protocole d'évaluation minutieux, analysant la gravité du problème et la quantité de produit concernée. Dans les cas les plus sérieux, une décision de rappel total du marché peut être prise. Pour les retours de moindre envergure, le client est invité à éliminer le produit et bénéficie d'un remboursement. Les retours plus conséquents sont rapatriés chez Henkel pour une batterie de tests approfondis. Selon les résultats obtenus, le produit est soit intégré dans un processus de recyclage, soit éliminé de manière responsable. Cette approche proactive permet non seulement de résoudre les problèmes immédiats, mais

aussi d'identifier les points d'amélioration potentiels dans la chaîne de production, alimentant ainsi un cycle d'amélioration continue.

Cartographie de la chaîne de valeur (Value Stream Mapping)

La cartographie de la chaîne de valeur, également connue sous le nom de VSM (Value Stream Mapping), est une technique essentielle de la production allégée (lean manufacturing). Cette méthode puissante permet d'analyser, de concevoir et de gérer efficacement le flux de matériaux et d'informations nécessaires pour livrer un produit au client final. En visualisant l'ensemble du processus de production, du fournisseur au consommateur, la VSM offre une vue d'ensemble claire et détaillée de toutes les étapes, qu'elles ajoutent de la valeur ou non.

Dans le contexte de la production de détergent en poudre chez Henkel, la cartographie de la chaîne de valeur joue un rôle crucial. Elle permet d'identifier les goulots d'étranglement, les retards et les inefficacités potentielles tout au long du processus de fabrication. Cette visualisation aide à optimiser le flux de production, à réduire les temps d'attente et à améliorer la qualité globale du produit.

La figure ci-dessous présente la cartographie de la chaîne de valeur pour le processus de production de détergent en poudre chez Henkel. Cette représentation visuelle englobe toutes les étapes que nous avons décrites précédemment, depuis la réception des matières premières jusqu'à la distribution du produit fini, en passant par les phases cruciales de mélange, de séchage et d'emballage.

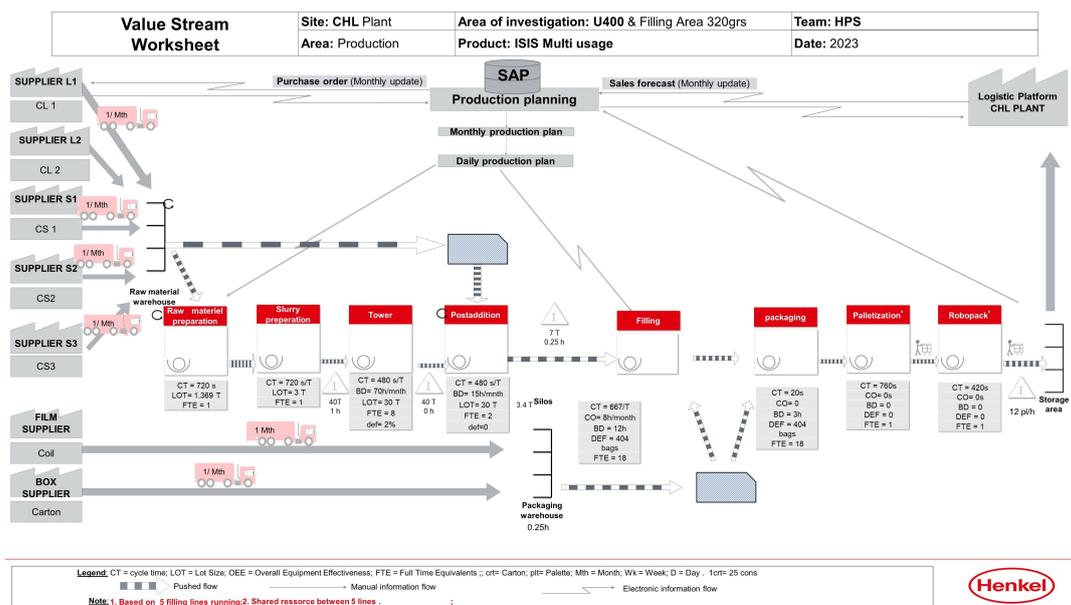


FIGURE 1.7 – Cartographie de la chaîne de valeur (VSM) du processus de fabrication de détergent en poudre chez Henkel (une autres figure plus claire est ajoutée au annexe)

Cette cartographie nous permet de visualiser clairement le flux de matériaux et d'informations à travers l'ensemble du processus de production. Elle met en évidence les interactions entre les différentes étapes, les temps de cycle, les stocks intermédiaires et les points de contrôle qualité. En analysant cette carte, Henkel peut identifier les opportunités d'amélioration, réduire les temps de production, minimiser les gaspillages et optimiser l'ensemble de la chaîne de valeur pour une production plus efficace et plus réactive aux besoins du marché.

1.4 Identification et analyse de la problématique

1.5 Analyse des Défauts

Un défaut est une déviation par rapport aux normes de qualité attendues d'un produit ou d'un service. Il peut se manifester sous différentes formes et à divers stades du processus de production ou de livraison. Les défauts peuvent affecter l'apparence, la fonctionnalité, la sécurité ou la conformité d'un produit, et peuvent résulter de plusieurs facteurs, y compris des erreurs humaines, des défaillances de machines, des matières premières de mauvaise qualité ou des conditions de production inadéquates.

Contexte et Importance de l'Analyse des Défauts

- Importance de l'analyse des défauts dans l'industrie

L'analyse des défauts est une composante essentielle du contrôle de qualité dans toute industrie. Dans le contexte de la production de détergent en poudre chez Henkel Algérie, cette analyse revêt une importance particulière pour garantir que les produits finis répondent aux normes de qualité et aux attentes des clients. L'identification et l'élimination des défauts permettent non seulement de maintenir la satisfaction des clients, mais aussi de réduire les coûts de production et de préserver la réputation de la marque.

- Impact des défauts sur le contrôle de qualité

Le contrôle de qualité est un processus crucial dans toute industrie, car il garantit que les produits ou services produits répondent aux normes requises. Les défauts dans le processus de production peuvent grandement affecter la qualité du produit final, entraînant l'insatisfaction des clients et la perte de revenus. Il est essentiel d'identifier et d'éliminer les défauts pour atteindre un contrôle de qualité efficace. Différents acteurs ont des points de vue différents sur l'impact des défauts sur le contrôle de la qualité, mais le consensus est que les défauts doivent être éliminés pour garantir des produits de qualité.

- **Impact sur la satisfaction des clients :** Les défauts dans le processus de production peuvent mener à des produits de mauvaise qualité, affectant ainsi la satisfaction des clients. Les consommateurs s'attendent à ce que les produits répondent à leurs besoins et exigences, et les défauts peuvent entraîner des échecs de produit, ce qui peut être frustrant et décevant pour les clients. Par exemple, une voiture avec des freins défectueux peut être dangereuse et causer des accidents, affectant la satisfaction des clients envers la marque. Par conséquent, éliminer les défauts est essentiel pour assurer la satisfaction des clients, ce qui peut entraîner une augmentation des ventes et des revenus.

- **Impact sur les coûts de production :** Les défauts peuvent augmenter les coûts de production, car ils nécessitent des ressources supplémentaires pour être identifiés et rectifiés. Par exemple, si un produit présente un défaut de fabrication, il devra être retravaillé ou mis au rebut, entraînant des coûts supplémentaires. Ces coûts peuvent s'accumuler et affecter la rentabilité de l'entreprise. Par conséquent, éliminer les défauts peut réduire les coûts de production, conduisant à une rentabilité accrue.

- **Impact sur la réputation de la marque :** Les défauts peuvent affecter la réputation de la marque, surtout s'ils sont répandus. Les clients associent les marques à la qualité, et les défauts peuvent entraîner des perceptions négatives de la marque. Par exemple, si une marque de smartphone a une histoire de batteries défectueuses, les clients peuvent éviter

d'acheter les produits de cette marque, affectant ainsi sa réputation. Par conséquent, éliminer les défauts peut protéger la réputation de la marque, conduisant à une fidélité accrue des clients et à une augmentation des revenus.

The Impact of Defects on Quality Control

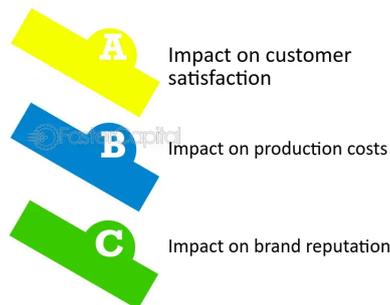


FIGURE 1.8 – L'impact des défaut sur la qualité [1]

Impact des défauts sur la performance de l'entreprise

Les défauts sont des écarts par rapport aux normes de qualité attendues d'un produit ou service. Ils peuvent se produire à n'importe quelle étape du processus de production ou de livraison et peuvent avoir un impact significatif sur la performance de l'entreprise. Les défauts peuvent affecter la satisfaction des clients, la réputation, la rentabilité et la compétitivité de l'entreprise. Ils peuvent également entraîner une augmentation des coûts, des déchets, des reprises, des retards et des responsabilités juridiques. Par conséquent, il est important pour les entreprises de prévenir, détecter et corriger les défauts le plus tôt et le plus efficacement possible.

1. **Coût de la qualité (COQ) :** C'est le coût total de l'assurance et du maintien de la qualité d'un produit ou service. Il comprend le coût de la prévention, de l'évaluation, des défaillances internes et des défaillances externes. Les coûts de prévention sont les coûts de planification et de mise en œuvre des activités de qualité pour prévenir les défauts. Les coûts d'évaluation sont les coûts d'inspection et de test du produit ou service pour s'assurer qu'il respecte les normes de qualité. Les coûts de défaillance interne sont les coûts de correction des défauts avant que le produit ou service n'atteigne le client. Les coûts de défaillance externe sont les coûts de correction des défauts après que le produit ou service a atteint le client. Le COQ peut être utilisé pour identifier les sources et les causes des défauts, et pour évaluer l'efficacité et l'efficacité des activités de qualité. Un COQ élevé indique un niveau élevé de défauts et un faible niveau de qualité. Un COQ faible indique un faible niveau de défauts et un haut niveau de qualité.
2. **Densité des défauts (DD) :** C'est le nombre de défauts par unité de production ou d'activité. Il peut être exprimé comme le rapport entre le nombre de défauts et la taille, le volume ou la complexité du produit ou service. Par exemple, la densité des défauts d'un produit logiciel peut être calculée comme le nombre de défauts par millier de lignes de code (KLOC). La densité des défauts peut être utilisée pour comparer la qualité de différents produits ou services, ou pour surveiller la qualité d'un produit ou service au fil du temps. Une densité de défauts élevée indique un niveau élevé de défauts et un faible niveau de qualité. Une densité de défauts faible indique un faible niveau de défauts et un haut niveau de qualité.
3. **Taux de défauts (DR) :** C'est le pourcentage d'unités ou de transactions défectueuses dans un échantillon ou une population. Il peut être calculé comme

le rapport entre le nombre d'unités ou de transactions défectueuses et le nombre total d'unités ou de transactions. Par exemple, le taux de défauts d'un processus de fabrication peut être calculé comme le nombre de produits défectueux divisé par le nombre de produits fabriqués. Le taux de défauts peut être utilisé pour estimer la probabilité de survenance de défauts dans un produit ou service, ou pour évaluer la performance d'un processus ou d'un fournisseur. Un taux de défauts élevé indique un niveau élevé de défauts et un faible niveau de qualité. Un taux de défauts faible indique un faible niveau de défauts et un haut niveau de qualité.

En prenant ces mesures, les entreprises peuvent améliorer leur qualité et réduire leurs défauts, et ainsi améliorer leur performance et leur compétitivité sur le marché. Les défauts ne sont pas seulement un coût, mais aussi une opportunité d'amélioration et d'innovation. En apprenant de leurs défauts, les entreprises peuvent créer de la valeur pour leurs clients et parties prenantes, et atteindre l'excellence dans leurs produits et services.

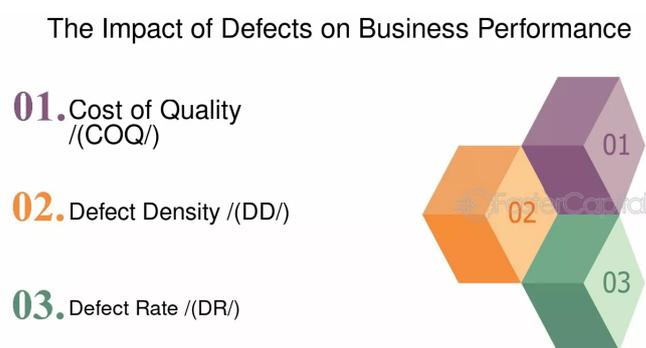


FIGURE 1.9 – L'impact des défaut sur la performances der l'entreprise [1]

Impact des défauts et des retours sur la satisfaction des clients

Les défauts et les retours peuvent avoir un impact significatif sur la satisfaction des clients, car ils affectent directement la qualité globale d'un produit ou d'un service. Lorsque les clients reçoivent des articles défectueux ou rencontrent des problèmes avec leur achat, cela entraîne non seulement de la frustration, mais aussi une érosion de la confiance dans la marque. Dans le marché concurrentiel d'aujourd'hui, où les clients ont de nombreuses options à choisir, assurer la satisfaction des clients est crucial pour que les entreprises prospèrent.

Du point de vue du client, les défauts et les retours peuvent être très gênants et chronophages. Imaginez acheter un nouveau smartphone pour découvrir qu'il a un écran défectueux ou un appareil photo défectueux. Non seulement cela perturbe la routine quotidienne du client, mais cela crée également une perception négative de la marque. Les clients s'attendent à ce que les produits fonctionnent parfaitement dès la sortie de la boîte, et toute déviation de cette attente peut entraîner une déception et une insatisfaction.

D'un autre côté, les défauts et les retours ont également un impact financier sur les entreprises. Les entreprises encourent des coûts associés au remplacement ou à la réparation des produits défectueux, aux frais d'expédition pour les retours et à la perte potentielle de ventes futures en raison de la détérioration de la réputation. De plus, la gestion des réclamations des clients et la gestion des retours nécessitent des ressources supplémentaires telles que des représentants du service client et du personnel logistique. Ces coûts peuvent rapidement s'accumuler et réduire les marges bénéficiaires d'une entreprise.

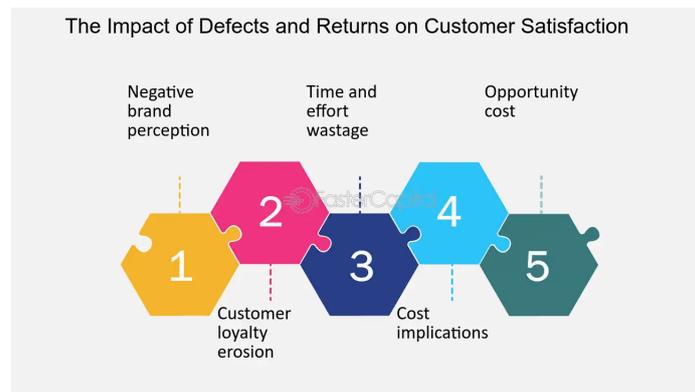


FIGURE 1.10 – L'Impact des Défauts et des Retours sur la Satisfaction des Clients - Coût Tangible de la Qualité : Assurer la Satisfaction des Clients [1]

Types de défauts dans la production de détergent en poudre

1. **Défauts de scellage** : Ces défauts se produisent lorsque le sachet de détergent n'est pas correctement scellé. Cela peut entraîner des fuites, une contamination du produit ou une exposition à l'humidité, ce qui peut altérer la qualité du détergent. Les causes courantes incluent une mauvaise adhésion des matériaux, des températures de scellage inappropriées ou des machines de scellage défectueuses.



FIGURE 1.11 – figure représentatif des défauts de scellage

2. **Défauts physiques et visuels** : Ces défauts incluent des anomalies telles que des déformations du sachet, des déchirures, des perforations ou des marques visibles qui peuvent affecter l'attrait visuel et la perception de qualité du produit.



FIGURE 1.12 – Figures représentatif des défauts visuelle

3. **Défauts d'impression** : Ces défauts se produisent lorsque les informations imprimées sur le sachet, telles que la date d'expiration ou les instructions d'utilisation, sont incorrectes, incomplètes ou illisibles. Cela peut être dû à des problèmes d'alignement, une mauvaise qualité de l'encre, des erreurs de programmation des machines d'impression ou des conditions environnementales défavorables.

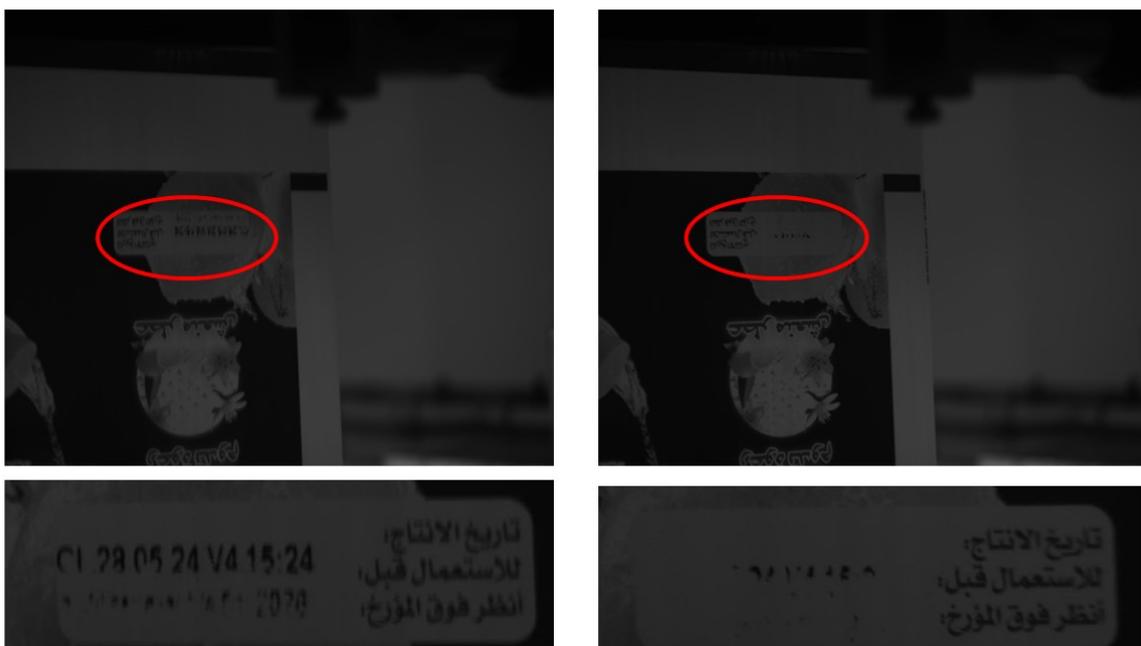


FIGURE 1.13 – défautes d'impression (étiquetages)

1.5.1 Analyse des Défauts chez Henkel

Dans le cadre de notre étude, nous avons collecté des données d'inspection de Henkel, comprenant des informations provenant de différentes régions, le nom des clients, le type de produit, et le nombre d'unités défectueuses. Les résultats montrent que le taux de défaut calculé est de 2,8 % de la production générale, dont seulement 0,5 % est détecté au niveau de l'usine. Le reste des défauts est identifié à partir des feuilles de réclamation et des résultats des inspections auprès des consommateurs.

Cependant, l'administration de Henkel estime que ce résultat est sous-évalué, car il n'y a pas de méthode fiable pour détecter la vraie valeur des défauts au niveau de l'usine. De plus, les clients ne se plaignent généralement pas lorsque le nombre d'unités défectueuses est faible.

1.5.2 Analyse des coûts de retour

Pour illustrer l'impact financier des retours de produits défectueux, nous présentons ci-dessous les coûts de retour en pourcentage par rapport au prix de vente pour différents formats de produits :

- **Sachet de 300g** : Le coût de retour représente 84% du prix de vente.
- **Sachet de 750g** : Le coût de retour représente 84% du prix de vente.
- **Sachet de 1.5kg** : Le coût de retour représente 91% du prix de vente.

Ces pourcentages montrent l'importance de détecter et de corriger les défauts au niveau de l'usine pour minimiser les coûts associés aux retours.

1.6 Problématique

Le principal objectif de notre projet est de réduire le nombre de réclamations et de créer une méthode pour détecter les défauts au niveau de l'usine. L'implémentation d'un système de contrôle qualité automatisé et intelligent permettra de détecter ces défauts en temps réel, d'améliorer l'efficacité du processus de production et de réduire les coûts associés aux retours et aux réclamations des clients. Ce système s'appuiera sur des techniques avancées de vision industrielle et d'apprentissage automatique pour identifier les anomalies et garantir la conformité des produits aux normes de qualité.

1.6.1 Objectifs spécifiques du projet

1. **Développement d'un système de détection des défauts de scellage** : Utilisation de caméras thermiques pour identifier les défauts de scellage sur les sachets de détergent en temps réel. Ce système permettra de détecter les fuites et les scellages défectueux avant que les produits ne quittent la chaîne de production.
2. **Mise en place d'un système de vérification de l'impression** : Développement d'un algorithme basé sur l'apprentissage automatique pour vérifier la qualité des impressions sur les sachets, y compris les dates d'expiration et les instructions d'utilisation. Ce

système identifiera les erreurs d'impression et garantira que toutes les informations sont lisibles et correctes.

1.7 Conclusion

L'implémentation de ce système de contrôle qualité automatisé permettra à Henkel Algérie de maintenir des standards de qualité élevés, d'optimiser les processus de production et de réduire les coûts associés aux défauts et aux retours. En garantissant des produits de haute qualité, Henkel renforcera sa position sur le marché et améliorera la satisfaction de ses clients.

Chapitre 2

Fondements Théoriques et Revue de Littérature

Ce chapitre vise à explorer les défauts courants dans la production de sachets de détergents en poudre, en se concentrant particulièrement sur les problèmes de soudure et d'étiquetage. Nous examinerons également les différentes techniques existantes pour traiter ces problèmes, afin de fournir un contexte solide pour notre étude chez Henkel. Nous commencerons par une vue d'ensemble des défauts de production des sachets, puis nous nous focaliserons sur les problèmes spécifiques de soudure et d'étiquetage. Ensuite, nous explorerons les techniques existantes pour traiter ces problèmes, en comparant les méthodes traditionnelles et avancées.

2.1 Défauts de production des pouches (sachets)

La production de sachets de détergents en poudre est un processus complexe qui peut être sujet à divers types de défauts, affectant la qualité et l'intégrité des produits finis. Une compréhension approfondie de ces défauts est cruciale pour améliorer les processus de production et garantir la satisfaction des consommateurs. Selon l'article [24], les défauts courants observés dans ce domaine comprennent les défauts de soudure, d'étiquetage, de remplissage, de matériau et de forme.

Défauts de soudure

Les défauts de soudure constituent l'une des principales préoccupations dans la production de sachets. Parmi les défauts de soudure, on trouve les fuites et les scellages incomplets. Les fuites surviennent lorsque les soudures ne sont pas complètement hermétiques, permettant ainsi au contenu de s'échapper. Cela peut être dû à une mauvaise application de la température, une pression inadéquate ou des contaminations au niveau de la zone de soudure. Les scellages incomplets, quant à eux, se produisent lorsque les soudures sont partiellement formées, entraînant une mauvaise fermeture du sachet et compromettant ainsi l'intégrité du produit.

Défauts d'étiquetage

Les défauts d'étiquetage représentent un autre défi majeur dans la production de sachets de détergents en poudre. Les impressions illisibles sont fréquentes lorsque les informations imprimées, telles que les dates de péremption, ne sont pas claires ou sont difficilement lisibles en raison de problèmes d'encre ou de mauvaise qualité d'impression. Les erreurs de date, incluant des dates incorrectes ou manquantes, créent des problèmes de traçabilité et de conformité, ce qui peut entraîner des rappels de produits et des non-conformités

réglementaires.

Défauts de remplissage

Les défauts de remplissage se manifestent par des sous-remplissages ou des sur-remplissages. Le sous-remplissage se produit lorsque les sachets contiennent moins de produit que la quantité spécifiée, pouvant résulter de calibrages incorrects des machines de remplissage. Le sur-remplissage, en revanche, entraîne des sachets contenant plus de produit que prévu, ce qui peut poser des problèmes de scellage et de transport, ainsi qu'une augmentation des coûts de production.

Défauts de matériau

Les défauts de matériau incluent des déchirures et des perforations. Les déchirures surviennent lorsque le matériau des sachets se déchire, compromettant ainsi la protection du produit et augmentant le risque de contamination. Les perforations involontaires, ou trous, peuvent se produire lors de la manipulation ou du transport, causant des fuites ou une exposition à l'air, ce qui affecte la qualité du produit.

Défauts de forme

Enfin, les défauts de forme sont également courants dans la production de sachets. Les plis apparaissent souvent lorsque les sachets présentent des plis non souhaités qui peuvent affecter l'intégrité de la soudure et l'apparence générale du produit. Les déformations des sachets, qui ne conservent pas leur forme prévue, posent des problèmes lors de l'emballage et du stockage, rendant difficile l'empilement ou l'affichage des produits sur les étagères. Ces défauts de production ont un impact significatif sur la qualité des sachets de détergents en poudre. La prochaine section se focalisera sur les causes spécifiques et les conséquences des défauts de soudure et d'étiquetage, ainsi que sur les techniques existantes pour les traiter.

2.2 Mise au point sur les problèmes de soudure et d'étiquetage

2.2.1 Défauts de soudure

Les défauts de soudure sont des problèmes critiques dans la production de sachets de détergents en poudre. Ils peuvent compromettre l'intégrité des produits et entraîner des pertes importantes. Cette section explore les différents types de défauts de soudure, leurs causes et leurs conséquences.

Types de défauts de soudure

Soudures incomplètes

Les soudures incomplètes se produisent lorsque la soudure n'est pas correctement formée sur toute la longueur du joint. Cela peut laisser des sections non scellées, permettant aux produits de s'échapper. Ce type de défaut est souvent le résultat d'une application inadéquate de la chaleur ou de la pression pendant le processus de soudage.

Soudures faibles

Les soudures faibles sont caractérisées par des joints soudés qui manquent de robustesse et peuvent se briser facilement sous une pression ou une manipulation normale. Les soudures faibles peuvent être dues à une température de soudage insuffisante, à une pression de soudage inappropriée ou à des matériaux de mauvaise qualité utilisés pour les sachets.

Contamination de la zone de soudure

La contamination de la zone de soudure se produit lorsque des particules étrangères, telles que de la poussière, des résidus de produit ou des huiles, sont présentes dans la zone où la soudure est effectuée. Ces contaminants peuvent empêcher une adhésion adéquate des surfaces soudées, entraînant des soudures défectueuses et des fuites potentielles.

Causes des défauts de soudure

Les défauts de soudure peuvent être attribués à plusieurs causes, notamment :

- **Température de soudage incorrecte** : Une température de soudage trop basse peut entraîner des soudures faibles, tandis qu'une température trop élevée peut endommager le matériau du sachet.
- **Pression inadéquate** : Une pression insuffisante ou excessive pendant le soudage peut causer des soudures incomplètes ou faibles.
- **Contamination du matériau** : La présence de particules étrangères dans la zone de soudure peut empêcher une bonne adhésion et causer des soudures défectueuses.
- **Usure de l'équipement** : Les équipements de soudage usés ou mal entretenus peuvent ne pas fonctionner correctement, entraînant des soudures de mauvaise qualité.

Conséquences des défauts de soudure

Les conséquences des défauts de soudure peuvent être graves et variées :

- **Fuites de produit** : Les sachets avec des soudures défectueuses peuvent fuir, entraînant une perte de produit et une contamination possible.
- **Contamination du produit** : Les soudures incomplètes ou faibles peuvent permettre à des contaminants extérieurs d'entrer dans les sachets, compromettant la qualité et la sécurité du produit.
- **Durée de conservation réduite** : Les soudures défectueuses peuvent affecter la durée de conservation des produits en permettant une entrée d'air ou d'humidité, ce qui peut accélérer la dégradation du détergent en poudre.

En conclusion, les défauts de soudure dans la production de sachets de détergents en poudre sont un problème majeur qui nécessite une attention particulière pour garantir la qualité et l'intégrité des produits finis. La prochaine section se concentrera sur les défauts d'étiquetage, leurs types, causes et conséquences.

2.2.2 Défauts d'étiquetage : types, causes et conséquences

Les défauts d'étiquetage sont un problème crucial dans la production de sachets de détergents en poudre, car ils affectent directement la lisibilité des informations importantes pour les consommateurs et les régulateurs. Cette section explore les différents types de défauts d'étiquetage, leurs causes et leurs conséquences.

Types de défauts d'étiquetage

Impressions floues ou illisibles

Les impressions floues ou illisibles se produisent lorsque les informations imprimées sur les sachets, telles que les dates de péremption ou les numéros de lot, ne sont pas clairement

visibles. Cela peut être dû à des problèmes d'encre, d'impression ou de qualité du matériel utilisé.

Erreurs de date ou de code

Les erreurs de date ou de code surviennent lorsque les informations imprimées sont incorrectes, manquantes ou mal codées. Cela inclut les dates de péremption erronées, les numéros de lot incorrects ou les codes-barres mal formés, entraînant des problèmes de traçabilité et de conformité.

Placement incorrect de l'étiquette

Le placement incorrect de l'étiquette se produit lorsque les étiquettes ne sont pas alignées correctement sur le sachet. Cela peut rendre les informations difficiles à lire ou à scanner, et affecter l'apparence générale du produit.

Causes des défauts d'étiquetage

Les défauts d'étiquetage peuvent être attribués à plusieurs causes, notamment :

- **Problèmes d'encre ou d'impression** : La qualité de l'encre utilisée, les paramètres d'impression incorrects ou l'usure des imprimantes peuvent entraîner des impressions floues ou illisibles.
- **Erreurs de programmation** : Des erreurs dans la programmation des imprimantes ou des systèmes de codage peuvent entraîner des informations incorrectes ou manquantes sur les étiquettes.
- **Mauvais alignement de l'équipement** : Un alignement incorrect des machines d'étiquetage peut provoquer un placement incorrect des étiquettes sur les sachets.

Conséquences des défauts d'étiquetage

Les conséquences des défauts d'étiquetage peuvent être graves et variées :

- **Confusion des consommateurs** : Des étiquettes incorrectes ou illisibles peuvent induire les consommateurs en erreur, affectant leur capacité à utiliser correctement le produit et à suivre les instructions de sécurité.
- **Problèmes de traçabilité** : Les erreurs de date ou de code peuvent rendre difficile le suivi des lots de production, ce qui est crucial en cas de rappel de produit ou de problème de qualité.
- **Non-conformité réglementaire** : Les défauts d'étiquetage peuvent entraîner des non-conformités avec les réglementations en vigueur, exposant l'entreprise à des sanctions ou à des rappels de produits coûteux.

2.2.3 Techniques existantes pour traiter les problèmes de soudure

Inspection visuelle manuelle

L'inspection visuelle manuelle est une méthode courante où un opérateur qualifié examine chaque soudure à l'œil nu pour identifier les défauts tels que les soudures incomplètes, faibles ou contaminées. Cette méthode, bien qu'efficace pour détecter certains défauts évidents, est limitée par la subjectivité et la variabilité humaine. Elle peut également être lente et inefficace pour les grandes productions en raison de sa nature laborieuse.

Tests de pression

Les tests de pression consistent à appliquer une pression interne au sachet pour vérifier l'intégrité des soudures. Cette technique permet de détecter les soudures faibles ou incomplètes en évaluant leur capacité à résister à une certaine pression. Cependant, ces tests peuvent être destructifs, car ils impliquent souvent la rupture des sachets défectueux, entraînant ainsi des pertes de produit.

Tests de fuite par immersion

Les tests de fuite par immersion impliquent de plonger les sachets dans l'eau et d'appliquer une pression pour observer la présence de bulles d'air s'échappant des soudures. Cette méthode est efficace pour détecter les fuites, mais elle est également destructrice et nécessite un traitement ultérieur des sachets pour les sécher. De plus, elle peut être moins adaptée à une inspection rapide et continue en ligne.

Systèmes de vision industrielle

Les systèmes de vision industrielle utilisent des caméras haute résolution et des logiciels sophistiqués pour inspecter les soudures en temps réel. Ces systèmes sont capables de détecter automatiquement les défauts de soudure, y compris les soudures incomplètes, faibles et les contaminations. Ils offrent une grande précision et peuvent être intégrés directement dans la ligne de production, permettant une inspection continue et rapide sans intervention humaine. Cette technologie réduit les erreurs humaines et augmente l'efficacité de la détection des défauts.

Radiographie industrielle

La radiographie industrielle utilise des rayons X pour inspecter les soudures à l'intérieur des sachets. Cette technique permet de détecter des défauts internes invisibles à l'œil nu, tels que des inclusions ou des porosités dans la zone de soudure. Bien que très efficace, la radiographie industrielle nécessite des équipements spécialisés et des mesures de sécurité appropriées pour protéger les opérateurs des radiations. Elle offre une solution non destructive pour l'inspection des soudures. [25]

Tests ultrasoniques

Les tests ultrasoniques utilisent des sondes ultrasonores pour évaluer l'intégrité des soudures en mesurant les échos renvoyés par les discontinuités internes. Cette technique non destructive est capable de détecter des défauts internes avec une grande précision. Les tests ultrasoniques sont particulièrement utiles pour les matériaux plastiques et métalliques utilisés dans les sachets de détergents en poudre. Cette méthode est rapide et peut être automatisée pour une inspection en ligne. [26, 27]

2.2.4 Techniques existantes pour traiter les problèmes d'étiquetage

Vérification visuelle manuelle

La vérification visuelle manuelle est une méthode où des opérateurs inspectent manuellement chaque étiquette pour s'assurer que les informations sont correctes et lisibles. Bien que cette méthode permette de détecter certains défauts, elle est sujette à des erreurs humaines et peut être inefficace pour des volumes de production élevés.

Utilisation de lecteurs de codes-barres

Les lecteurs de codes-barres sont utilisés pour scanner les codes-barres imprimés sur les sachets et vérifier leur lisibilité et leur exactitude. Cette méthode est efficace pour vérifier rapidement les codes-barres, mais elle ne détecte pas nécessairement d'autres types de défauts d'étiquetage comme les impressions floues ou les erreurs de date.

Systèmes de vision industrielle avec OCR (Reconnaissance Optique de Caractères)

Les systèmes de vision industrielle équipés de logiciels OCR utilisent des caméras haute résolution pour lire et vérifier automatiquement les informations imprimées sur les étiquettes. Ces systèmes peuvent détecter des défauts d'impression, des erreurs de date et des placements incorrects avec une grande précision. Ils offrent une inspection rapide et continue en ligne, améliorant ainsi l'efficacité et la fiabilité du contrôle qualité.

Technologies Laser

Les systèmes laser sont utilisés pour coder et marquer les emballages avec des informations telles que les dates de péremption et les numéros de lot. Ces systèmes offrent une grande précision et une durabilité des marquages, même sur des surfaces difficiles. Les lasers peuvent également vérifier la lisibilité des marquages en temps réel, assurant ainsi que les informations restent claires et lisibles tout au long de la chaîne d'approvisionnement.

2.2.5 Comparaison des différentes approches

Pour évaluer l'efficacité des différentes techniques de traitement des problèmes de soudure, il est utile de comparer les méthodes traditionnelles et avancées en fonction de plusieurs critères clés. Cette comparaison permet de mieux comprendre les avantages et les inconvénients de chaque approche et d'identifier la méthode la plus appropriée pour une production donnée.

Comparaison des différentes approches pour les problèmes de soudure

Le tableau (2.1) représente une comparaison des différentes techniques de traitement des problèmes de soudure en fonction de plusieurs critères clés.

TABLE 2.1 – Comparaison des différentes techniques de traitement des problèmes de soudure

Critère	Vision Industrielle	Tests de Pression	de	Tests de Fuite par Immersion	Vérification Manuelle
Précision et fiabilité	Très élevée	Moyenne		Moyenne	Moyenne
Vitesse d'inspection	Très élevée	Moyenne		Moyenne	Faible
Coûts d'implémentation	Élevés	Modérés		Modérés	Faibles
Coûts d'exploitation	Réduits à long terme	Modérés		Modérés	Élevés
Capacité de détection	Tous types de défauts	Fuites et pressions	et	Fuites	Visuels

Comparaison des différentes approches pour les problèmes d'étiquetage

Le tableau (2.2) compare les différentes techniques de traitement des problèmes d'étiquetage selon plusieurs critères essentiels.

TABLE 2.2 – Comparaison des différentes techniques de traitement des problèmes d'étiquetage

Critère	Vision Industrielle avec OCR	Lecteurs de Codes-Barres	Vérification Manuelle	Technologies Laser
Précision de la lecture	Très élevée	Élevée	Moyenne	Très élevée
Vitesse de traitement	Très élevée	Élevée	Faible	Très élevée
Capacité à gérer différents types d'étiquettes et d'impressions	Très élevée	Moyenne	Faible	Élevée
Flexibilité et adaptabilité	Très élevée	Moyenne	Faible	Élevée
Coûts d'implémentation	Élevés	Modérés	Faibles	Élevés

La vision industrielle est choisie pour traiter les problèmes de soudure et d'étiquetage en raison de sa précision élevée, de sa rapidité d'inspection en temps réel, et de sa capacité à détecter divers défauts. Bien que les coûts d'implémentation initiaux soient élevés, elle offre des coûts d'exploitation réduits à long terme grâce à l'automatisation, tout en étant flexible et adaptable aux différents types de produits et d'étiquettes.

2.3 vision industrielle

La vision industrielle, appelée "Machine Vision" par les Anglo-Saxons, est une discipline appliquée de la vision artificielle qui emploie le traitement numérique d'images pour automatiser des tâches. Elle utilise des moyens électroniques et informatiques pour doter des machines et robots de capacités sensorielles analogues à la vision humaine. Autrement dit, elle utilise des caméras comme capteurs pour détecter des défauts sur des objets manufacturés, inspecter des pièces en cours de fabrication, les compter, les trier, les classer, les mesurer, etc., à partir de leur apparence visuelle.

Il faut néanmoins se garder d'imaginer une analogie trop profonde entre les performances de la vision humaine, capable d'explorer la structure spatiale de son environnement et d'y reconnaître des objets complexes et variés, connus ou inconnus, en deux dimensions comme en trois, et les capacités bien plus modestes des systèmes de vision industriels [28].

La vision industrielle peut être définie comme l'application de la vision par ordinateur à des problématiques de production. Son principe est de doter les machines de production de la capacité de voir afin d'automatiser les tâches de contrôle qualité ou de contrôle de processus. Cette automatisation permet d'augmenter les performances et les cadences de production, de

rendre la production plus fiable, d'améliorer la qualité des produits, d'assurer leur traçabilité, et de garantir la sécurité [29].

Elle est utilisée dans l'industrie pour l'inspection des matériaux, la reconnaissance des objets, la reconnaissance des formes, l'analyse des composants électroniques, la reconnaissance des signatures, la reconnaissance optique des caractères et la reconnaissance de l'argent [30, 31]. Le processus de vision industrielle commence par une analyse détaillée des besoins et du projet, en développant une solution. La procédure commence par l'imagerie, puis l'analyse automatique de l'image et l'extraction d'informations [32, 33]. Le terme est le plus souvent associé à ces compétences dans le domaine de l'automatisation industrielle, bien qu'il soit également utilisé dans d'autres contextes tels que la sécurité et la navigation automobile.

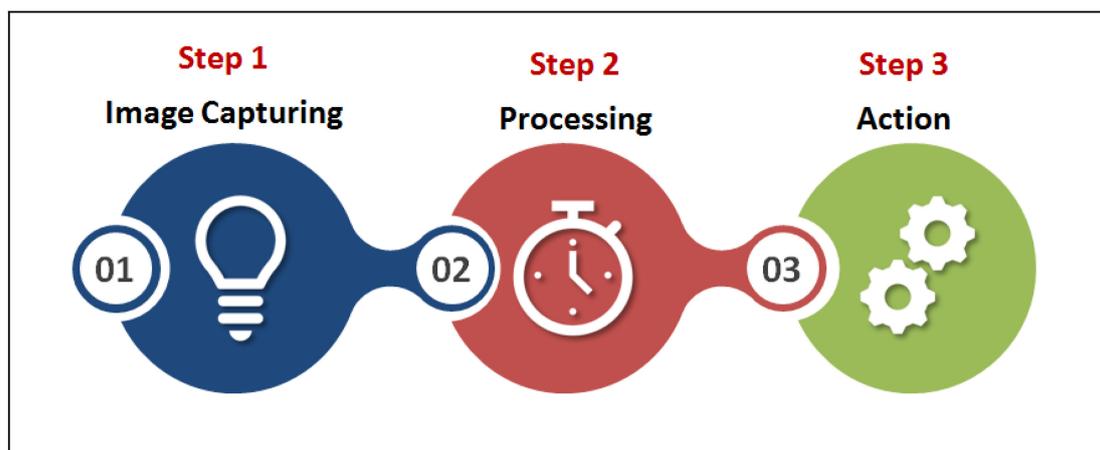


FIGURE 2.1 – Étapes de la machine vision

2.3.1 Historique et évolution

L'histoire et l'évolution de la vision par ordinateur, qui constitue la base de la vision industrielle, ont été marquées par des avancées significatives depuis ses débuts dans les années 1970. Initialement perçue comme une composante de la perception visuelle dans l'ambitieux programme visant à imiter l'intelligence humaine et à doter les robots de comportements intelligents, la vision par ordinateur a été considérée par les pionniers de l'intelligence artificielle et de la robotique comme un pas facile vers la résolution de problèmes plus complexes tels que le raisonnement de haut niveau et la planification [34].

Dès le départ, la distinction entre la vision par ordinateur et le traitement numérique des images existant déjà était claire : il s'agissait de récupérer la structure tridimensionnelle du monde à partir d'images pour avancer vers une compréhension complète de la scène. Les tentatives initiales de compréhension de la scène impliquaient l'extraction des contours et l'inférence de la structure 3D d'un objet ou d'un "monde de blocs" à partir de la structure topologique des lignes 2D, avec plusieurs algorithmes de marquage de lignes développés pendant cette période [34].

Dans les années 1980, l'attention s'est portée sur des techniques mathématiques plus sophistiquées pour l'analyse quantitative des images et des scènes. L'utilisation de pyramides d'images pour réaliser des tâches telles que le mélange d'images et la recherche de correspondances grossières à fines a commencé à se répandre. Des versions continuent de pyramides utilisant le concept de traitement en espace d'échelle ont également été

développées, et à la fin des années 1980, les ondelettes ont commencé à remplacer ou à compléter les pyramides d'images régulières dans certaines applications [35].

Cette évolution reflète une progression de la vision par ordinateur, allant de simples tâches de reconnaissance à des analyses visuelles profondes et complexes, et établissant les fondements sur lesquels la vision industrielle s'est construite. La vision industrielle a ensuite appliqué ces principes avancés au contexte spécifique de l'automatisation industrielle, transformant la manière dont la production, l'inspection de qualité, et le tri de produits sont réalisés dans les environnements de fabrication.

2.3.2 La Vision par Ordinateur dans l'Industrie 4.0

L'Industrie 4.0, également connue sous le nom de quatrième révolution industrielle, se réfère à l'intégration de technologies numériques avancées dans le secteur manufacturier. Elle englobe des concepts tels que l'Internet des Objets (IoT), les systèmes cyber-physiques et l'analyse des données. L'objectif principal de l'Industrie 4.0 est de créer des usines intelligentes capables de s'auto-optimiser, de s'adapter et d'apprendre de manière autonome.

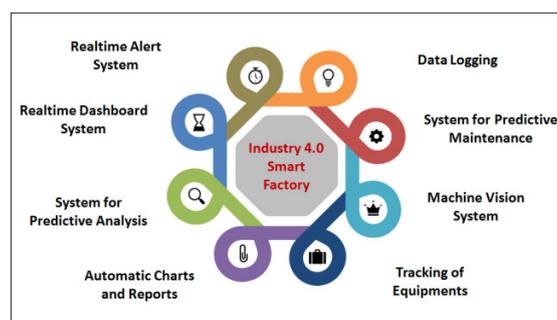


FIGURE 2.2 – intégration de la machine vision dans l'industrie 4.0 [2]

La vision par ordinateur joue un rôle crucial dans la mise en œuvre de l'Industrie 4.0 en fournissant des capacités avancées de capture et d'analyse de données visuelles. Voici comment elle s'intègre aux principaux composants de l'Industrie 4.0 :

1. **Assurance Qualité Automatisée** : En utilisant des caméras et des algorithmes de traitement d'image, la vision par ordinateur permet une inspection en temps réel des produits, garantissant que seuls les produits conformes aux normes de qualité quittent la chaîne de production.
2. **Maintenance Prédicative** : Les systèmes de vision peuvent détecter les signes précoces de défaillance des équipements, permettant ainsi une maintenance proactive avant que des pannes ne surviennent. Cela réduit les temps d'arrêt et les coûts de réparation.
3. **Guidage Robotique** : Les robots équipés de systèmes de vision par ordinateur peuvent naviguer dans des environnements complexes, assembler des composants avec précision et manipuler des objets de manière autonome.
4. **Optimisation des Processus** : En analysant les flux de production et en identifiant les goulots d'étranglement, les systèmes de vision par ordinateur aident à améliorer l'efficacité des processus. Les données visuelles peuvent être intégrées aux systèmes de gestion de la production pour une optimisation continue.

5. **Sécurité et Surveillance** : La vision par ordinateur est utilisée pour surveiller les environnements de travail et garantir la sécurité des travailleurs. Elle peut détecter les comportements dangereux et alerter les responsables en cas d'incident.

2.3.3 Avantages de la Vision par Ordinateur dans l'Industrie 4.0

1. **Amélioration de la Qualité** : Détection précise des défauts et des variations, assurant des normes de qualité élevées.
2. **Augmentation de la Productivité** : Automatisation des tâches répétitives et réduction des temps d'arrêt.
3. **Réduction des Coûts** : Prévention des erreurs coûteuses et réduction des besoins en main-d'œuvre pour les tâches d'inspection.
4. **Flexibilité Accrue** : Adaptation rapide aux changements dans les lignes de production et personnalisation des produits.

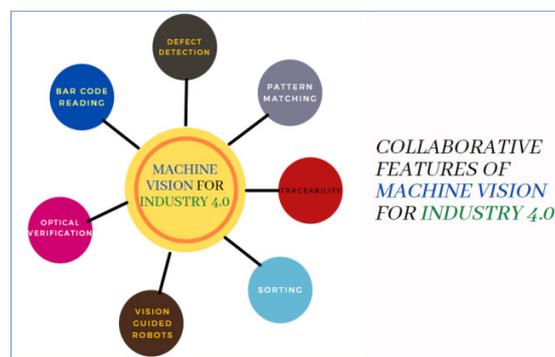


FIGURE 2.3 – utilisation de la machine vision dans l'industrie 4.0 [2]

2.4 Fondamentaux de l'Apprentissage Automatique

L'apprentissage automatique est une branche spécifique de l'intelligence artificielle qui enseigne à une machine comment apprendre, alors que l'intelligence artificielle (IA) en général est la science qui vise à émuler les capacités humaines [32]. Une méthode de l'IA appelée apprentissage automatique permet aux ordinateurs d'apprendre à partir de leurs expériences passées. Contrairement aux méthodes traditionnelles, les algorithmes d'apprentissage automatique ne reposent pas sur une équation prédéterminée comme modèle. Ils "apprennent" plutôt directement à partir des données en utilisant des techniques computationnelles. Au fur et à mesure que le volume d'exemples d'apprentissage augmente, les algorithmes s'améliorent de manière adaptative dans leurs fonctions.

Cet aspect de l'apprentissage automatique est crucial car il souligne l'importance de l'accumulation de données et de l'expérience dans l'amélioration des performances des modèles. Les algorithmes deviennent plus efficaces et précis à mesure qu'ils traitent davantage de données, ce qui les rend capables de réaliser des tâches de plus en plus complexes [33]. Cette capacité à apprendre à partir des données et à s'améliorer de manière incrémentale est au cœur de ce qui rend l'apprentissage automatique si puissant et révolutionnaire dans divers domaines, allant de la reconnaissance vocale et de l'image à la prise de décision dans les systèmes autonomes.

2.4.1 Catégorisation de l'Apprentissage Automatique

Les techniques d'apprentissage automatique se divisent principalement en quatre catégories : l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage semi-supervisé et l'apprentissage par renforcement. Chacune de ces catégories possède des caractéristiques uniques et des applications spécifiques, différenciées essentiellement par la façon dont les données sont utilisées pour entraîner les modèles.

2.4.1.1 Apprentissage supervisé

L'une des techniques d'apprentissage automatique est l'apprentissage supervisé, dans lequel nous formons le système d'apprentissage automatique à l'aide d'échantillons de données étiquetées puis nous l'observons prédire le résultat. Afin d'anticiper les événements futurs, les algorithmes d'apprentissage automatique supervisé utilisent des exemples étiquetés pour appliquer ce qu'ils ont appris dans le passé à de nouvelles données. La méthode d'apprentissage crée une fonction déduite pour prévoir les valeurs de sortie en examinant un ensemble de données d'apprentissage connues [35]. Après une formation adéquate, le système peut fournir des cibles pour toute nouvelle entrée. Afin d'identifier les failles et de corriger le système, le modèle si nécessaire peut également comparer ses résultats avec ceux qui sont corrects et prévus. Afin de comprendre les ensembles de données et d'apprendre à

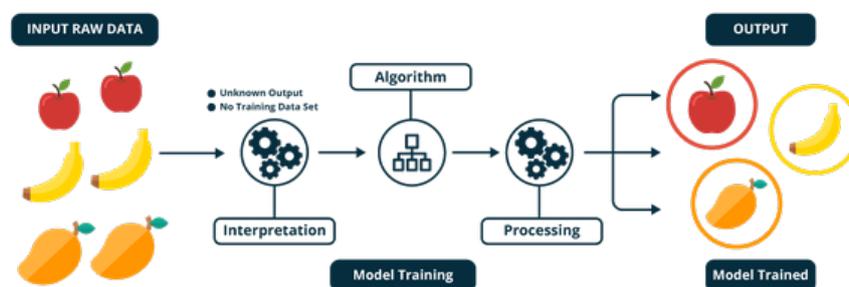


FIGURE 2.4 – apprentissage supervisé [3]

connaître chacun d'entre eux, le système construit un modèle à l'aide de données étiquetées. Après la formation et le traitement, le modèle est testé en recevant un échantillon de données pour voir s'il prédit la sortie avec précision ou non [36]. La mise en correspondance des données d'entrée et de sortie est l'objectif de l'apprentissage supervisé. Un élève qui apprend sous la supervision de l'enseignant est assimilé à l'apprentissage supervisé car il est basé sur la supervision. Le filtrage des spams est une illustration de l'apprentissage supervisé.

L'apprentissage supervisé peut être regroupé en deux catégories d'algorithmes :

Classification : Lorsque la variable de sortie est catégorique, c'est-à-dire qu'il existe deux classes telles que Oui-Non, Homme-Femme, Vrai-Faux, détection de spam, etc., des méthodes de classification sont utilisées. Plusieurs algorithmes de classification sont utilisés : Random Forest, Decision Trees, Logistic Regression, Support Vector Machines.

Régression : Les procédures de régression sont appliquées s'il existe une corrélation entre les

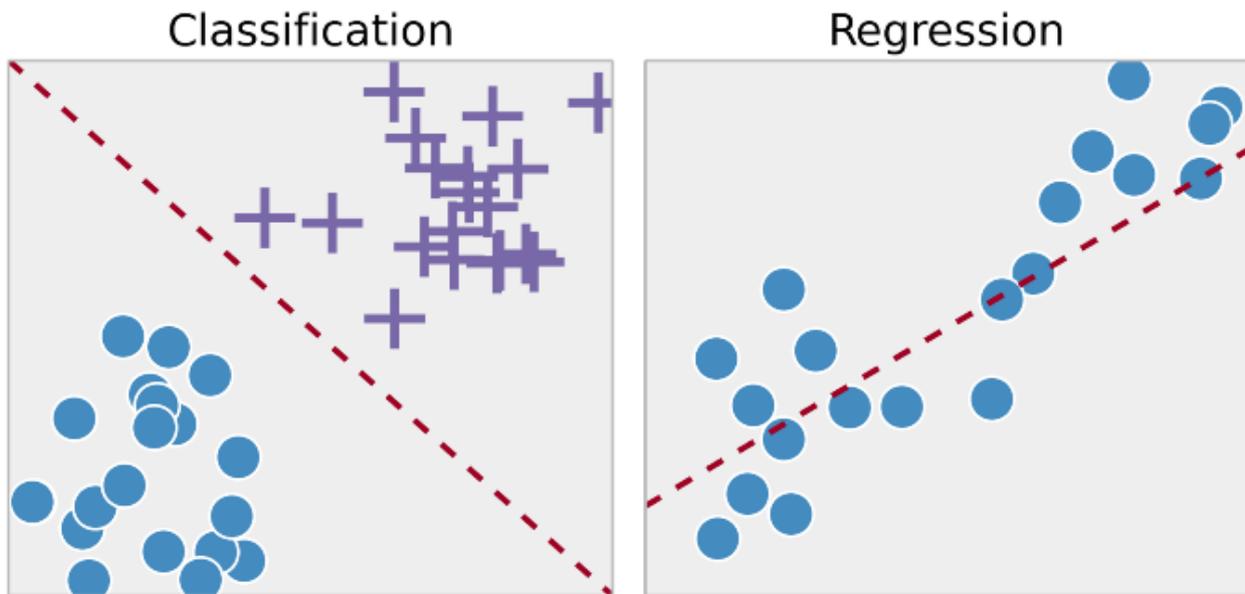


FIGURE 2.5 – Catégorie de l'apprentissage supervisé [4]

variables d'entrée et de sortie. Elles sont utilisées pour prévoir des variables continues telles que la météo, les tendances du marché, etc. Plusieurs algorithmes de régression bien connus qui relèvent de l'apprentissage supervisé sont énumérés ci-dessous : Régression linéaire, arbres de régression, régression non linéaire, régression linéaire bayésienne, régression polynomiale.

2.4.1.2 Apprentissage semi-supervisé :

L'apprentissage semi-supervisé est une méthode qui combine des éléments des approches supervisées et non supervisées. Cette technique est utilisée lorsque l'on dispose d'un grand volume de données non étiquetées et d'un petit ensemble de données étiquetées. L'idée principale est d'utiliser les informations fournies par les données étiquetées pour guider le processus d'apprentissage effectué sur les données non étiquetées. Cette approche permet de bénéficier des avantages de l'apprentissage supervisé, même lorsque les données étiquetées sont rares ou coûteuses à obtenir [37].

Par exemple, dans le domaine de la reconnaissance d'images, l'apprentissage semi-supervisé peut être utilisé pour améliorer la précision de la classification en utilisant un petit nombre d'images étiquetées pour aider à classer un grand ensemble d'images non étiquetées. Cela est particulièrement utile dans les situations où l'étiquetage des données nécessite une expertise spécifique ou est chronophage [38].

L'apprentissage semi-supervisé est donc particulièrement efficace dans les contextes où les données étiquetées sont insuffisantes pour entraîner un modèle robuste, mais où suffisamment de données non étiquetées sont disponibles pour enrichir et affiner le modèle d'apprentissage [39, 40].

2.4.1.3 Apprentissage non supervisé :

L'apprentissage non supervisé est un type d'apprentissage dans lequel un ordinateur recueille des informations sans aucune intervention humaine. La machine est formée à l'aide d'un

ensemble de données non étiquetées, non classées ou non catégorisées, et l'algorithme doit répondre indépendamment à ces données [37]. L'objectif de l'apprentissage non supervisé est de réorganiser les données entrantes en nouvelles caractéristiques ou en une collection d'objets ayant des liens entre eux. Il n'y a pas de résultat prédéfini dans l'apprentissage non supervisé.

Lorsque les données d'apprentissage ne sont ni catégorisées ni étiquetées, les techniques d'apprentissage automatique non supervisé sont utilisées. L'apprentissage non supervisé étudie comment les systèmes peuvent extrapoler une fonction à partir de données non étiquetées pour décrire une structure cachée. Le système n'a jamais la garantie que le résultat est correct. Au lieu de cela, il déduit ce que le résultat devrait être sur la base d'ensembles de données. La machine recherche dans le vaste volume de données des informations utiles.

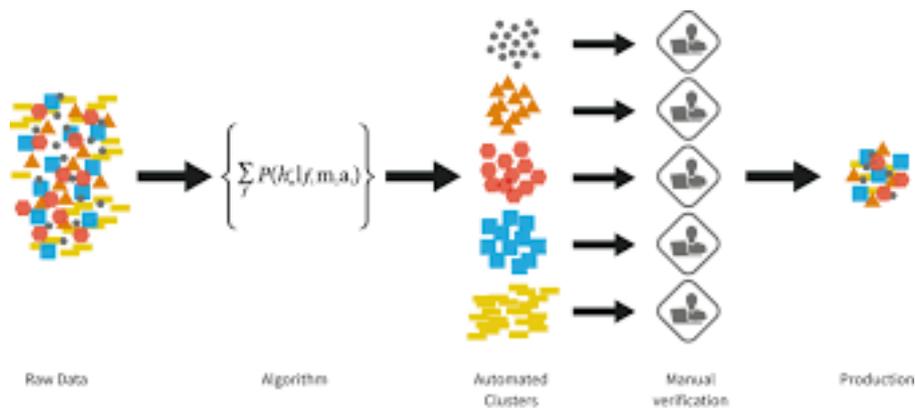


FIGURE 2.6 – Apprentissage non supervisé [5]

Clustering : Le regroupement est une technique d'apprentissage non supervisé qui permet de rassembler les éléments dans des groupes ou clusters de telle sorte que les éléments au sein d'un même groupe partagent un grand nombre de similitudes et présentent peu ou pas de similitudes avec ceux des autres groupes. Cette classification est réalisée en identifiant des caractéristiques communes parmi les données grâce à l'analyse de clustering.

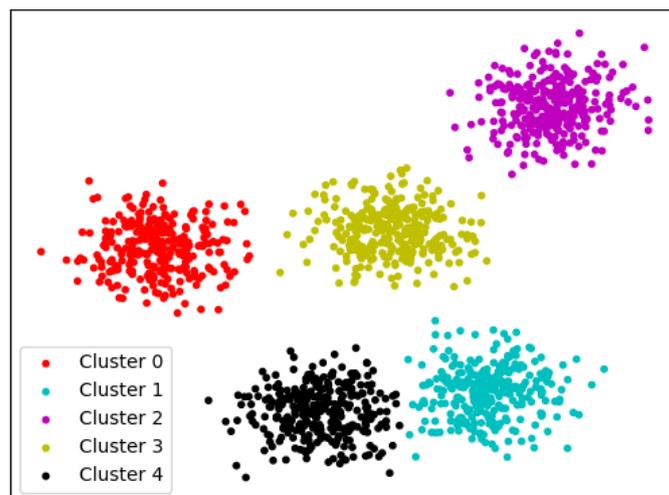


FIGURE 2.7 – clustering [6]

Réduction de dimensionnalité : Cette stratégie vise à exprimer des exemples à travers un plus petit nombre d'attributs. Cependant, chacun de ces attributs peut être une fonction non

linéaire de toutes les variables de l'exemple d'origine, par exemple le produit des variables. De plus, contrairement aux techniques de sélection de variables, les algorithmes de réduction de dimensionnalité ne sont généralement pas supervisés ; ils se basent simplement sur la distribution spatiale des exemples pour déterminer quels attributs sont pertinents.

Règles d'association : La règle d'association est une technique d'apprentissage non supervisé utilisée pour découvrir des relations entre les variables dans de grandes bases de données. Elle permet d'identifier des groupes d'éléments qui apparaissent fréquemment ensemble dans les transactions. Cette méthode est très utile en marketing pour améliorer l'efficacité des stratégies promotionnelles, en illustrant par exemple que les personnes qui achètent du pain achètent aussi souvent du beurre ou de la confiture. Ce phénomène est couramment exploré à travers l'analyse de panier de la ménagère.

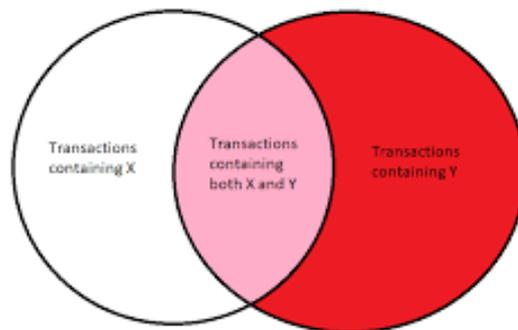


FIGURE 2.8 – Règles d'associations [7]

Le tableau 2.3 représente une comparaison entre les différents types d'apprentissages

TABLE 2.3 – Comparaison entre l'apprentissage supervisé et non supervisé

Caractéristique	Apprentissage Supervisé	Apprentissage Non Supervisé
Objectif	Prédire une sortie en fonction des entrées.	Découvrir des patterns ou des regroupements dans les données.
Type de données	Nécessite des données d'entrée avec des étiquettes.	Utilise des données d'entrée sans étiquettes.
Complexité des données	Moins complexe car les étiquettes guident l'apprentissage.	Plus complexe, nécessite souvent plus de données pour identifier les structures.
Méthodes courantes	Régression, classification.	Clustering, réduction de dimensionnalité.
Exemples d'algorithmes	Régression linéaire, réseaux de neurones, SVM.	K-means, analyse en composantes principales (PCA), autoencodeurs.
Applications	Prévision météo, reconnaissance de la parole, systèmes de recommandation.	Segmentation de marché, organisation de grandes bibliothèques de documents, compression de données.
Avantages	Précis lorsque les étiquettes sont correctes. Facile à évaluer grâce aux données de test.	Peut travailler avec des données non structurées. Utile pour découvrir l'inconnu dans les ensembles de données.
Inconvénients	Nécessite une grande quantité de données étiquetées, ce qui peut être coûteux.	Moins précis, difficile à évaluer les performances du modèle.

2.4.1.4 Apprentissage par Renforcement :

Le fonctionnement des algorithmes d'apprentissage par renforcement (RL) peut être comparé à l'apprentissage humain, notamment dans le contexte de la découverte culinaire. Imaginons un individu visitant un nouveau restaurant. Il goûte différents plats, rejetant ceux qu'il n'apprécie pas et notant ceux qu'il aime pour de futures visites.

De manière similaire, en RL, un agent évolue dans un environnement simulé. Il dispose d'un ensemble d'actions possibles, chacune influençant l'environnement. Les actions bénéfiques sont récompensées, tandis que les actions préjudiciables sont pénalisées. Au fil du temps, l'agent apprend à favoriser les actions générant des récompenses.

La stratégie optimale développée par l'agent, constituée d'une série d'états et d'actions, est appelée politique.

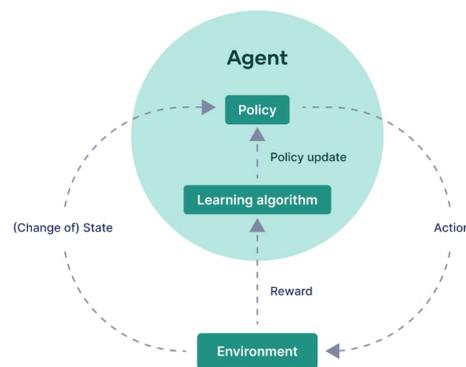


FIGURE 2.9 – Apprentissage par renforcement [8]

2.4.2 Processus d'Apprentissage et d'Évaluation des Modèles

Processus d'Apprentissage en Machine Learning :

La compréhension approfondie des besoins métiers constitue l'étape initiale et cruciale de tout projet de machine learning. Cette phase, bien que fondamentale, mérite d'être soulignée : une définition claire du problème à résoudre est indispensable avant d'entamer toute démarche de résolution. Pour ce faire, il est essentiel d'établir une collaboration étroite avec le responsable du projet (ou "project owner"). Cette coopération vise à saisir pleinement ses objectifs et ses attentes. L'enjeu est double : d'une part, transposer ces informations en une formulation mathématique adaptée au machine learning, et d'autre part, élaborer une stratégie préliminaire pour atteindre les buts fixés. Cette approche permet non seulement de garantir une compréhension mutuelle entre les experts techniques et les parties prenantes du projet, mais aussi de poser les bases solides d'une solution pertinente et efficace.

Collecte des données :

Préparation des Données pour l'Apprentissage Automatique :

La préparation des données pour l'apprentissage automatique commence par la collecte de données. Au cours de l'étape de collecte de données, vous collectez des données pour former et ajuster le futur modèle ML. Ce faisant, gardez à l'esprit le type, le volume et la qualité des données : ces facteurs détermineront la meilleure stratégie de préparation des données.

L'apprentissage automatique utilise trois types de données : structurées, non structurées et semi-structurées.

- Les données structurées sont organisées d'une manière spécifique, généralement sous forme de tableau ou de feuille de calcul. Les exemples de données structurées couvrent les informations collectées à partir de bases de données ou de systèmes transactionnels.
- Les données non structurées comprennent les images, les vidéos, les enregistrements audio et d'autres informations qui ne suivent pas les modèles de données conventionnels.
- Les données semi-structurées ne suivent pas le format d'un modèle de données tabulaire. Pourtant, il n'est pas complètement désorganisé, car il contient des éléments structurels, comme des balises ou des métadonnées qui facilitent son interprétation. Les exemples incluent des données aux formats XML ou JSON.

La structure des données détermine l'approche optimale de préparation des données pour l'apprentissage automatique. Les données structurées, par exemple, peuvent être

facilement organisées en tableaux et nettoyées via la déduplication, le remplissage des valeurs manquantes ou la normalisation des formats de données.

En revanche, l'extraction de caractéristiques pertinentes à partir de données non structurées nécessite des techniques plus complexes, telles que le traitement du langage naturel ou la vision par ordinateur.

L'approche optimale de la préparation des données pour l'apprentissage automatique est également affectée par le volume de données de formation. Un grand ensemble de données peut nécessiter un échantillonnage, ce qui implique la sélection d'un sous-ensemble de données pour former le modèle en raison de limitations de calcul. Un plus petit, à son tour, peut obliger les scientifiques des données à prendre des mesures supplémentaires pour générer plus de données basées sur les points de données existants.

La qualité des données collectées est également cruciale. L'utilisation de données inexactes ou biaisées peut affecter la production de ML, ce qui peut avoir des conséquences importantes, en particulier dans des domaines tels que la finance, la santé et la justice pénale. Il existe des techniques qui permettent de corriger les données pour les erreurs et les biais. Cependant, ils peuvent ne pas fonctionner sur un ensemble de données qui est intrinsèquement biaisé. Une fois que vous savez ce qui fait de « bonnes » données, vous devez décider comment les collecter et où les trouver. Il existe plusieurs stratégies pour cela :

- **Collecte de données à partir de sources internes** : si vous disposez d'informations stockées dans votre entrepôt de données d'entreprise, vous pouvez les utiliser pour entraîner des algorithmes de ML. Ces données peuvent inclure des transactions de vente, des interactions avec les clients, des données provenant de plateformes de médias sociaux et d'autres sources.
- **Collecte de données à partir de sources externes** : vous pouvez vous tourner vers des sources de données accessibles au public, telles que des portails de données gouvernementaux, des référentiels de données universitaires et des communautés de partage de données, telles que Kaggle, UCI Machine Learning Repository ou Google Dataset Search.
- **Web scraping** : Cette technique consiste à extraire des données de sites Web à l'aide d'outils automatisés. Cette approche peut être utile pour collecter des données à partir de sources qui ne sont pas accessibles par d'autres moyens, tels que des critiques de produits, des articles de presse et des médias sociaux.
- **Enquêtes** : cette approche peut être utilisée pour collecter des points de données spécifiques auprès d'un public cible spécifique. Il est particulièrement utile pour collecter des informations sur les préférences ou le comportement des utilisateurs.

Parfois cependant, ces stratégies ne produisent pas suffisamment de données. Vous pouvez compenser le manque de points de données avec ces techniques :

- **Augmentation des données** : qui permet de générer plus de données à partir d'échantillons existants en les transformant de diverses manières, par exemple, en les faisant pivoter, en les traduisant ou en les mettant à l'échelle.
- **Apprentissage actif** : qui permet de sélectionner l'échantillon de données le plus informatif pour l'étiquetage par un expert humain.
- **L'apprentissage par transfert** : qui consiste à utiliser des algorithmes ML pré-formés appliqués pour résoudre une tâche connexe comme point de départ pour former un nouveau modèle ML, suivi d'un réglage fin du nouveau modèle sur de nouvelles données.

- **Partage de données collaboratif** : qui implique de travailler avec d'autres chercheurs et organisations pour collecter et partager des données dans un but commun.

Nettoyage et préparation des données : Les valeurs manquantes sont un problème courant dans l'apprentissage automatique. Il peut être géré par imputation (pensez : remplir les valeurs manquantes avec des données prédites ou estimées), par interpolation (dérivation des valeurs manquantes à partir des points de données environnants) ou par suppression (supprimer simplement des lignes ou des colonnes avec des valeurs manquantes d'un ensemble de données).

Gestion des valeurs aberrantes :

Les valeurs aberrantes sont des points de données qui diffèrent considérablement du reste de l'ensemble de données. Les valeurs aberrantes peuvent se produire en raison d'erreurs de mesure, d'erreurs de saisie de données ou simplement parce qu'elles représentent des observations inhabituelles ou extrêmes. Dans un ensemble de données sur les salaires des employés, par exemple, une valeur aberrante peut être un employé qui gagne beaucoup plus ou moins que les autres. Les valeurs aberrantes peuvent être gérées en les supprimant, en les transformant pour réduire leur impact, en les winsorisant (pensez : remplacer les valeurs extrêmes par les valeurs les plus proches qui se situent dans la plage de distribution normale) ou en les traitant comme une classe de données distincte.

Suppression des doublons :

Une autre étape du processus de préparation des données pour l'apprentissage automatique consiste à supprimer les doublons. Non seulement les doublons faussent les prédictions de ML, mais ils gaspillent également de l'espace de stockage et augmentent le temps de traitement, en particulier dans les grands ensembles de données. Pour supprimer les doublons, les scientifiques des données ont recours à diverses techniques d'identification des doublons (comme la correspondance exacte, la correspondance approximative, le hachage ou le couplage d'enregistrements). Une fois identifiés, ils peuvent être supprimés ou fusionnés. Cependant, dans les ensembles de données déséquilibrés, les doublons peuvent en fait être les bienvenus pour obtenir une distribution normale.

Traitement des données non pertinentes :

Les données non pertinentes font référence aux données qui ne sont pas utiles ou applicables à la résolution du problème. La gestion de données non pertinentes peut aider à réduire le bruit et à améliorer la précision des prédictions. Pour identifier les données non pertinentes, les équipes de données ont recours à des techniques telles que l'analyse en composantes principales, l'analyse de corrélation ou s'appuient simplement sur leur connaissance du domaine. Une fois identifiés, ces points de données sont supprimés de l'ensemble de données.

Traitement des données incorrectes :

La préparation des données pour l'apprentissage automatique doit également inclure la gestion de données incorrectes et erronées. Les techniques courantes de traitement de ces données incluent la transformation des données (modification des données afin qu'elles répondent aux critères définis) ou la suppression totale des points de données incorrects.

Gestion des données déséquilibrées :

Un jeu de données déséquilibré est un jeu de données dans lequel le nombre de points de données dans une classe est nettement inférieur au nombre de points de données dans une autre classe. Cela peut aboutir à un modèle biaisé qui donne la priorité à la classe majoritaire,

tout en ignorant la classe minoritaire. Pour résoudre le problème, les équipes de données peuvent recourir à des techniques telles que le rééchantillonnage (soit suréchantillonner la classe minoritaire, soit sous-échantillonner la classe majoritaire pour équilibrer la distribution des données), la génération de données synthétiques (générant synthétiquement des points de données supplémentaires pour la classe minoritaire), le coût-apprentissage sensible (attribuer un poids plus élevé à la classe minoritaire pendant la formation), apprentissage d'ensemble (combinant plusieurs modèles formés sur différents sous-ensembles de données à l'aide de différents algorithmes), et autres.

Ces activités permettent de s'assurer que les données de formation sont exactes, complètes et cohérentes. Bien qu'il s'agisse d'une grande réussite, cela ne suffit pas encore pour produire un modèle ML fiable. Ainsi, la prochaine étape du parcours de préparation des données pour l'apprentissage automatique consiste à s'assurer que les points de données de l'ensemble de données de formation sont conformes à des règles et à des normes spécifiques. Et cette étape du processus de gestion des données est appelée transformation des données.

Transformation des données :

Au cours de l'étape de transformation des données, vous convertissez les données brutes dans un format adapté aux algorithmes d'apprentissage automatique. Cela, à son tour, garantit des performances et une précision algorithmiques supérieures.

Nos experts en préparation de données pour l'apprentissage automatique citent les techniques courantes de transformation de données suivantes :

Mise à l'échelle :

Dans un jeu de données, différentes entités peuvent utiliser différentes unités de mesure. Par exemple, un ensemble de données immobilières peut inclure des informations sur le nombre de pièces dans chaque propriété (allant de une à dix) et le prix (allant de 50 000 \$ à 1 000 000 \$). Sans mise à l'échelle, il est difficile d'équilibrer l'importance des deux fonctionnalités. L'algorithme peut accorder trop d'importance à la fonctionnalité avec des valeurs plus grandes - dans ce cas, le prix - et pas assez à la fonctionnalité avec des valeurs apparemment plus petites. La mise à l'échelle aide à résoudre ce problème en transformant tous les points de données de manière à ce qu'ils correspondent à une plage spécifiée, généralement entre 0 et 1. Vous pouvez désormais comparer différentes variables sur un pied d'égalité.

Normalisation :

Une autre technique utilisée dans la préparation des données pour l'apprentissage automatique est la normalisation. C'est similaire à la mise à l'échelle. Cependant, alors que la mise à l'échelle modifie la plage d'un ensemble de données, la normalisation modifie sa distribution.

Codage :

Les données catégorielles ont un nombre limité de valeurs, par exemple, des couleurs, des modèles de voiture ou des espèces animales. Étant donné que les algorithmes d'apprentissage automatique fonctionnent généralement avec des données numériques, les données catégorielles doivent être codées pour pouvoir être utilisées comme entrée. Ainsi, l'encodage signifie convertir des données catégorielles en un format numérique. Vous avez le choix entre plusieurs techniques d'encodage, notamment l'encodage à chaud, l'encodage ordinal et l'encodage d'étiquettes.

Discretisation :

La discrétisation est une approche de préparation des données pour l'apprentissage automatique qui permet de transformer des variables continues, telles que le temps, la température ou le poids, en variables discrètes. Prenons un ensemble de données contenant des informations sur la taille des personnes. La taille de chaque personne peut être mesurée comme une variable continue en pieds ou en centimètres. Cependant, pour certains algorithmes ML, il peut être nécessaire de discrétiser ces données en catégories, par exemple, "court", "moyen" et "grand". C'est exactement ce que fait la discrétisation. Cela permet de simplifier l'ensemble de données de formation et de réduire la complexité du problème. Les approches courantes de la discrétisation couvrent la discrétisation basée sur le clustering et la discrétisation basée sur l'arbre de décision.

Réduction de la dimensionnalité :

La réduction de dimensionnalité consiste à limiter le nombre de caractéristiques ou de variables dans un ensemble de données et à ne conserver que les informations pertinentes pour résoudre le problème. Prenons un ensemble de données contenant des informations sur l'historique des achats des clients. Il comporte la date d'achat, l'article acheté, le prix de l'article et le lieu où l'achat a eu lieu. En réduisant la dimensionnalité de cet ensemble de données, nous omettons toutes les caractéristiques sauf les plus importantes, par exemple l'article acheté et son prix. La réduction de la dimensionnalité peut être effectuée avec une variété de techniques, certaines d'entre elles étant l'analyse en composantes principales, l'analyse discriminante linéaire et l'intégration de voisins stochastiques à distribution t.

Transformation logarithmique :

Une autre façon de préparer les données pour l'apprentissage automatique, la transformation logarithmique, consiste à appliquer une fonction logarithmique aux valeurs d'une variable dans un ensemble de données. Il est souvent utilisé lorsque les données d'entraînement sont fortement asymétriques ou présentent une large plage de valeurs. L'application d'une fonction logarithmique peut aider à rendre la distribution des données plus symétrique.

En parlant de transformation des données, nous devrions également mentionner l'ingénierie des fonctionnalités. Bien qu'il s'agisse d'une forme de transformation de données, c'est plus qu'une technique ou une étape du processus de préparation des données pour l'apprentissage automatique. Il représente la sélection, la transformation et la création d'entités dans un jeu de données. L'ingénierie des fonctionnalités implique une combinaison de techniques statistiques, mathématiques et informatiques, y compris l'utilisation de modèles ML, pour créer des fonctionnalités qui capturent les informations les plus pertinentes dans les données.

Il s'agit généralement d'un processus itératif qui nécessite de tester et d'évaluer différentes techniques et combinaisons de fonctionnalités afin de trouver la meilleure approche pour résoudre un problème.

Séparation des données : La prochaine étape du processus de préparation des données pour l'apprentissage automatique consiste à diviser toutes les données recueillies en sous-ensembles - le processus connu sous le nom de fractionnement des données. En règle générale, les données sont décomposées en un jeu de données de formation, de validation et de test.

Ensemble de données de formation : utilisé pour apprendre à un modèle d'apprentissage automatique à reconnaître les modèles et les relations entre les variables d'entrée et cibles. Cet ensemble de données est généralement le plus volumineux.

Jeu de données de validation : un sous-ensemble de données utilisé pour évaluer les

performances du modèle pendant la formation. Il aide à affiner le modèle en ajustant les hyperparamètres (pensez : les paramètres du processus de formation qui sont définis manuellement avant la formation, comme le taux d'apprentissage, la force de régularisation ou le nombre de couches cachées). L'ensemble de données de validation permet également d'éviter le surajustement des données de formation.

Jeu de données de test : un sous-ensemble de données utilisé pour évaluer les performances du modèle formé. Son objectif est d'évaluer la précision du modèle sur de nouvelles données inédites. L'ensemble de données de test n'est utilisé qu'une seule fois - après que le modèle a été formé et affiné sur les ensembles de données de formation et de validation.

En divisant les données, nous pouvons évaluer les performances d'un modèle d'apprentissage automatique sur des données qu'il n'a jamais vues auparavant. Sans fractionnement, il y a de fortes chances que le modèle fonctionne mal sur de nouvelles données. Cela peut se produire parce que le modèle vient peut-être de mémoriser les points de données au lieu d'apprendre des modèles et de les généraliser à de nouvelles données.

Il existe plusieurs approches de fractionnement des données, et le choix de la meilleure dépend du problème à résoudre et des propriétés de l'ensemble de données. Nos experts en préparation des données pour l'apprentissage automatique disent que cela nécessite souvent une certaine expérimentation de la part de l'équipe de données pour déterminer la stratégie de fractionnement la plus efficace. Voici les plus courants :

Échantillonnage aléatoire : où, comme son nom l'indique, les données sont réparties de manière aléatoire. Cette approche est souvent appliquée à de grands ensembles de données représentatifs de la population modélisée. Alternativement, il est utilisé lorsqu'il n'y a pas de relations connues dans les données qui nécessitent une approche plus spécialisée.

Échantillonnage stratifié : où les données sont divisées en sous-ensembles en fonction des étiquettes de classe ou d'autres caractéristiques, suivi d'un échantillonnage aléatoire de ces sous-ensembles. Cette stratégie est appliquée aux ensembles de données déséquilibrés avec le nombre de valeurs dans une classe dépassant de manière significative le nombre de valeurs dans les autres. Dans ce cas, l'échantillonnage stratifié permet de s'assurer que les ensembles de données d'apprentissage et de test ont une distribution similaire des valeurs de chaque classe.

Échantillonnage basé sur le temps : où les données collectées jusqu'à un certain moment constituent un ensemble de données d'apprentissage, tandis que les données collectées après le point de consigne sont formées dans un ensemble de données de test. Cette approche est utilisée lorsque les données ont été collectées sur une longue période, par exemple dans des ensembles de données financières ou médicales, car elle permet de s'assurer que le modèle peut faire des prédictions précises sur les données futures.

Validation croisée : où les données sont divisées en plusieurs sous-ensembles, ou plis. Certains plis sont utilisés pour former le modèle, tandis que les autres sont utilisés pour l'évaluation des performances. Le processus est répété plusieurs fois, chaque pli servant de données de test au moins une fois. Il existe plusieurs techniques de validation croisée, par exemple, la validation croisée k-fold et la validation croisée leave-one-out. La validation croisée fournit généralement une estimation plus précise des performances du modèle que l'évaluation sur un seul jeu de données de test.

Entraînement du modèle L'entraînement d'un modèle d'apprentissage automatique consiste à ajuster les paramètres du modèle afin qu'il puisse apprendre à faire des prédictions précises à partir des données d'entrée.

Évaluation des Modèles L'évaluation des performances est une étape cruciale dans tout projet de machine learning. Cette phase essentielle vise à prédire la capacité du modèle à généraliser ses apprentissages sur de nouvelles données, non utilisées lors de son entraînement. En d'autres termes, il s'agit d'estimer la précision avec laquelle le modèle traitera des données inédites, également appelées données "hors échantillon". Pour réaliser cette évaluation de manière rigoureuse et objective, nous disposons d'un ensemble de métriques spécifiques. Ces outils de mesure permettent de quantifier différents aspects de la performance du modèle. Dans la partie suivante de notre analyse, nous allons explorer en détail ces différentes métriques, leur signification, ainsi que les méthodes employées pour les calculer. Cette approche méthodique nous permettra non seulement de juger de l'efficacité de notre modèle, mais aussi d'identifier les axes potentiels d'amélioration, assurant ainsi la robustesse et la fiabilité de notre solution d'apprentissage automatique.

Matrice de confusion Dans le cadre d'une classification supervisée, où la variable cible peut prendre plusieurs modalités (y_1, y_2, \dots, y_q), l'évaluation de la performance d'un classifieur repose sur un outil fondamental : la matrice de confusion.

Cette matrice est construite en comparant les prédictions du modèle (\hat{y}) avec les véritables classes (y) d'un ensemble d'échantillons de test. Plus précisément, la matrice de confusion, notée M , représente la distribution conjointe des variables y (classes réelles) et \hat{y} (classes prédites) [41].

En d'autres termes, cette matrice offre une vue d'ensemble de la performance du classifieur en montrant :

1. Le nombre de prédictions correctes pour chaque classe (sur la diagonale de la matrice).
2. Les types d'erreurs commises par le modèle (hors de la diagonale).

Chaque cellule $M_{i,j}$ de la matrice indique le nombre d'échantillons de la classe réelle i que le modèle a classés dans la catégorie j .

Cette représentation permet non seulement de quantifier la précision globale du modèle, mais aussi d'identifier les classes pour lesquelles le classifieur performe bien ou mal, offrant ainsi des pistes d'amélioration ciblées.

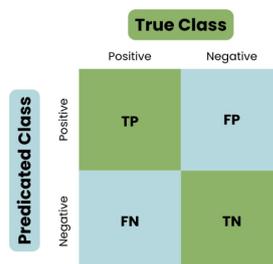


FIGURE 2.10 – Matrice de confusion [9]

La matrice de confusion est un outil essentiel pour évaluer les performances d'un classifieur. Dans le cas d'une classification binaire, où la variable cible y prend les valeurs 0 (classe "négative") ou 1 (classe "positive"), la matrice de confusion se compose de quatre éléments clés :

- **Vrais Positifs (VP ou TP)** : Nombre de cas positifs correctement identifiés.
- **Vrais Négatifs (VN ou TN)** : Nombre de cas négatifs correctement identifiés.
- **Faux Positifs (FP ou FP)** : Nombre de cas négatifs incorrectement classés comme positifs.
- **Faux Négatifs (FN ou FN)** : Nombre de cas positifs incorrectement classés comme négatifs.

Cette décomposition permet une analyse détaillée des performances du modèle, mettant en lumière non seulement sa précision globale, mais aussi ses tendances spécifiques en termes de types d'erreurs commises. Les éléments hors diagonale de la matrice représentent les erreurs de classification, tandis que les éléments sur la diagonale indiquent les prédictions correctes.

Mesures d'évaluation

1. Accuracy (Exactitude)

Ratio entre prédictions correctes et total des observations. Bien qu'intuitive, elle peut être trompeuse pour des ensembles de données asymétriques.

44

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

2. Precision (Précision)

Ratio des vrais positifs par rapport à tous les résultats positifs prédits. Elle indique la fiabilité des prédictions positives.

$$Precision = \frac{TP}{TP + FP} \quad (2.2)$$

3. Recall (Rappel ou Sensibilité)

Ratio des vrais positifs par rapport à tous les cas réellement positifs. Il mesure la capacité du modèle à identifier tous les cas positifs.

$$Recall = \frac{TP}{TP + FN} \quad (2.3)$$

4. F1-Score

Moyenne harmonique de la précision et du rappel. Particulièrement utile pour les distributions de classes inégales, il offre un équilibre entre précision et rappel.

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (2.4)$$

Ces mesures permettent une évaluation complète des performances du classifieur, chacune apportant un éclairage différent sur ses capacités.

Démarche d'évaluation Le but de l'évaluation est d'estimer au mieux les performances d'un classifieur sur de nouvelles données, c'est-à-dire lorsque le classifieur sera utilisé en pratique. L'évaluation doit donc reposer sur une démarche permettant de capturer la capacité du classifieur à se généraliser à de nouvelles données. C'est pourquoi, plutôt que d'utiliser un seul et même jeu de données pour la phase d'apprentissage et la phase d'évaluation, on préfère, via des techniques d'échantillonnage, construire et évaluer un classifieur sur des données indépendantes tirées d'une même population. On parle alors d'évaluation par validation croisée, dont les deux principales variantes sont :

Validation croisée simple

Cette démarche convient lorsque l'on travaille avec peu de données. Elle consiste simplement à découper les données en deux échantillons, D_a pour l'apprentissage du modèle et D_t pour l'évaluation, respectant la distribution de la variable de classe. Typiquement, on fait en sorte que D_a comporte 60% des données et D_t les 40% restant. Ainsi, on construit d'abord le classifieur à partir des observations dans D_a , puis on construit la matrice de confusion M en prédisant la classe de chaque observation dans D_t .

K-Validation croisée

La K-validation croisée est la démarche la plus recommandée, à condition d'avoir suffisamment de données pour pouvoir l'appliquer. Dans un premier temps, ayant choisi une valeur pour k (typiquement, on prend $k = 10$), on partitionne les données en k morceaux de même taille (D_1, D_2, \dots, D_k). Ensuite, on construit k classifieurs, chacun à partir des observations issues d'un morceau D_i , où $i \in [1; k]$. Enfin, on construit pour chaque classifieur une matrice de confusion M_i sur la base des observations issues des $k - 1$ autres morceaux (D_j , où $j \in [1; k]$).

Diagnostic biais / variance

Lorsque l'erreur mesurée sur les données d'évaluation est trop élevée, il est important d'essayer de déterminer l'origine du problème, afin de savoir comment la faire diminuer. Un diagnostic possible se base sur la décomposition de l'erreur en biais et variance. Celui-ci repose sur la comparaison de l'erreur mesurée sur les données d'apprentissage et les données d'évaluation.

1. Biais

Si l'erreur est aussi élevée sur les données d'apprentissage que sur les données d'évaluation, alors on peut supposer un problème de biais trop élevé. Cela signifie que le modèle est victime de sous-apprentissage. Pour pallier un problème de biais élevé, on peut tenter d'incorporer des descripteurs supplémentaires ou bien changer de type de modèle.

2. Variance

Si l'erreur est significativement moins élevée sur les données d'apprentissage que sur les données d'évaluation, alors on peut supposer un problème de variance trop élevée. Cela signifie que le modèle est victime de sur-apprentissage. Pour pallier un problème de variance élevée, on peut tenter de collecter plus d'individus ou bien tenter de réduire le nombre de variables.

2.5 Apprentissage Profond Appliqué à la Machine Vision

L'apprentissage profond (Deep Learning) est une branche avancée de l'intelligence artificielle qui utilise des architectures de réseaux de neurones pour modéliser des données complexes. Cette méthode d'apprentissage s'inspire directement du cerveau humain et repose sur l'utilisation des réseaux de neurones artificiels (Artificial Neural Networks : ANN). Le

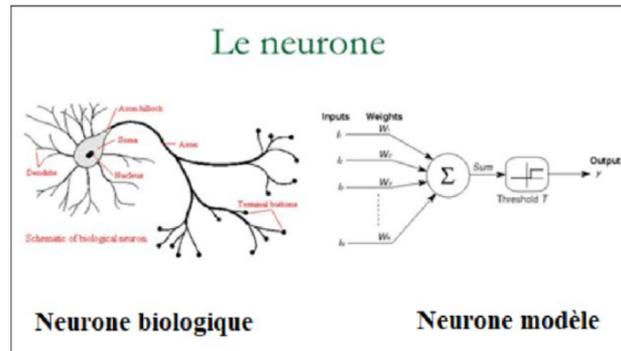


FIGURE 2.11 – Schéma représentant l’analogie entre le neurone biologique et le neurone artificielles [10]

Deep Learning est un ensemble de techniques d'apprentissage automatique qui tentent de modéliser des phénomènes de haut niveau en utilisant des architectures composées de multiples transformations non linéaires [42].

1. **Perceptron** Le perceptron représente la forme la plus élémentaire des réseaux neuronaux et constitue l'un des premiers algorithmes d'apprentissage automatique développés pour la classification binaire. Il consiste en un seul neurone avec plusieurs entrées, une unité de traitement, et d'une sortie.

Le fonctionnement du perceptron repose sur la sommation pondérée de ses entrées pour produire une sortie unique, qui peut être utilisée pour classer les entrées en deux catégories linéairement séparables. Chaque entrée du vecteur d'entrée « A » est pondérée par un poids synaptique correspondant du vecteur de poids « W ». Cette somme est ensuite passée à travers une fonction d'activation, qui décide de l'activation de la sortie en fonction d'un seuil prédéfini, imitant le fonctionnement des neurones biologiques.

Le vecteur d'entrée :

$$A = \{a_1, a_2, \dots, a_m\} \in \mathbb{R}^m$$

Le vecteur de poids synaptique :

$$W = \{w_1, w_2, \dots, w_m\} \in \mathbb{R}^m$$

La somme pondérée, plus un biais est calculée comme suit :

$$Z_j = \sum_{i=1}^p w_{ij}h(x_i) + b_j, \quad j = 1, 2, \dots, h$$

2. **Perceptron Multicouche** Le Perceptron Multicouche (PMC), ou réseau de neurones à plusieurs couches, est une extension du modèle de perceptron simple, permettant de modéliser des frontières de décision non linéaires grâce à plusieurs couches de neurones. Le PMC est composé typiquement de trois types de couches.

- **Couches d'entrée** : Déterminent la dimensionnalité des données traitées.
- **Couches de sortie** : Le nombre de neurones indique le nombre de classes pour la tâche de classification.
- **Couches cachées** : Leur configuration et le nombre de neurones sont des paramètres essentiels qui influencent la capacité du réseau à apprendre des représentations complexes.

Le PMC utilise l'algorithme de rétro-propagation pour son apprentissage, ajustant dynamiquement les poids synaptiques afin de minimiser l'écart entre les sorties prédites et les valeurs réelles. Ce processus repose sur la méthode de descente de gradient, où le gradient de l'erreur est propagé à travers le réseau pour permettre la mise à jour des poids de manière efficace. Pendant la rétro-propagation, les informations circulent en sens inverse dans le réseau, calculant le gradient nécessaire pour ajuster les poids synaptiques de tous les neurones à travers les différentes couches. L'objectif principal est de minimiser l'erreur quadratique moyenne E entre le vecteur de sortie estimé y et le vecteur de sortie réelle y_r , améliorant ainsi continuellement la précision et la performance du modèle.

Les poids synaptiques sont ensuite modifiés tel que :

$$\Delta w^{(t+1)} = -\alpha \frac{\partial E}{\partial w} + \mu \Delta w^{(t)}$$

$$E = \frac{1}{2} \sum_{i=1}^N (y_{r_i} - y_i)^2$$

3. Réseaux de Neurones Convolutifs (CNN)

Les réseaux de neurones convolutifs (CNN) sont conçus pour traiter des données qui ont une grille topologique, comme les images. La structure et le fonctionnement des CNN sont adaptés pour capturer les motifs spatiaux hiérarchiques dans les données d'entrée grâce à une architecture composée de couches convolutives, de couches de pooling, et de couches entièrement connectées.

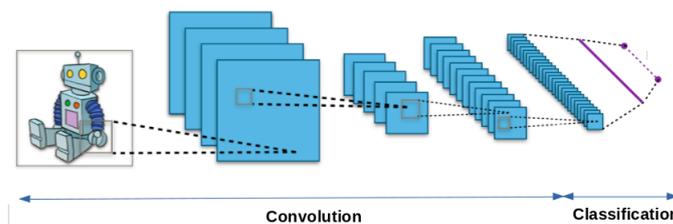


FIGURE 2.12 – Réseaux de Neurones Convolutifs (CNN) [11]

Architecture d'un CNN

a. Couche Convulsive :

La couche convulsive est la première étape d'un CNN et elle est responsable de l'extraction des caractéristiques de l'image. Elle applique des filtres (ou noyaux) sur les images d'entrée pour produire des cartes de caractéristiques. Ces filtres sont capables de détecter des motifs simples comme des bords, des textures, et des coins, et des motifs plus complexes dans les couches profondes.

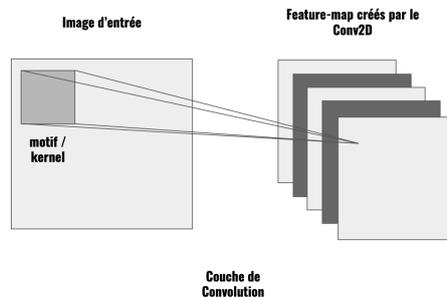


FIGURE 2.13 – Couche Convulsive [12]

b. Fonction d'Activation (ReLU) :

Après chaque couche convulsive, une fonction d'activation non-linéaire, typiquement ReLU (Rectified Linear Unit), est appliquée pour introduire de la non-linéarité dans le modèle. La fonction ReLU transforme les valeurs négatives en zéro, ce qui aide à accélérer la convergence du réseau et à éviter le problème de l'explosion ou de la disparition des gradients.

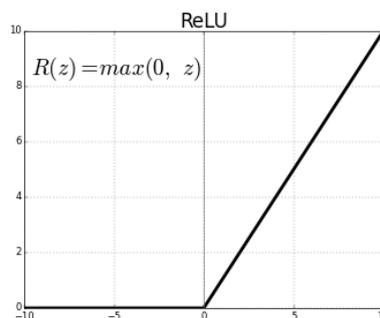


FIGURE 2.14 – Fonction d'Activation (ReLU) [13]

c. Couche de Pooling :

La couche de pooling suit généralement une couche convulsive. Son rôle est de réduire la dimensionnalité des cartes de caractéristiques tout en préservant les informations essentielles. Le pooling peut être de type max pooling, qui sélectionne la valeur maximale dans chaque région de la carte de caractéristiques, ou average pooling, qui calcule la moyenne des valeurs.

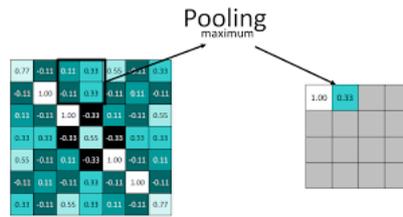


FIGURE 2.15 – Couche de Pooling [14]

d. **Couche Entièrement Connectée (FC) :**

Les couches entièrement connectées apparaissent vers la fin du CNN. Elles prennent les cartes de caractéristiques aplaties des couches précédentes et les utilisent pour prédire la classe de l'image. Chaque neurone dans une couche entièrement connectée est connecté à chaque neurone de la couche précédente, ce qui permet de combiner les caractéristiques extraites pour effectuer la classification.

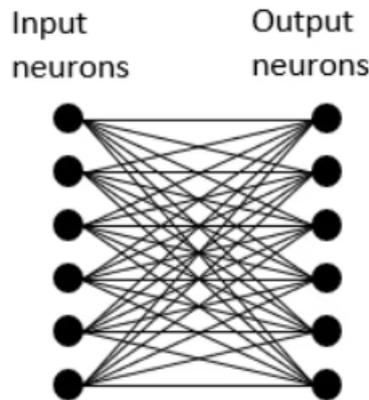


FIGURE 2.16 – Schéma d'une couche entièrement connectée avec 6 classes [15]

2.6 Apprentissage par transfert (Transfer Learning)

L'apprentissage par transfert est une technique puissante en apprentissage automatique, particulièrement utile en vision par ordinateur, où l'on transfère les connaissances acquises lors de la résolution d'un problème source pour résoudre un problème cible différent mais lié. L'idée est de réutiliser un modèle pré-entraîné sur une tâche avec généralement un large dataset (ex : classification sur ImageNet) et de l'adapter à une nouvelle tâche spécifique disposant souvent de moins de données annotées. Cela permet de tirer parti des features génériques apprises par le modèle source pour obtenir de bonnes performances sur la tâche cible malgré des données limitées, réduisant le besoin de partir de zéro.

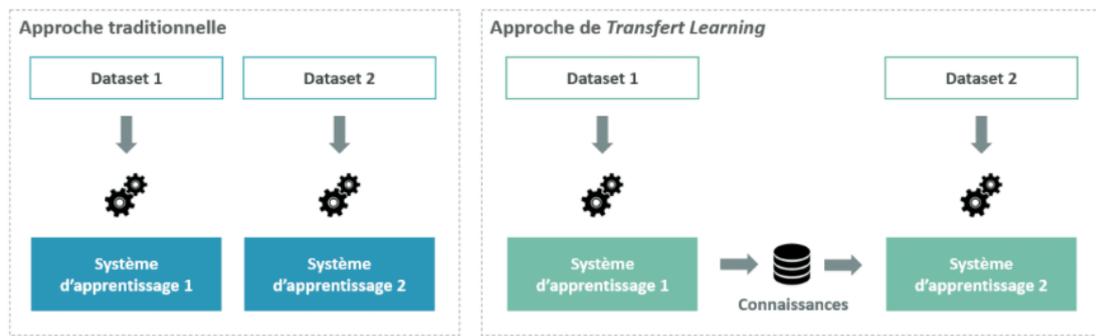


FIGURE 2.17 – Apprentissage par transfert [16]

On distingue trois principaux types d'apprentissage par transfert :

Apprentissage par transfert inductif : Dans ce scénario d'apprentissage par transfert, le domaine source et le domaine cible sont identiques, mais les tâches associées diffèrent. L'objectif est d'exploiter les connaissances acquises lors de la résolution de la tâche source pour faciliter l'apprentissage de la nouvelle tâche cible. Les algorithmes développés dans ce contexte cherchent à tirer parti des biais inductifs inhérents au domaine commun, permettant ainsi un transfert efficace des connaissances entre les tâches. Cette approche vise à améliorer les performances sur la tâche cible en capitalisant sur l'expérience acquise dans le même domaine, mais pour une tâche différente [43].

Apprentissage par transfert non supervisé : Ce type de transfert d'apprentissage s'apparente au transfert inductif, mais se concentre sur les tâches non supervisées dans le domaine cible. Il se caractérise par :

- Une similarité entre les domaines source et cible.
- Des tâches différentes entre ces domaines.
- L'absence de données étiquetées dans les deux domaines.

Cette approche vise à exploiter les connaissances acquises dans le domaine source pour améliorer les performances sur une tâche non supervisée dans le domaine cible, malgré l'absence d'étiquettes. Elle relève le défi de transférer des connaissances utiles entre des tâches distinctes et des domaines apparentés, sans s'appuyer sur des données étiquetées [43].

Apprentissage par transfert transductif : Dans ce type il existe des similitudes entre les tâches source et cible, mais les domaines correspondants sont différents. Dans ce paramètre, le domaine source a beaucoup de données étiquetées, tandis que le domaine cible n'en a pas. Cela peut être classé en sous-catégories, en se référant aux paramètres où les espaces caractéristiques sont différents ou les probabilités marginales [43].

Avantages du Transfer Learning :

Le transfert learning (TL) se distingue de l'apprentissage traditionnel par sa capacité à utiliser les connaissances acquises dans un domaine source pour améliorer l'apprentissage dans un domaine cible, souvent avec des contraintes différentes sur les données disponibles. Voici les principales relations :

Flexibilité des données étiquetées :

- L'apprentissage traditionnel nécessite généralement des données étiquetées dans le domaine d'intérêt.
- Le TL peut fonctionner avec des étiquettes disponibles ou non dans les domaines source et cible, selon le type de TL utilisé.

Diversité des tâches :

- L'apprentissage traditionnel se concentre souvent sur une seule tâche à la fois.
- Le TL, notamment dans sa forme inductive, peut aborder l'apprentissage multi-tâche et l'apprentissage autodidacte.

Adaptation aux domaines :

- L'apprentissage traditionnel fonctionne généralement dans un seul domaine.
- Le TL, en particulier le TL transductif, permet une adaptation entre différents domaines.

Types de problèmes :

- L'apprentissage traditionnel se limite souvent à la régression et à la classification supervisées.
- Le TL élargit le champ d'application à des scénarios non supervisés comme le clustering.

Efficacité d'apprentissage :

- L'apprentissage traditionnel peut nécessiter de grandes quantités de données étiquetées.
- Le TL peut être plus efficace en utilisant les connaissances transférées, réduisant potentiellement le besoin de données étiquetées dans le domaine cible.

TABLE 2.4 – Tableau comparatif des types de transfert learning

Type de TL	Domaines connexes	Domaine source étiqueté	Domaine cible étiqueté	Tâches
TL inductif	Apprentissage multi-tâche	Disponible	Non disponible	Régression, Classification
TL transductif	Adaptation de domaine	Disponible	Non disponible	Régression, Classification
TL non supervisé	-	Non disponible	Non disponible	Clustering

2.6.1 Fine-tuning pour l'adaptation des modèles :

D'après l'article [44] Une approche courante pour réaliser le *transfer learning* est le *fine-tuning* du modèle pré-entraîné sur la tâche cible. Les premières couches du réseau, qui capturent généralement des features de bas niveau (bords, textures...), sont souvent gelées, tandis que les dernières couches plus spécifiques à la tâche sont réentraînées avec les données cibles. L'étendue du *fine-tuning* (nombre de couches réentraînées et à quel point) est un hyperparamètre clé.

Il existe principalement deux types de *fine-tuning* :

Fine-Tuning des Dernières Couches

Cette approche consiste à geler les premières couches du réseau et à réentraîner seulement les couches supérieures. Cela permet de mieux s'adapter à la tâche cible tout en nécessitant moins de données et en étant plus robuste au sur-apprentissage quand les données cibles sont limitées. Pour la classification d'images, une méthode répandue est d'utiliser le réseau pré-entraîné comme extracteur de features fixe (en ôtant la dernière couche) et d'entraîner juste un nouveau classifieur par-dessus.

Fine-Tuning de Toutes les Couches

Cette méthode implique de réentraîner l'intégralité du réseau, ce qui permet de mieux adapter le modèle aux spécificités de la tâche cible. Cependant, elle nécessite une quantité de données plus importante pour éviter le sur-apprentissage. Pour des tâches plus complexes comme la détection d'objets, un schéma courant est de geler le *backbone* (typiquement un réseau de classification comme ResNet) et de *fine-tuner* les couches hautes (par exemple, RPN et classifieurs dans Faster R-CNN).

Un *fine-tuning* extensif permet de mieux s'adapter à la tâche cible mais nécessite plus de données, tandis qu'un *fine-tuning* léger est plus robuste au sur-apprentissage quand les données cibles sont limitées. En fonction de la nature et de la complexité de la tâche, ainsi que de la quantité de données disponibles, le choix entre ces deux approches de *fine-tuning* doit être soigneusement considéré.

Pour de la classification d'images, une approche répandue est d'utiliser le réseau pré-entraîné comme extracteur de features fixe (en ôtant la dernière couche) et d'entraîner juste un nouveau classifieur par-dessus. Pour des tâches plus complexes comme la détection d'objets, un schéma courant est de geler le backbone (typiquement un réseau de classification comme ResNet) et de fine-tuner les couches hautes (ex : RPN, classifieurs dans Faster R-CNN).

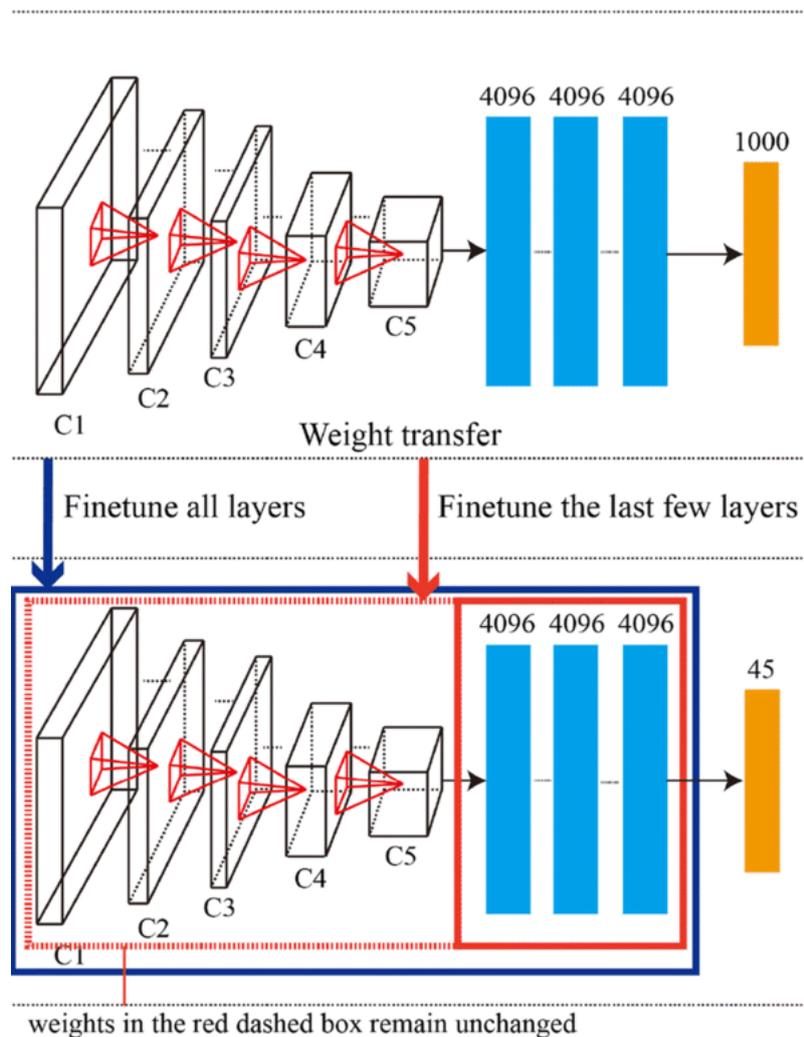


FIGURE 2.18 – Explication de fine-tuning [17]

2.6.2 Modèles pré-entraînés populaires :

De nombreux modèles pré-entraînés sur de larges datasets comme ImageNet sont disponibles dans les bibliothèques de deep learning et forment d'excellents points de départ pour le transfer learning.

VGGNet : Proposé en 2014 par Simonyan et al. du Visual Geometry Group, a marqué une évolution majeure des réseaux de neurones convolutifs. Le modèle VGG-16 se compose de 13 couches convolutionnelles et 3 couches entièrement connectées, utilisant uniquement des filtres 3×3, une innovation par rapport aux filtres plus larges d'AlexNet et ZF Net. Avec 138 millions de paramètres et 500 Mo de stockage, VGG-16 a démontré sa puissance en atteignant un taux d'erreur de 7,3% dans le top 5 à l'ILSVRC. Les chercheurs ont également conçu le VGG-19, une version plus profonde du modèle.

ResNet : ResNet (Residual Neural Network) est une architecture de réseau profond comprenant 152 couches, développée par Microsoft en 2015. Ce modèle a remporté l'ILSVRC 2015 avec un taux d'erreur de 3,6 %, surpassant même la précision humaine. Les couches résiduelles de ResNet calculent les changements par rapport à l'entrée, qui sont ensuite ajoutés à cette entrée pour produire la sortie finale. ResNet-50 est l'un des premiers modèles à intégrer la normalisation par batch. L'architecture de ResNet se compose de deux types de blocs : les blocs convolutionnels (CONV) et les blocs d'identité.

Inception/GoogLeNet : GoogLeNet, également connu sous le nom d'Inception-v1, a remporté l'ILSVRC 2014 avec un taux d'erreur de 6,7 % dans le top 5. Développé par des chercheurs de Google en 2014, ce modèle se distingue par sa pile de 22 couches de convolution et l'introduction d'un nouveau module appelé Inception Module. Le modèle utilise des calculs parallèles de convolutions de différentes tailles ainsi que du max-pooling, dont les sorties sont concaténées à la fin de chaque module Inception. Grâce à cette architecture innovante, le nombre de paramètres a été réduit de 60 millions (comme dans AlexNet) à seulement 4 millions.

MobileNetV2 : L'architecture MobileNetV2 est conçue pour bien fonctionner sur les appareils mobiles. Elle repose sur une structure résiduelle inversée, où les connexions résiduelles se trouvent entre les couches de bottleneck. La couche d'expansion intermédiaire utilise des convolutions en profondeur légères pour filtrer les entités, servant ainsi de source de non-linéarité. Dans son ensemble, l'architecture de MobileNetV2 comprend une couche initiale entièrement convolutionnelle avec 32 filtres, suivie de 19 couches de bottleneck résiduelles.

EfficientNet : EfficientNet est une architecture et une méthode de mise à l'échelle qui ajuste uniformément toutes les dimensions de profondeur, largeur et résolution en utilisant un coefficient composé. Contrairement à la pratique conventionnelle qui ajuste ces facteurs de manière arbitraire, la méthode de mise à l'échelle d'EfficientNet applique un ensemble fixe de coefficients pour une mise à l'échelle uniforme de la largeur, de la profondeur et de la résolution du réseau. Le réseau de base, EfficientNet-B0, repose sur les blocs résiduels de bottleneck inversé de MobileNetV2, avec l'ajout des blocs de squeeze-and-excitation.

YOLO (You Only Look Once) : est l'un des algorithmes de détection d'objets les plus populaires, largement utilisé par les chercheurs du monde entier. Présenté pour la première fois en 2015 dans un article de Joseph Redmon et al., YOLO utilise les caractéristiques de l'image entière pour prédire chaque boîte englobante. Il prédit également simultanément toutes les boîtes englobantes pour toutes les classes présentes dans une image. Cette approche permet au réseau de raisonner globalement sur l'ensemble de l'image et tous les objets qu'elle

contient. La conception de YOLO permet un apprentissage de bout en bout et offre des vitesses de traitement en temps réel tout en maintenant une précision moyenne élevée [45].

Faster R-CNN : Ce réseau, proposé par Shaoqing Ren et al. en 2016, se compose de deux modules principaux :

- Premier module : Un réseau convolutif profond qui crée la carte de caractéristiques convolutives. Cette carte est utilisée par un module RPN (Réseau de Proposition de Région), qui prend cette carte (quelle que soit sa taille) et produit un ensemble de propositions d'objets rectangulaires, chacune accompagnée d'un score de précision.
- Second module : Le détecteur Fast R-CNN, qui utilise les régions proposées par le module précédent, comme décrit dans l'architecture Fast R-CNN [46].

RetinaNet : est un réseau unifié composé d'un réseau dorsal (backbone) et de deux sous-réseaux spécifiques à des tâches. Le réseau dorsal est chargé de calculer une carte de caractéristiques convolutives sur l'ensemble de l'image d'entrée et consiste en un réseau convolutif standard. Le premier sous-réseau effectue une classification d'objets par convolution sur la sortie du réseau dorsal ; le second sous-réseau réalise une régression des boîtes englobantes par convolution. Les deux sous-réseaux présentent une conception simple que nous proposons spécifiquement pour la détection dense en une étape. Bien qu'il existe de nombreux choix possibles pour les détails de ces composants, la plupart des paramètres de conception ne sont pas particulièrement sensibles aux valeurs exactes, comme le montrent les expériences. **EfficientDet** : EfficientDet est une architecture d'apprentissage profond novatrice, a été présentée en 2019 par Mingxing Tan et Quoc V. Le. Cette approche vise à résoudre le défi complexe d'optimisation de la taille du modèle et de l'efficacité computationnelle, tout en maintenant, voire en améliorant, les performances globales du réseau. L'innovation clé d'EfficientNet réside dans sa méthode de mise à l'échelle, qui cherche à équilibrer trois dimensions cruciales : la profondeur, la largeur et la résolution du réseau. Contrairement aux approches traditionnelles qui ajustent ces dimensions séparément, EfficientNet propose une mise à l'échelle composée, harmonisant les trois aspects de manière systématique et équilibrée. Cette méthode s'est avérée particulièrement efficace dans divers domaines d'application [47]. Le tableau (Table 2.5) présente une classification des modèles pré-entraînés selon l'objectif principal de chaque modèle.

TABLE 2.5 – Modèles pré-entraînés et leurs utilisations

Classification	Détection d'Objets	Segmentation d'Images
VGGNet	YOLO (You Only Look Once)	Mask R-CNN
ResNet	Faster R-CNN	
Inception/GoogLeNet	RetinaNet	
MobileNetV2	EfficientDet	
EfficientNet		

2.6.3 Évaluation des modèles pré-entraînés pour la détection d'objet

Perte d'entropie croisée : La perte d'entropie croisée est une fonction de perte largement utilisée pour les problèmes de classification, où elle compare les probabilités prédites du modèle avec les étiquettes réelles. Pour la détection d'objets, la perte d'entropie croisée peut être appliquée à la fois au score d'objectivité et au score de classe de chaque cadre englobant.

Le score d'objectivité indique si le cadre de sélection contient un objet ou non, et le score de classe indique à quelle classe d'objet il appartient. La perte d'entropie croisée pénalise le modèle pour avoir fait des prédictions incorrectes et l'encourage à apprendre les probabilités correctes. Cependant, la perte d'entropie croisée ne tient pas compte du chevauchement spatial entre les cadres de délimitation prédits et réels, ce qui est important pour la détection d'objets.

Perte d'IoU : La perte d'IoU est une fonction de perte qui mesure directement le chevauchement spatial entre les cadres de délimitation prévus et réels. IoU signifie intersection sur union, et il est défini comme le rapport de la zone d'intersection à la zone d'union de deux boîtes de délimitation. La perte d'IoU est calculée comme 1 moins l'IoU, de sorte qu'une perte plus faible signifie un chevauchement plus élevé. La perte d'IoU peut être utilisée pour optimiser les coordonnées des cadres de sélection, tandis que la perte d'entropie croisée peut être utilisée pour les scores d'objectivité et de classe. La perte d'IoU a l'avantage d'être plus alignée sur la métrique d'évaluation de la détection d'objets, qui est généralement basée sur IoU. Cependant, la perte d'IoU peut être instable et sensible aux valeurs aberrantes, car elle peut avoir de grands gradients lorsque le chevauchement est faible ou nul.

Perte focale : La perte focale est une version modifiée de la perte d'entropie croisée, conçue pour résoudre le problème du déséquilibre de classe dans la détection d'objets. Le déséquilibre de classe signifie que certaines classes sont plus fréquentes ou plus faciles à détecter que d'autres, ce qui peut amener le modèle à se concentrer davantage sur elles et à ignorer les classes rares ou difficiles. La perte focale introduit un facteur d'échelle qui réduit la perte pour les exemples bien classés et augmente la perte pour les exemples mal classés. De cette façon, la perte focale peut équilibrer la contribution des différentes classes et améliorer les performances de détection. La perte focale peut être appliquée à la fois aux scores d'objectivité et de classe des cadres englobants. La perte focale a l'avantage d'être plus résistante au déséquilibre de classe et au bruit, mais elle introduit également un hyperparamètre supplémentaire qui doit être ajusté.

Perte de GIoU : La perte GIoU est une extension de la perte IoU, qui vise à surmonter certaines de ses limites. GIoU signifie intersection généralisée sur l'union, et il est défini comme l'IoU moins l'aire de la plus petite boîte englobante qui contient à la fois les boîtes de délimitation prédites et vraies. La perte GIoU est calculée comme 1 moins le GIoU, de sorte qu'une perte plus faible signifie un chevauchement plus élevé et une boîte de confinement plus petite. La perte de GIoU peut gérer les cas où l'IoU est nulle ou proche de zéro, en pénalisant le modèle de création de grandes boîtes englobantes. La perte de GIoU peut également capturer la forme et l'orientation des cadres de sélection, ce que la perte d'IoU ne peut pas. La perte de GIoU peut améliorer la précision et la stabilité de la détection d'objets, mais elle peut également être plus coûteuse en calcul que la perte IoU.

Ces métriques aident à déterminer dans quelle mesure le modèle pré-entraîné a transféré ses connaissances à la nouvelle tâche et au nouveau jeu de données. Le suivi de ces métriques pendant le fine-tuning garantit que le modèle s'adapte bien à la tâche spécifique. Si les métriques ne sont pas satisfaisantes, un fine-tuning supplémentaire, des ajustements d'hyperparamètres ou l'exploration de modèles pré-entraînés alternatifs peuvent être nécessaires.

Tous ces modèles pré-entraînés peuvent être utilisés comme point de départ en les fine-tuning sur de nouvelles tâches ou domaines avec généralement beaucoup moins de données que pour un apprentissage from scratch. Ils tirent leur force de leur capacité à capturer des représentations génériques transférables. Le transfer learning a révolutionné la vision par ordinateur en permettant d'obtenir rapidement de bonnes performances même en disposant

de peu de données. Cependant, son efficacité dépend de la similarité entre tâche/ domaine source et cible. Lorsque ceux-ci sont très différents (ex : imagerie médicale), un apprentissage from scratch peut être préférable. Enfin, l'apprentissage par transfert non supervisé est un domaine de recherche très actif qui pourrait permettre de tirer parti de larges volumes de données non annotées.

2.6.4 L'apprentissage automatique VS l'apprentissage par transfert

Le tableau (Table 2.6) représente une comparaison entre les caractéristiques d'apprentissage automatique classiques et l'apprentissage par transferts

TABLE 2.6 – Comparaison entre Machine Learning et Transfer Learning - Caractéristiques

Caractéristique	Machine Learning	Transfer Learning
Principe Fondamental	Création de modèles à partir de zéro, en apprenant des données spécifiques à la tâche.	Utilisation de connaissances préexistantes à partir de modèles pré-entraînés pour accélérer le processus d'apprentissage sur de nouvelles tâches similaires.
Utilisation Initiale des Modèles	Démarre l'apprentissage à partir de zéro sur des données spécifiques à la tâche.	Commence avec des modèles pré-entraînés sur des ensembles de données vastes et généraux, comme ImageNet ou des modèles linguistiques pour NLP.
Quantité de Données Requise	Nécessite souvent de grandes quantités de données pour obtenir des performances optimales.	Réduit la quantité de données requises en utilisant des modèles pré-entraînés comme point de départ, ce qui est efficace pour les tâches avec des données limitées.
Temps d'Entraînement	Peut nécessiter des temps de calcul significatifs pour entraîner un modèle à partir de zéro.	Accélère le processus d'entraînement en adaptant des modèles pré-entraînés, réduisant ainsi le temps nécessaire pour atteindre la compétence sur une nouvelle tâche.
Adaptabilité à de Nouvelles Tâches	Moins adaptable sans beaucoup de données spécifiques à la tâche.	Très adaptable, même avec peu de données, en utilisant des connaissances pré-existantes des modèles pré-entraînés.

2.7 Technologies des Caméras Industrielles

Les technologies des caméras industrielles jouent un rôle crucial dans l'inspection de la qualité des produits et des processus dans divers secteurs industriels. Le choix de la caméra appropriée est déterminant pour optimiser les processus de production et garantir la qualité des produits. Ce choix dépend de plusieurs facteurs, tels que le type de produit à inspecter, la vitesse de production, ainsi que les exigences de résolution et de précision.

Types de Caméras Les caméras industrielles peuvent être classées en plusieurs catégories standard en fonction de leurs caractéristiques et de leurs applications :

Caméras 2D

- **Caméras Linéaires** : Utilisent une rangée de capteurs pour reconstituer une image par balayage. Elles permettent d'atteindre des vitesses et des résolutions élevées, dépassant souvent 8000 pixels sur une seule ligne, ce qui les rend idéales pour la visualisation d'objets en mouvement ou de grande dimension et pour la supervision de processus continus tels que la production textile. Une application typique est la visualisation d'objets cylindriques, où la rotation de l'objet permet une visualisation plane de sa surface.
- **Caméras Matricielles** : Capturent des images en deux dimensions à un moment donné. Elles sont souvent utilisées dans des applications où l'objet est statique ou se déplace lentement, nécessitant une analyse détaillée de la surface des objets.

Caméras 3D

- **Caméras à Temps de Vol (ToF)** : Mesurent le temps que met un signal lumineux à revenir à la caméra après réflexion pour créer des images 3D précises. Elles sont utilisées pour des applications nécessitant des mesures de distance et des reconstructions 3D, comme la robotique, la logistique et l'inspection des objets complexes.
- **Stéréocaméras** : Utilisent deux ou plusieurs capteurs pour capturer des images sous différents angles et reconstruire la géométrie 3D. Elles sont souvent utilisées pour la navigation des robots, la surveillance et les applications nécessitant une vision en profondeur.

Caméras Spectrales

- **Caméras Hyperspectrales** : Capturent des images à travers un large spectre de longueurs d'onde, permettant d'identifier les matériaux et compositions chimiques des objets inspectés. Ces caméras sont utilisées dans des secteurs comme l'agriculture de précision, la surveillance environnementale et l'inspection des produits alimentaires.
- **Caméras Multispectrales** : Capturent des images dans plusieurs bandes spectrales discrètes pour identifier des caractéristiques spécifiques des matériaux. Elles sont utilisées dans des applications agricoles, environnementales et médicales pour identifier des caractéristiques spécifiques des matériaux.

Caméras Thermiques

- **Caméras Thermiques** : Détectent les rayonnements infrarouges émis par les objets et les convertissent en images thermiques. Elles sont utilisées pour surveiller la température des objets, détecter les points chauds, et identifier les problèmes de surchauffe dans les équipements électriques et mécaniques. Elles sont également utilisées dans les applications de sécurité et de surveillance.

Caméras à Haute Vitesse

- **Caméras à Haute Vitesse** : Capturent des images à des taux de rafraîchissement extrêmement élevés, souvent des milliers d'images par seconde. Elles sont utilisées

pour analyser des événements très rapides, tels que des explosions, des impacts, ou des processus de production à grande vitesse. Ces caméras sont essentielles pour les applications de recherche et développement, ainsi que pour les analyses détaillées de mouvements rapides.

le tableau 2.7 présente un résumé des différentes classes de caméras industrielles, leur fonctionnement et leurs applications typiques, facilitant ainsi le choix de la technologie de caméra la plus appropriée pour une application donnée.

TABLE 2.7 – Classification et Applications des Caméras Industrielles

Classe de Caméra	Fonctionnement	Applications Typiques
Caméras 2D	Utilisent une rangée de capteurs pour reconstituer une image par balayage. Capturent des images en deux dimensions à un moment donné.	Production textile, visualisation d'objets cylindriques. Inspection détaillée des surfaces d'objets statiques
Caméras 3D	Mesurent le temps que met un signal lumineux à revenir à la caméra après réflexion pour créer des images 3D précises. Utilisent deux ou plusieurs capteurs pour capturer des images sous différents angles et reconstruire la géométrie 3D.	Robotique, logistique, inspection d'objets complexes. Navigation des robots, surveillance, applications nécessitant une vision en profondeur
Caméras Spectrales	Capturent des images à travers un large spectre de longueurs d'onde pour identifier les matériaux et compositions chimiques. Capturent des images dans plusieurs bandes spectrales discrètes pour identifier des caractéristiques spécifiques des matériaux.	Agriculture de précision, surveillance environnementale, inspection des produits alimentaires. Agriculture, environnement, médical pour identifier des caractéristiques spécifiques des matériaux
Caméras Thermiques	Détectent les rayonnements infrarouges émis par les objets et les convertissent en images thermiques pour surveiller la température.	Surveillance de la température des équipements électriques et mécaniques, applications de sécurité
Caméras à Haute Vitesse	Capturent des images à des taux de rafraîchissement très élevés, permettant l'analyse d'événements rapides.	Analyse des explosions, impacts, processus de production à grande vitesse

Chapitre 3

Conception et Développement du Système de Contrôle Qualité

Ce chapitre présente le développement d'un système de contrôle qualité destiné à l'inspection des impressions de dates et des scellages des sachets de détergent. En utilisant des techniques avancées de vision industrielle et d'apprentissage automatique, l'objectif est d'assurer la conformité et la qualité des produits. Nous débutons par une description détaillée du dispositif de caméra utilisé pour l'inspection des impressions de dates, couvrant l'architecture du système et les caractéristiques techniques de la caméra industrielle. Nous expliquons ensuite le fonctionnement du système, incluant la capture d'images, la détection de la zone d'impression, et la reconnaissance optique de caractères (OCR). Le chapitre se poursuit avec le développement des modèles d'inspection, abordant la collecte et la préparation des données, le prétraitement, l'augmentation et l'annotation des images, ainsi que la division des données en ensembles d'entraînement, de validation et de test. Pour l'inspection des scellages, deux approches principales sont explorées : la détection d'objets et la segmentation d'instances avec analyse des lignes de scellage. Chaque approche est décrite en termes de modèles utilisés, de processus de développement et de résultats obtenus. Nous comparons les performances des modèles YOLOv8 et YOLOv9, en soulignant pourquoi YOLOv8 a été retenu pour sa précision et sa stabilité. Enfin, ce chapitre se conclut par une évaluation des différentes approches pour la détection des soudures, démontrant l'efficacité et la robustesse du système développé pour l'inspection de la qualité des impressions de dates et des scellages.

3.1 Partie1 :Inspection de l'impression des Dates

3.1.1 Architecture du système

Description Technique du Dispositif de Caméra : Le système d'inspection de l'impression des dates est placé à l'arrière de la ligne Volpack, exactement à l'endroit où le film est préparé avant d'être formé en sachet. À cette étape, l'opération de datation est effectuée, ce qui rend cet emplacement idéal pour l'inspection car le film est encore plat, facilitant ainsi la détection de la zone d'impression et de la date.

Pour cette partie, nous utilisons la caméra industrielle monochrome **MER2-503-36U3M/C** :, qui possède les caractéristiques suivantes :

TABLE 3.1 – Caractéristiques de la caméra MER2-503-36U3M/C

Caractéristiques	Détails
Capteur	CMOS à obturateur global Sony IMX264
Résolution	2448 × 2048 pixels
Taille de pixel	3,45 μ m × 3,45 μ m
Taux de trame maximal	36 fps
Dimensions	29mm × 29mm × 29mm
Consommation d'énergie	< 2,7W à 5V
Connexion	USB3 Vision et GenICam
Modes d'acquisition	Cadre unique, continu, déclenchement logiciel/externe
Modes d'exposition	Standard et minimal
Fonctionnalités supplémentaires	LUTs, ensembles utilisateur, minuterie, compteur

La caméra MER2-503-36U3M/C est particulièrement adaptée pour les applications de vision industrielle grâce à ses performances élevées et sa faible consommation d'énergie. Ses caractéristiques détaillées incluent :

- **Mode d'acquisition** : Acquisition de cadre unique, acquisition continue, acquisition par déclenchement logiciel, acquisition par déclenchement externe
- **Modes d'exposition** : Temps d'exposition standard et minimal
- Support des **LUTs**, ensembles utilisateur stockables, minuterie et compteur
- Réduction du **bruit** et **netteté** pour les modèles monochromes
- Compatible avec les logiciels tiers tels que **HALCON**, **MERLIC** et **LabVIEW**

La caméra est intégrée dans la ligne de conditionnement ou de remplissage sans source de lumière supplémentaire nécessaire pour la ligne Volpack spécifique, bien que dans d'autres lignes, une source de lumière **LED** pourrait être nécessaire pour éclairer la zone d'impression.

Fonctionnement du Système d'inspection Le système de caméra est intégré de manière stratégique dans la ligne de conditionnement pour capturer des images des sachets juste avant leur formation finale. La caméra est positionnée de sorte que chaque film passe sous la caméra avant d'être formé en sachet, ce qui permet une inspection optimale de l'impression des dates.



FIGURE 3.1 – Position de la caméra dans la ligne volpack

Le processus d'inspection comprend plusieurs étapes clés :

1. **Capture d'Image** : La caméra capture des images haute résolution des impressions de dates sur les films plats.
2. **Détection de la Zone d'Impression** : Le système commence par détecter la zone d'impression sur le film. Cela est essentiel pour s'assurer que l'analyse se concentre sur la zone pertinente.
3. **Reconnaissance Optique de Caractères (OCR)** : Une fois la zone d'impression détectée, un algorithme OCR est utilisé pour inspecter la lisibilité du code imprimé. Cela permet de vérifier si la date est imprimée de manière lisible et conforme aux spécifications.

3.1.2 Développement des Modèles D'inspection d'impression

Collecte et Préparation des Données

1. Collecte des Données

La collecte de données est une étape cruciale pour le développement de modèles d'apprentissage automatique précis et robustes. Pour ce faire, nous avons utilisé la caméra industrielle monochrome MER2-503-36U3M/C et le logiciel Galaxy Viewer de Daheng. Les données collectées comprennent des images optiques haute résolution des impressions de dates sur les films plats, avant leur formation en sachets.

2. **Conditions d'Enregistrement** : Pour ajuster la mise au point des caméras, nous avons utilisé le Galaxy Viewer. Nous avons pris des images individuelles à des intervalles définis en mesurant la vitesse de la ligne Volpack. La vitesse de la ligne dépend de la gamme de sachets produits :

TABLE 3.2 – Vitesse de production et intervalle de capture d'images pour différentes gammes de sachets

Gamme de sachets	Vitesse de production (sachets par minute)	Intervalle de capture d'images
300g	60	1 seconde
750g	35	1,7 secondes

3. **Formats de Capture** Les images ont été capturées aux formats BMP, JPG, et PNG. Le format BMP est recommandé pour une meilleure qualité et des détails plus fins, cependant, il nécessite plus de stockage et de temps de traitement. En effet, la taille des fichiers BMP est environ 40 fois plus grande que celle des formats JPG et PNG, ce qui peut poser des défis en termes de stockage et de rapidité de traitement.

Grâce à ce processus, nous avons réussi à collecter un total de 7800 photos pour traiter le problème de l'impression des dates.

4. **Prétraitement des Données** : Après la collecte des images, nous sommes passés à l'étape de prétraitement. Cette étape comprend plusieurs sous-étapes importantes :
 - **Nettoyage des Données** : La première étape du prétraitement consiste à nettoyer le jeu de données en éliminant manuellement les images de mauvaise qualité. Les principales causes de mauvaise qualité sont les défaillances et les décalages dus aux intervalles de 1 seconde ou 1,7 seconde, qui peuvent provoquer des erreurs. Il est donc crucial de surveiller étroitement le processus de collecte pour minimiser ces erreurs, bien que cela ne les élimine pas complètement. Cette phase de nettoyage vise à réduire autant que possible les images décalées.
 - **Redimensionnement** : Les images ont été redimensionnées à une résolution de 640X640 pour réduire la taille des fichiers et faciliter le traitement.
 - **Étirage de Contraste** : Cette technique de prétraitement améliore la qualité des images en augmentant le contraste, ce qui facilite la détection des défauts d'impression.
 - **Augmentation des Données** : Après le prétraitement, des opérations morphologiques et d'étirement de contraste ont été appliquées. Pour augmenter la diversité du jeu de données, une augmentation par un facteur de trois a été réalisée. Les méthodes utilisées pour l'augmentation sont les suivantes :

- **Rotation** : Pour compenser les changements d'angle de capture causés par les vibrations de la machine.
- **Flou** : Pour simuler les images floues dues aux vibrations de la machine.

5. Annotation des Données :

Les données doivent être annotées pour entraîner les modèles à reconnaître les défauts. Cette section décrit la méthode d'annotation utilisée.

- **Méthode d'Annotation** : L'annotation des données a été réalisée manuellement pour identifier les défauts d'impression sur les images collectées. Toutes les 7800 images, y compris les images floues ayant passé le processus de nettoyage, ont été annotées.
- **Outils d'Annotation** : L'outil utilisé pour l'annotation est Roboflow. Des boîtes englobantes ont été dessinées autour des zones pertinentes sur chaque image.

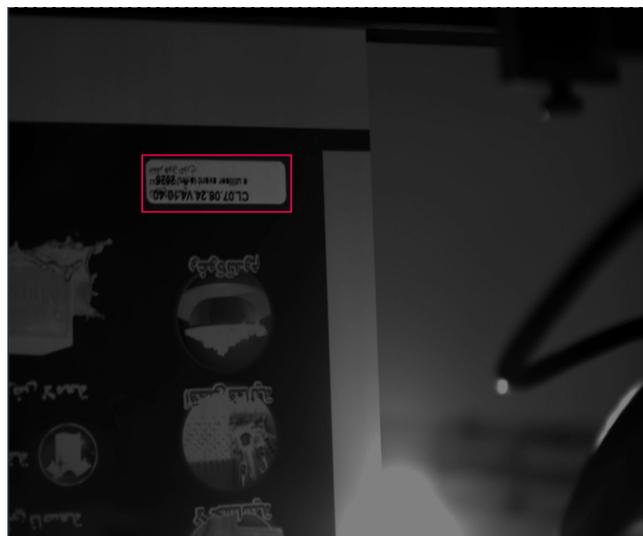


FIGURE 3.2 – Exemple d'annotation d'images avec Roboflow. Cette figure montre un exemple de l'interface Roboflow utilisée pour annoter les images avec des boîtes englobantes autour des zones d'intérêt.

6. Partition des Données :

La division des données en ensembles d'entraînement, de validation, et de test est une étape cruciale pour évaluer les performances des modèles.

Stratégie de Division : L'outil Roboflow a été utilisé pour diviser les données. Cette division permet de tester les modèles sur différentes parties du jeu de données, améliorant ainsi leur robustesse.

Proportions : Les données ont été divisées de manière à utiliser 70% pour l'entraînement, 15% pour la validation, et 15% pour les tests. Cette répartition permet de maximiser l'utilisation des données tout en assurant une évaluation rigoureuse des modèles.

TABLE 3.3 – Répartition des données pour l’entraînement, la validation et les tests.

Ensemble	Pourcentage	Nombre d’images
Entraînement	70%	3859
Validation	15%	827
Test	15%	827

3.1.3 Choix des Algorithmes et Évaluation des Modèles

Dans cette partie, j’introduis l’approche suivie pour résoudre le problème de détection de la zone d’impression sur les nouveaux emballages Henkel. Après avoir exploré différentes options, notamment les méthodes classiques de traitement d’image, nous avons finalement opté pour la segmentation d’instance utilisant des modèles pré-entraînés de la famille YOLO. Ce choix a été motivé par les défis spécifiques posés par les nouveaux designs d’emballage, qui présentent une zone brillante juste en dessous de la zone d’impression. Cette particularité rend difficile, voire impossible, la détection précise de la région d’intérêt en utilisant uniquement les techniques traditionnelles de traitement d’image. En effet, mes premiers essais avec des méthodes telles que le seuillage, la détection de contours et la segmentation basée sur les régions n’ont pas donné de résultats satisfaisants. La zone brillante interférait systématiquement avec la détection, conduisant à une segmentation imprécise de la zone d’impression.

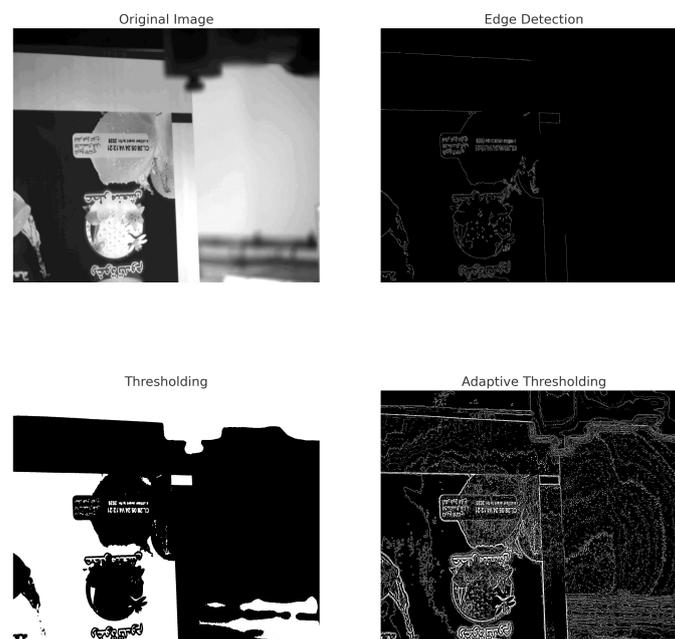


FIGURE 3.3 – Comparaison des méthodes classiques de traitement d’image pour la détection de la zone d’impression

Pour la segmentation et l’identification de la zone d’impression, nous avons utilisé l’apprentissage par transfert et le réglage fin d’un modèle de segmentation : YOLOv8m. Les avantages des modèles YOLO incluent la disponibilité de nombreuses versions adaptées à différents besoins. Dans notre cas, nous avons utilisé la version médium de YOLOv8 (YOLOv8m) pour la segmentation. Ce modèle a été choisi pour son équilibre entre vitesse et précision, ce qui est crucial pour notre application spécifique.

Le modèle YOLOv8m Seg utilisé dans notre étude présente les caractéristiques présenter

dans le tableau (table 3.4) :

TABLE 3.4 – YOLOv8m Segmentation

Caractéristiques	YOLOv8m Seg
Taille des images	640 pixels
mAP (box) 50-95	49.9
mAP (mask) 50	40.8
Paramètres (Mask)	27.3 millions
FLOPs (Box)	110.2

Ces caractéristiques démontrent la capacité du modèle YOLOv8m Seg à effectuer une segmentation précise tout en maintenant une complexité computationnelle raisonnable, ce qui est essentiel pour notre tâche d'identification de la zone d'impression

3.1.4 Evaluation du Modèle

L'évaluation du modèle YOLOv8 a été effectuée à l'aide de plusieurs indicateurs de performance, incluant les courbes de perte d'entraînement et de validation, les courbes de performance et la matrice de confusion.

Courbes de perte d'entraînement et de validation :

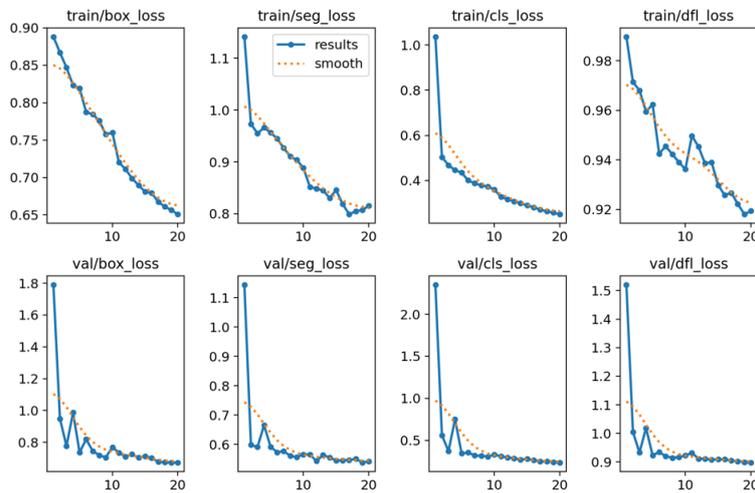


FIGURE 3.4 – Courbes de perte d'entraînement et de validation

Courbes de performance pour la validation et l'entraînement :

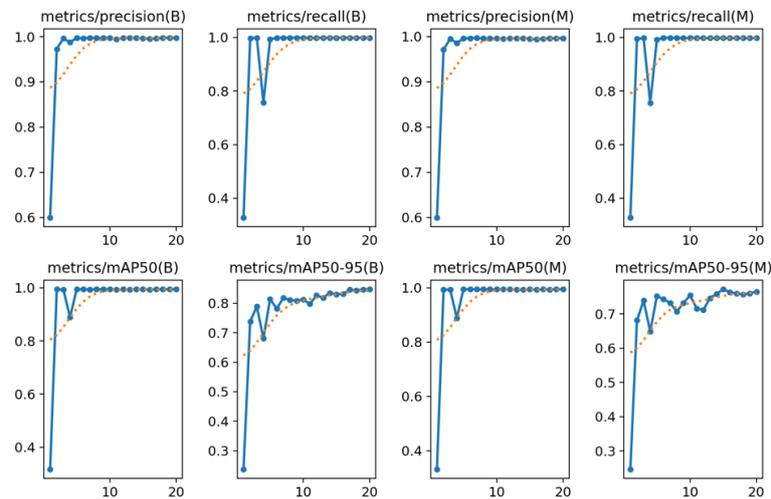


FIGURE 3.5 – Courbes de performance pour la validation et l'entraînement

Les résultats montrent une amélioration significative des performances du modèle au fil des époques. La diminution des pertes et l'augmentation des métriques de précision, de rappel et de mAP suggèrent que le modèle YOLOv8-l est efficace pour détecter les lignes de soudure des sachets de détergent. Cependant, la variabilité des pertes en validation indique que des ajustements supplémentaires peuvent être nécessaires pour améliorer la généralisation du modèle.

Résultats de l'entraînement du modèle YOLOv8 pour la segmentation de la zone d'impression :

Les résultats de l'entraînement du modèle YOLOv8 pour la segmentation de la zone d'impression sont présentés dans les figures (3.6) à (3.10) Ces courbes montrent la performance du modèle à travers diverses métriques, illustrant son efficacité et sa précision.

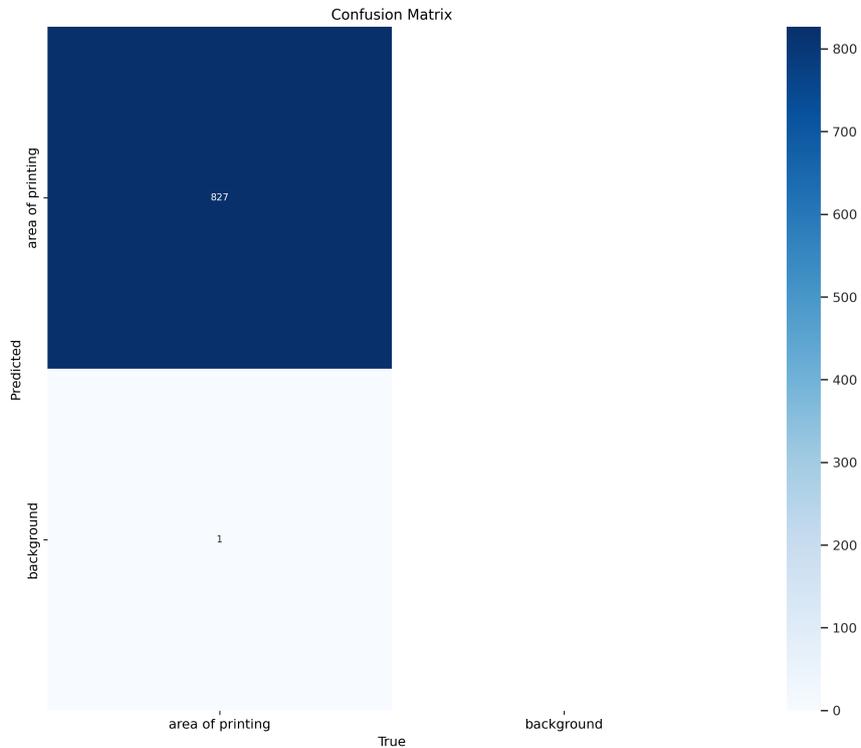
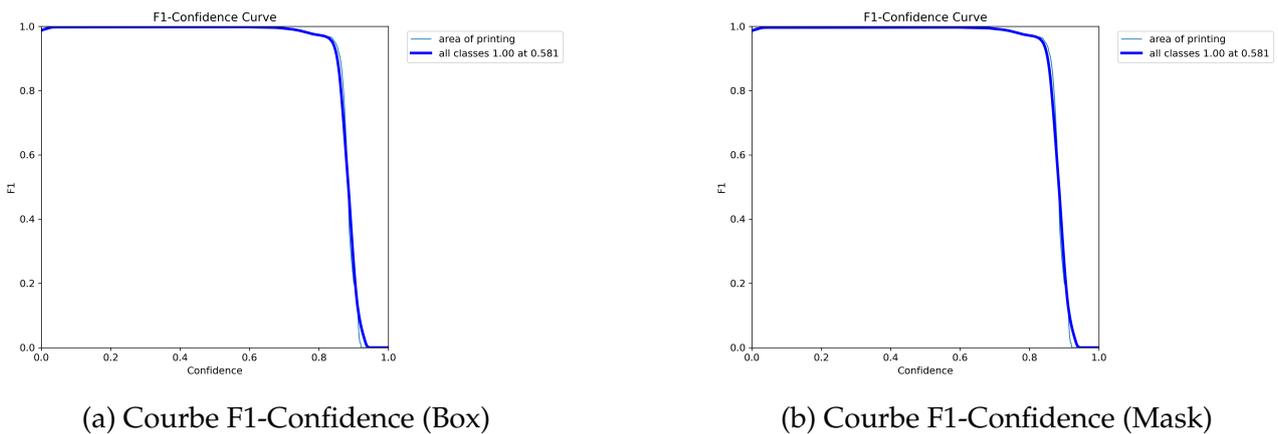


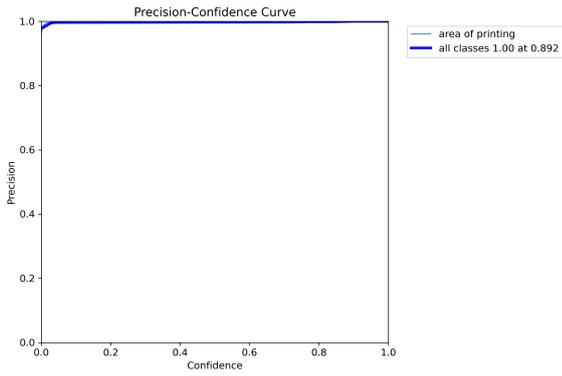
FIGURE 3.6 – Matrice de confusion



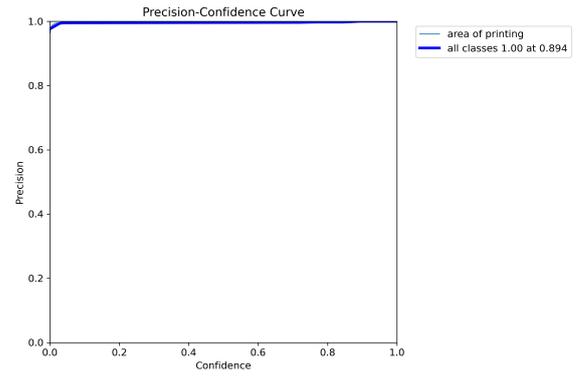
(a) Courbe F1-Confidence (Box)

(b) Courbe F1-Confidence (Mask)

FIGURE 3.7 – Courbes F1-Confidence pour box et mask

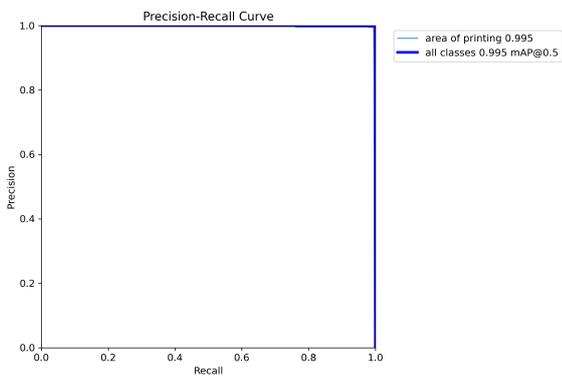


(a) Courbe Precision-Confidence (Box)

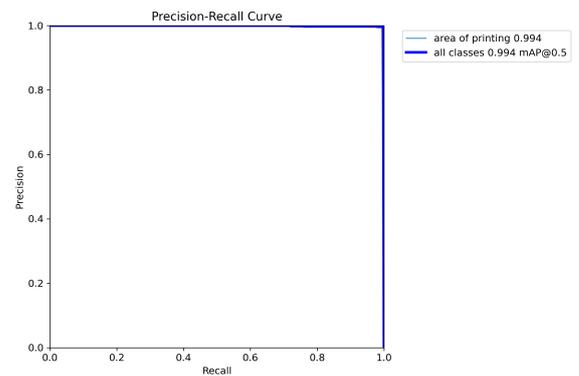


(b) Courbe Precision-Confidence (Mask)

FIGURE 3.8 – Courbes Precision-Confidence pour box et mask

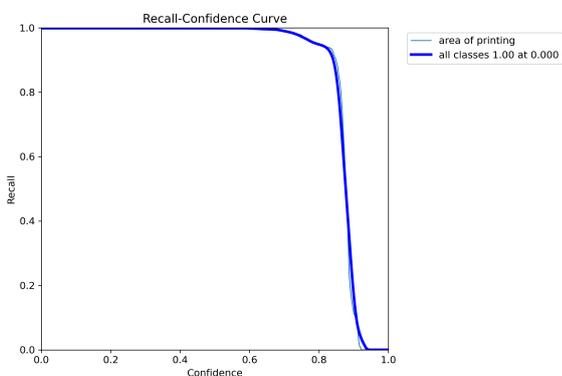


(a) Courbe Precision-Recall (Box)

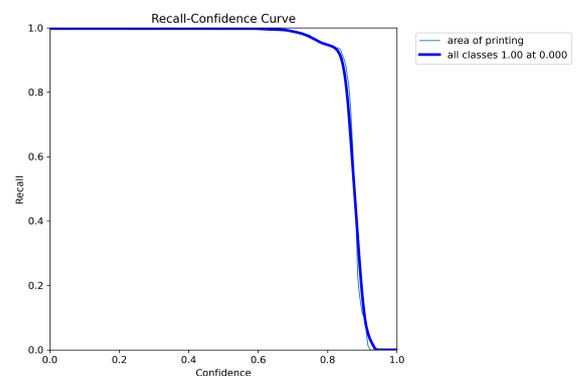


(b) Courbe Precision-Recall (Mask)

FIGURE 3.9 – Courbes Precision-Recall pour box et mask



(a) Courbe Recall-Confidence (Box)



(b) Courbe Recall-Confidence (Mask)

FIGURE 3.10 – Courbes Recall-Confidence pour box et mask,

Résultats de prédiction

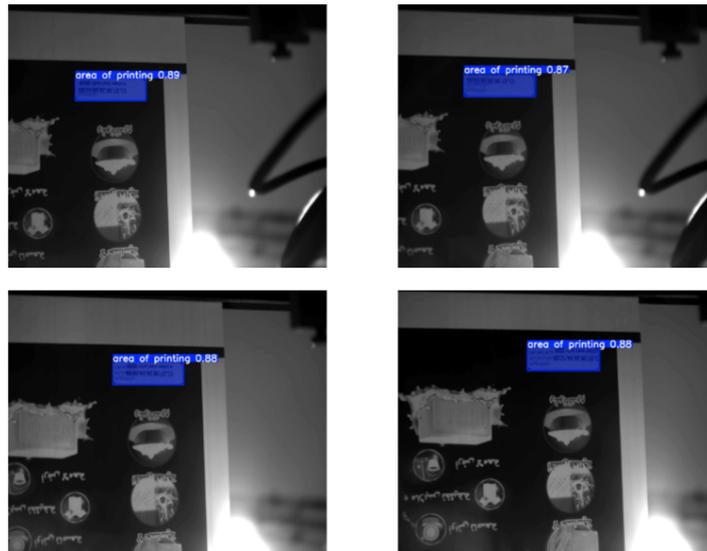


FIGURE 3.11 – Prédications du Modèle YOLOv8 montrant une segmentation précise des zones d'impression sur différentes images

Les résultats des courbes de performance et de la matrice de confusion démontrent la robustesse et l'efficacité du modèle YOLOv8 dans ses tâches de détection et de segmentation. Les valeurs élevées de précision, rappel et F1 à travers différents niveaux de confiance, ainsi que la forte performance de la matrice de confusion, indiquent que le modèle est bien entraîné et performant, avec une excellente capacité de généralisation aux données de validation. Ces résultats suggèrent que le modèle peut être utilisé de manière fiable dans des applications industrielles pour l'inspection et la segmentation des zones d'impression.

résultats de segmentation Après avoir appliqué la segmentation avec le modèle YOLOv8, la figure (3.12) représente les résultats obtenus.

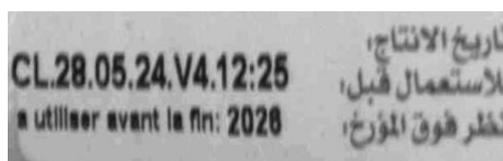


FIGURE 3.12 – Résultat de segmentation de modèle YoloV8

3.1.5 Reconnaissance Optique de Caractères (OCR)

Après la segmentation de la zone d'impression, une fonction OCR est utilisée pour inspecter la lisibilité du code imprimé. Cela permet de vérifier si la date est imprimée de manière lisible et conforme aux spécifications.

Voici une description de la fonction OCR utilisée :

- La fonction utilise une expression régulière pour rechercher et vérifier la lisibilité de la date et de l'heure imprimées sur les sachets.

- Elle vérifie la présence d'un motif spécifique qui inclut :
 - o Le préfixe "Cl."
 - o Une date au format JJ.MM.AA.
 - o Un indicateur de Volpack commençant par "V" suivi d'un chiffre de 1 à 9
 - o Une heure au format HH :MM
- La fonction s'assure que le format de la date et de l'heure est correct en tentant de les analyser avec la bibliothèque datetime.
- Si tous les critères sont satisfaits, la fonction considère que le texte est lisible et conforme. La figure 3.13 représente la décision de l'OCR sur la qualité de l'étiquetage.

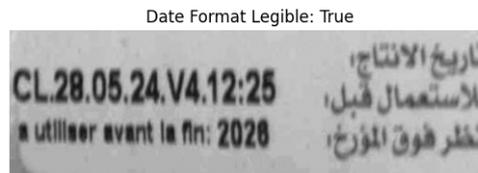


FIGURE 3.13 – Application de l'OCR sur la zone segmenté

Cette approche permet de vérifier non seulement la présence des informations requises, mais aussi leur format et leur validité, assurant ainsi une inspection approfondie de la qualité d'impression.

3.2 Partie2 :Inspection du Scellage

3.2.1 Architecture du Système

Description Technique des Dispositifs à Double Caméras

Pour l'inspection thermique des scellages, le système utilise des dispositifs à double caméras, comprenant une caméra thermique FLIR A35 et une caméra visuelle pour l'intégration au convoyeur. La caméra thermique FLIR A35 est choisie pour sa haute résolution thermique et sa capacité à détecter les variations de température avec précision, ce qui est crucial pour identifier les défauts de scellage.

Caractéristique	Valeur
Résolution	320 x 256 pixels
Plage de Température	-40°C à +550°C
Sensibilité Thermique (NETD)	< 50 mK
Taux de Rafraîchissement	60 Hz
Interface	GigE Vision

Intégration au Convoyeur : Pour l'inspection thermique des scellages, nous appliquons les principes de l'imagerie thermique pour détecter les défauts dans les lignes de scellage en fonction de la distribution de la chaleur sur le scellage. Une caméra thermique, spécifiquement la FLIR A35, est placée sur le convoyeur pour inspecter les sachets en mouvement.

La distance et le positionnement de la caméra sont cruciaux pour garantir une inspection efficace. Voici les détails du positionnement et de l'intégration de la caméra sur le convoyeur :

Positionnement des Caméras :

- **Hauteur et Angle :** La caméra thermique FLIR A35 est placée à une hauteur optimale pour capturer une vue claire et stable des lignes de scellage. L'angle de la caméra est ajustable selon le besoin pour assurer que toute la largeur du scellage est dans le champ de vision, ce qui permet une inspection complète et précise.
- **Distance et Timing :** La distance entre la caméra et la ligne de scellage n'est pas aléatoire. Elle est définie pour capturer les sachets exactement 2 secondes après l'opération de scellage. nous avons également défini la région d'intérêt (ROI) de la caméra pour voir le sachet après ces deux secondes, ce qui rend la distance entre la caméra et le sachet détecté d'environ 40 cm. Cela permet de détecter les variations thermiques sur le scellage sans interférences, en assurant que la chaleur du scellage soit bien distribuée et révélatrice de défauts potentiels.

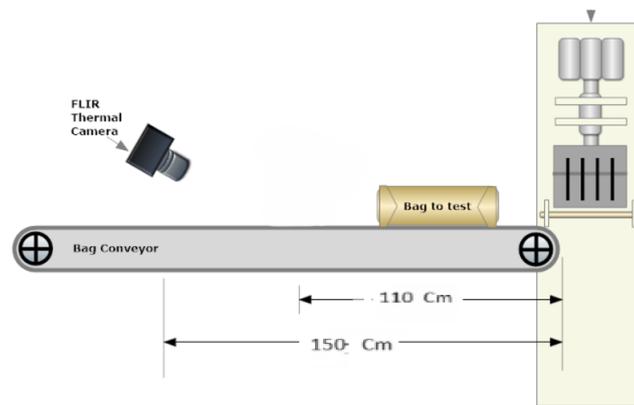


FIGURE 3.14 – Position de la caméra par rapport à la chaîne de conditionnement

Transmission des Données :

- **Capture et Transmission** : Les données thermiques capturées par la caméra sont transmises en temps réel à un système d'analyse dédié. Ce système évalue la qualité du scellage en se basant sur les variations thermiques détectées, permettant une inspection rapide et précise.
- **Analyse en Temps Réel** : La transmission continue des données assure que toute anomalie détectée peut être immédiatement signalée, permettant des interventions rapides pour corriger les défauts de scellage.

Fonctionnement du Système : Le fonctionnement du système repose sur une série d'étapes coordonnées pour garantir une inspection précise et fiable des scellages :

Acquisition des Images :

- La caméra thermique FLIR A35 capture l'image thermique des sachets en mouvement sur le convoyeur. La caméra est placée à une distance calculée pour capturer les sachets exactement 2 secondes après l'opération de scellage, ce qui permet de voir la distribution de la chaleur sur le scellage.

Inspection des Sachets :

- Les images thermiques capturées sont transmises en temps réel à un système d'analyse. Ce système utilise des algorithmes de traitement d'image pour analyser la distribution thermique et identifier les zones présentant des anomalies.

Signalement des Anomalies :

- Lorsqu'une anomalie est détectée, le système signale immédiatement l'erreur pour permettre une intervention rapide. Les données sont également enregistrées pour une analyse ultérieure et pour améliorer les processus de qualité.

Avantages et Inconvénients des Caméras Thermiques : Les caméras thermiques offrent de nombreux avantages pour l’inspection des scellages, mais elles présentent également certains inconvénients. Voici un tableau récapitulatif des principaux avantages et inconvénients :

TABLE 3.5 – Avantages et inconvénients de la caméra thermique FLIR A35

Avantages	Inconvénients
Détection précise des variations thermiques	Coût élevé des caméras thermiques
Inspection en temps réel	Nécessité d’un étalonnage régulier
Indépendance par rapport aux conditions d’éclairage	Sensibilité aux interférences environnementales
Capacité à détecter des défauts invisibles à l’œil nu	Maintenance et support technique requis
Intégration facile avec les systèmes d’analyse automatisés	Complexité de l’analyse des données thermiques
Amélioration de la qualité des produits	

Cette configuration permet de garantir une inspection rigoureuse et continue des scellages, assurant ainsi la qualité et l’intégrité des produits en sortie de ligne de production.

3.2.2 Collecte et Préparation des Données :

Collecte des Données La collecte des données est une étape cruciale dans le développement des modèles d’apprentissage automatique pour l’inspection thermique des scellages. Cette section détaille les spécifications des caméras, les conditions d’enregistrement, et les types de données recueillies.

Conditions d’Enregistrement :

- **Environnement de Production :** Les données sont collectées dans un environnement de production réel, où les sachets de détergent passent sur un convoyeur après l’opération de scellage.
- **Réglages de la Caméra :** La caméra thermique est positionnée de manière à capturer les sachets exactement 2 secondes après l’opération de scellage, à une distance d’environ 150 cm de la ligne de scellage. La région d’intérêt (ROI) de la caméra est définie pour voir le sachet après ces deux secondes, avec une distance d’environ 40 cm entre la caméra et le sachet détecté.
- **Fréquence de Capture :** En utilisant MATLAB pour connecter la caméra FLIR AX5, les images sont acquises grâce à l’outil d’acquisition d’images, qui supporte la transmission GigE Vision et fonctionne parfaitement avec le modèle FLIR AX5. Toutefois, cet outil ne supporte pas la capture automatisée d’une seule image à des intervalles de temps définis. Les options disponibles étaient l’acquisition continue de vidéos ou la mise en place d’un déclencheur matériel (capteur ultrasonique ou cellule photoélectrique). Dans mon cas, la seconde option n’était pas disponible, donc nous avons dû capturer des vidéos continues.

Développement d’une Interface Utilisateur pour l’Acquisition d’Images : Pour faciliter le processus d’acquisition des images à partir des vidéos, nous avons développé une interface utilisateur. Cette interface simplifie la gestion et l’extraction des images en fournissant des

outils interactifs pour configurer et contrôler le processus de capture. Voici les fonctionnalités principales de l'interface :

- **Sélection et Chargement de Vidéo :** L'interface permet de sélectionner un fichier vidéo à partir de l'ordinateur. Une fois la vidéo sélectionnée, elle est chargée et prête pour la capture des images.
- **Définition du Répertoire de Capture :** L'utilisateur peut définir le répertoire où les images capturées seront stockées. Cela permet une organisation claire et structurée des données.
- **Configuration des Paramètres de Capture :**
 - o **Temps de Départ :** L'utilisateur peut spécifier le moment à partir duquel commencer la capture des images (en secondes).
 - o **Capture :** L'utilisateur définit l'intervalle de temps entre chaque image capturée (en secondes).
 - o **Mode de Capture :** L'interface offre deux modes de capture : continuer jusqu'à la fin de la vidéo ou capturer jusqu'à un moment spécifique.
- **Contrôle de la Lecture Vidéo :** Des boutons permettent de contrôler la lecture de la vidéo (démarrer, mettre en pause, avancer ou reculer). Cela facilite la navigation et la sélection des moments clés pour la capture des images.
- **Capture d'Images :**
 - o **Capture Continue :** L'interface permet de capturer des images à intervalles réguliers à partir du temps de départ défini.
 - o **Capture Unique :** En raison des variations possibles dans les intervalles de temps sur la ligne de production, une fonction de capture unique est disponible. Cette fonctionnalité permet de capturer une image unique à partir du point actuel dans la vidéo, ce qui est utile pour corriger les décalages temporels.
- **Affichage en Temps Réel :** L'image actuelle de la vidéo est affichée en temps réel dans l'interface, avec un indicateur de temps. Cela permet à l'utilisateur de visualiser et de vérifier les images avant et pendant la capture. La figure 3.15 montre l'interface.



FIGURE 3.15 – Interface

L'interface utilisateur développée simplifie grandement le processus d'acquisition des images à partir des vidéos. Elle assure une collecte de données flexible et robuste, adaptée aux environnements de production réels où les conditions peuvent varier. Les images capturées fournissent des informations cruciales pour le développement de modèles d'apprentissage automatique capables de détecter les défauts de scellage en temps réel.

Prétraitement des Données : Cette section décrit les différentes étapes du prétraitement des données, en se concentrant principalement sur le nettoyage des données en raison de son importance cruciale dans le contexte d'une ligne de production.

Nettoyage des Données : Sur une ligne de production, le principal défi est le décalage temporel après avoir défini l'intervalle de capture, ce qui peut entraîner des erreurs de synchronisation. Les étapes de nettoyage incluent :

- **Suppression des Images Floues :** Identification et suppression des images floues ou mal capturées pour garantir que seules les images de haute qualité sont utilisées pour l'entraînement des modèles.
- **Correction des Artéfacts :** Correction des artéfacts ou des imperfections dans les images, telles que les taches ou les rayures, qui pourraient déformer les données.
- **Uniformité des Images :** Assurer que toutes les images capturées ont des dimensions et une résolution uniformes pour maintenir la cohérence dans l'entraînement des modèles.
- **Gestion des Décalages Temporels :** Sur une ligne de production, la possibilité de décalage temporel est élevée en raison des arrêts de la machine ou des changements de chemin de scellage. Ces décalages peuvent entraîner des erreurs de synchronisation dans la capture des images. Une attention particulière est portée à l'alignement correct des images en fonction de l'intervalle de temps défini pour garantir une représentation précise des conditions de production.

Les étapes de normalisation et d'augmentation des données seront ajustées selon les besoins spécifiques des modèles d'apprentissage automatique utilisés. Cependant, le nettoyage des données reste la priorité pour s'assurer que les images sont utilisables et représentatives des scénarios de production réels.

3.2.3 Développement des Modèles d'inspection de scellage

Conception des Modèles

Afin de développer un système d'inspection thermique des scellages performant et complet, il est crucial d'évaluer et de comparer plusieurs approches de modélisation. Deux approches principales ont été sélectionnées pour ce projet : la détection d'objets et la segmentation d'instance. Chaque approche offre des avantages distincts et permet de traiter différents aspects des défauts de scellage, garantissant ainsi une inspection précise et exhaustive.

Ces approches complémentaires permettent de développer un système robuste et polyvalent, capable de couvrir tous les aspects des défauts de scellage et de s'assurer que les modèles utilisés sont rigoureusement évalués. Dans les sections suivantes, nous détaillerons chacune de ces approches, les modèles utilisés, le processus de développement, et les résultats obtenus.

Approche 1 : Détection d'Objets (Object Detection)

Cette approche vise à utiliser des modèles de détection d'objets pour identifier et classer deux types principaux de scellages dans les images : "scellage correct" et "scellage incorrect". Les modèles de détection d'objets permettent de localiser et d'identifier les défauts de scellage de manière efficace.

Modèles Utilisés : Pour cette approche nous avons choisi d'utiliser les modèles YOLOv8 et YOLOv9. Ces modèles sont connus pour leur rapidité et leur précision dans la détection d'objets en temps réel.

Processus de Développement :

- Préparation des Données :

- Les images collectées et annotées avec des boîtes de délimitation (bounding boxes) sont utilisées comme données d'entraînement.



FIGURE 3.16 – annotation avec roboflow

- Les images sont redimensionnées à une taille standard (par exemple, 640x640 pixels) pour être compatibles avec les modèles YOLO.
- **Augmentation des Données :** Pour améliorer la robustesse des modèles, des techniques d'augmentation des données telles que la rotation, le recadrage et le changement de luminosité ont été appliquées.
- **Entraînement des Modèles :**
 - Les modèles YOLOv8 et YOLOv9 sont entraînés sur les données annotées.
 - Des techniques de transfert de connaissances (transfer learning) sont utilisées pour améliorer l'efficacité de l'entraînement.
- **Évaluation des Modèles :**

- Les modèles sont évalués en utilisant des ensembles de données de validation et de test pour mesurer leur performance.
- Les métriques d'évaluation incluent la précision, le rappel, et le mAP (mean Average Precision) à différents seuils d'IoU (Intersection over Union).

Visualisation des Performances des Modèles : Les graphiques dans la figure (3.17) et (3.18) illustrent les différentes métriques d'évaluation des modèles YOLOv8 et YOLOv9 au cours des 20 époques d'entraînement :

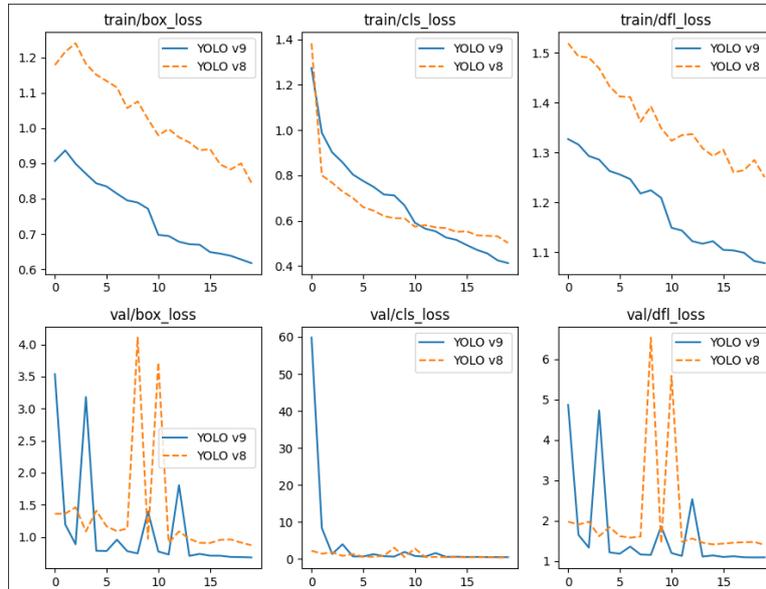


FIGURE 3.17 – Graphiques des métriques d'évaluation des modèles

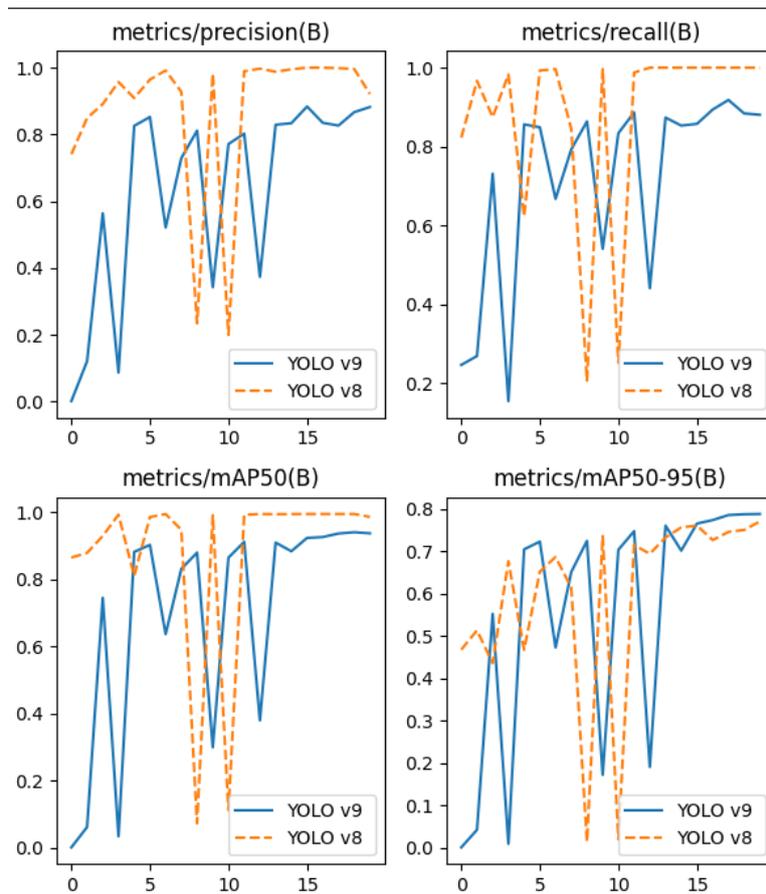


FIGURE 3.18 – Graphiques des métriques d'évaluation des modèles

Résultats : Les résultats des modèles sont comparés pour déterminer lequel offre les meilleures performances pour l'inspection des scellages. Voici un résumé des performances des deux modèles :

TABLE 3.6 – Comparaison des performances entre YOLO v9 et YOLO v8

Métrique	YOLO v9	YOLO v8	Interprétation
train/box_loss	Commence à 1.2, finit à 0.6 (50% de réduction)	Commence à 1.1, finit à 0.9 (18% de réduction)	YOLO v9 montre une amélioration significative dans la précision de la prédiction des boîtes englobantes.
train/cls_loss	Commence à 1.4, finit à 0.4 (71% de réduction)	Commence à 1.2, finit à 0.6 (50% de réduction)	YOLO v9 s'améliore plus significativement en précision de classification.
train/df_l_loss	Commence à 1.5, finit à 1.1 (27% de réduction)	Commence à 1.4, finit à 1.3 (7% de réduction)	YOLO v9 montre une meilleure amélioration dans la distribution de confiance sur les classes.

Métrique	YOLO v9	YOLO v8	Interprétation
val/box_loss	Très variable, moyenne 1.5	Très variable, moyenne 1.2	YOLO v8 a une perte de validation de boîte englobante moyenne plus basse, indiquant une meilleure généralisation.
val/cls_loss	Très variable, moyenne 10	Très variable, moyenne 15	Les deux modèles sont instables, mais YOLO v8 montre une tendance plus cohérente.
val/df_l_loss	Très variable, moyenne 2	Très variable, moyenne 2.5	YOLO v9 performe légèrement mieux dans la distribution de confiance en validation.
metrics/precision(B)	Très variable, moyenne 0.6	Plus stable, moyenne 0.9	YOLO v8 a une précision plus élevée et plus stable.
metrics/recall(B)	Très variable, moyenne 0.7	Plus stable, moyenne 0.9	YOLO v8 a un rappel plus élevé et plus stable.
metrics/mAP50(B)	Très variable, moyenne 0.5	Plus élevé et plus stable, moyenne 0.7	YOLO v8 montre une meilleure précision moyenne et stabilité.
metrics/mAP50-95(B)	Très variable, moyenne 0.4	Plus élevé et plus stable, moyenne 0.6	YOLO v8 performe mieux à travers différents seuils IoU.

D'après la comparaison des métriques de performance, YOLOv8 montre une meilleure stabilité et une précision plus élevée à travers différents critères d'évaluation par rapport à YOLOv9. Par conséquent, nous continuerons avec YOLOv8 pour détecter les classes défectueuses dans notre système d'inspection.

Exemples de Détections : Les figures (3.19) et (3.20) montrent des exemples de détections effectuées par le modèle YOLOv8 sur des images extérieures au jeu de données original.

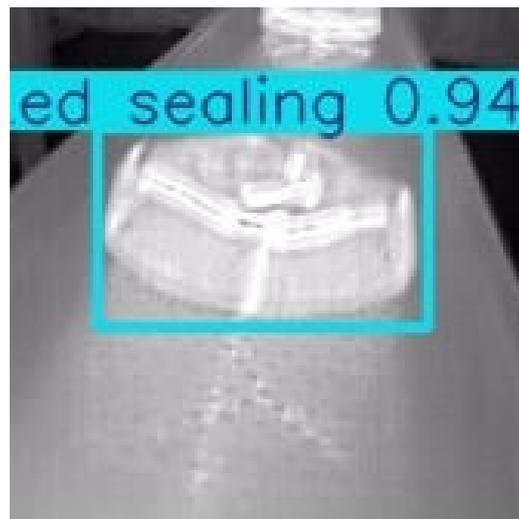


FIGURE 3.19 – Détection d'un scellage incorrect avec une précision de 0.94



FIGURE 3.20 – Détection d'un scellage correct avec une précision de 0.88

Conclusion : Les résultats de cette première approche indiquent que YOLOv8 est le modèle le plus performant pour l'inspection thermique des scellages, offrant à la fois rapidité et précision.

Approche 2 : Segmentation d'Instances et Analyse des Lignes de Scellage Objectif

Cette approche vise à combiner la segmentation d'instances pour isoler précisément la zone de scellage des sachets de détergent et une analyse des lignes pour évaluer la qualité du scellage. L'objectif est de détecter avec précision les zones de scellage et de déterminer si le scellage est correct ou défectueux.

Modèle Utilisé

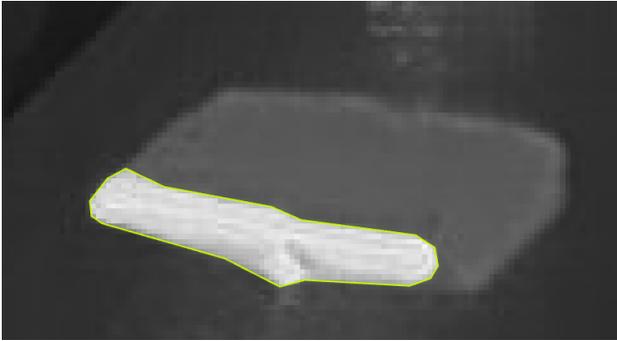
Le modèle utilisé pour cette approche est YOLOv8 Large (YOLOv8-l), qui offre un bon compromis entre la précision et la vitesse d'inférence. Les caractéristiques du modèle sont les suivantes :

- Taille des images : 640x640 pixels
- mAPval 50-95 : 52.9
- Vitesse CPU ONNX : 375.2 ms
- Vitesse A100 TensorRT : 2.39 ms
- Nombre de paramètres : 43.7 M
- FLOPs : 165.2 B

Processus de Développement : Le processus de développement de cette approche comprend deux étapes principales :

Préparation des Données : L'annotation des images a été réalisée à l'aide de RoboFlow, en utilisant l'option polygone pour couvrir précisément la forme de la zone de scellage. Cela permet d'obtenir des annotations précises pour la segmentation d'instances. Afin de pallier le manque de données, une augmentation des données a été effectuée avec RoboFlow. Le jeu de données initial a été augmenté jusqu'à trois fois en appliquant diverses techniques d'augmentation, telles que :

- Édition de la Saturation : Ajustement de la saturation des images entre -25
- Blur : Application d'un flou jusqu'à 2.5 pixels pour simuler des variations de mise au point.
- Noise : Ajout de bruit à jusqu'à 0.1
- Égalisation d'Histogramme Adaptative : Amélioration du contraste des images pour mieux révéler les détails dans les zones de scellage.



fvf<dth.png

FIGURE 3.21 – annotation avec roboflow

Résultats et Interprétation des Performances du Modèle

TABLE 3.7 – Évolution des pertes du modèle YOLOv8 Large (YOLOv8-l) pour la détection de la qualité du scellage

Type de Pertes	Valeur Initiale	Valeur Finale	Interprétation
Train/Box Loss	1.7	1.1	Diminution régulière, amélioration de la localisation des boîtes.
Train/Seg Loss	3.2	2.2	Diminution modérée, amélioration de la segmentation des zones de scellage.
Train/Cls Loss	1.6	0.6	Amélioration significative de la capacité à classer correctement les scellages.
Train/Dfl Loss	1.5	1.1	Légère amélioration de la distribution de confiance du modèle sur les classes prédites.
Val/Box Loss	1.5	1.3	Variabilité, mais tendance à la diminution, meilleure généralisation aux données de validation.
Val/Seg Loss	3.1	2.6	Variabilité, mais tendance à la diminution, meilleure segmentation sur les données de validation.
Val/Cls Loss	2.5	1.0	Diminution, meilleure performance de classification sur les données non vues.
Val/Dfl Loss	1.8	1.5	Diminution, meilleure distribution de confiance sur les données de validation.

TABLE 3.8 – Performances des métriques du modèle YOLOv8 Large (YOLOv8-l) pour la détection de la qualité du scellage

Métrique	Valeur Initiale	Valeur Finale	Interprétation
Precision (B)	0.1	0.9	Augmentation notable, meilleure précision pour détecter les scellages.
Recall (B)	0.5	0.8	Amélioration continue, meilleur rappel pour détecter les scellages.
mAP50 (B)	0.08	0.82	Amélioration substantielle, meilleure précision moyenne du modèle.
mAP50-95 (B)	0.03	0.46	Progrès important, meilleure performance sur une gamme de seuils IoU.
Precision (M)	0.1	0.88	Augmentation notable, meilleure précision pour segmenter les scellages.
Recall (M)	0.5	0.78	Amélioration continue, meilleur rappel pour segmenter les scellages.
mAP50 (M)	0.09	0.79	Amélioration substantielle, meilleure précision moyenne pour les masques.
mAP50-95 (M)	0.03	0.38	Progrès important, meilleure performance sur une gamme de seuils IoU pour les masques.

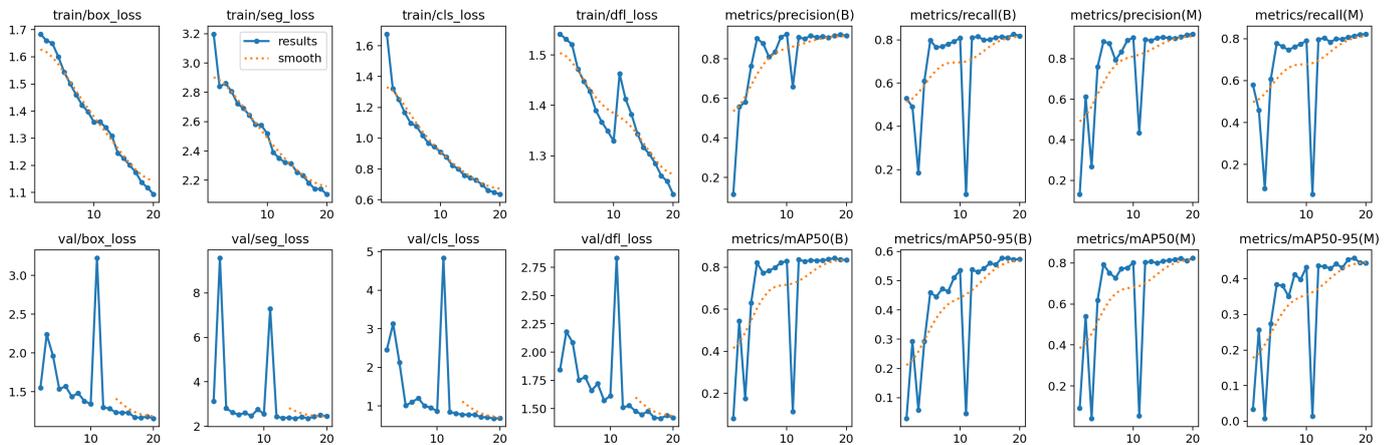


FIGURE 3.22 – Graphiques des resultat

Commentaire : Les résultats montrent une amélioration significative des performances du modèle au fil des époques. La diminution des pertes et l’augmentation des métriques de précision, de rappel et de mAP suggèrent que le modèle YOLOv8-l est efficace pour détecter les lignes de soudure des sachets de détergent. Cependant, la variabilité des pertes en validation indique que des ajustements supplémentaires peuvent être nécessaires pour améliorer la généralisation du modèle.

Algorithme de Détection de la Qualité du Scellage Après avoir segmenté les zones de scellage avec précision en utilisant le modèle YOLOv8-l, l’étape suivante consiste à analyser les lignes de scellage pour déterminer leur qualité. Cet algorithme utilise des techniques de traitement d’image pour évaluer la qualité du scellage.

1. **Détection des Bords** : La première étape consiste à utiliser l'algorithme de Canny pour détecter les bords dans les zones de scellage. L'image est d'abord convertie en niveaux de gris, puis l'algorithme de Canny est appliqué pour identifier les contours des lignes de scellage. Cette étape permet de mettre en évidence les caractéristiques importantes des zones de scellage.
2. **Opérations Morphologiques** : Une fois les bords détectés, des opérations morphologiques, telles que la dilatation, sont appliquées pour renforcer les bords et éliminer les petits bruits. La dilatation morphologique est réalisée à l'aide d'un noyau spécifique, qui permet d'épaissir les contours des lignes de scellage. Cette étape contribue à améliorer la qualité des bords détectés et à faciliter la détection des lignes dans les étapes suivantes.
3. **Détection des Lignes** : L'étape suivante consiste à utiliser la transformation de Hough pour détecter les lignes dans les zones de scellage. La transformation de Hough est une technique puissante pour identifier les lignes droites dans une image. Elle est appliquée sur l'image résultante des opérations morphologiques, qui a été prétraitée pour mettre en évidence les bords. La transformation de Hough permet de localiser les lignes de scellage de manière précise.
4. **Analyse des Lignes Détectées** : La dernière étape de l'algorithme consiste à analyser les lignes détectées pour déterminer la qualité du scellage. Le nombre de lignes détectées est utilisé comme critère principal pour évaluer la qualité du scellage. Si une seule ligne est détectée dans la zone de scellage, cela indique que le scellage est correct et de bonne qualité. En revanche, si plusieurs lignes sont détectées ou si aucune ligne n'est détectée, cela suggère que le scellage est défectueux ou incomplet.



FIGURE 3.23 – Détection d'un scellage incorrect

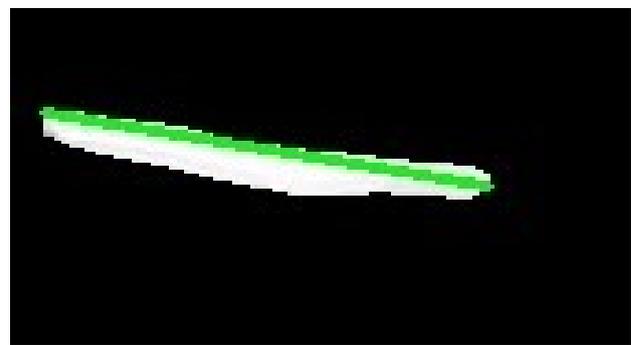


FIGURE 3.24 – Détection d'un scellage correct

Cette analyse des lignes détectées permet de prendre une décision automatisée sur la qualité du scellage, en se basant sur des critères objectifs. L'algorithme peut ainsi identifier les sachets de détergent correctement scellés et ceux présentant des défauts de scellage.

Évaluation des Approches :

Ces approches complémentaires permettent de développer un système robuste et polyvalent pour l'inspection thermique des scellages, en s'assurant que tous les aspects des défauts de scellage sont couverts et que les modèles utilisés sont évalués de manière exhaustive.

Concept de chaque approche

TABLE 3.9 – Concept de chaque approche

Approche	Objectif	Modèles Utilisés	Processus de Développement
Détection d'Objets (Object Detection)	Identifier et classer les scellages comme "correct" ou "incorrect".	YOLOv8, YOLOv9	Préparation des données, Entraînement des modèles, Évaluation des modèles
Segmentation d'Instances et Analyse des Lignes de Scellage	Segmenter précisément la zone de scellage et analyser les lignes pour évaluer la qualité du scellage.	YOLOv8 Large (YOLOv8-l)	Préparation des données, Entraînement des modèles, Analyse des lignes de scellage

Caractéristiques des approches

TABLE 3.10 – Caractéristiques des approches

Approche	Avantages	Inconvénients
Détection d'Objets (Object Detection)	<ul style="list-style-type: none"> - Implémentation simple, rapidité d'exécution. - Efficacité pour la classification des scellages en "correct" et "incorrect". 	<ul style="list-style-type: none"> - Peut manquer de précision pour les petits défauts. - Nécessite une grande quantité de données annotées.
Segmentation d'Instances et Analyse des Lignes de Scellage	<ul style="list-style-type: none"> - Haute précision pour la localisation des défauts. - Analyse détaillée des lignes de scellage. - Capacité d'adaptation aux variations de forme et de texture des scellages. 	<ul style="list-style-type: none"> - Plus complexe à implémenter. - Temps de calcul plus long. - Nécessite une annotation précise et une augmentation des données pour de bons résultats.

Ces tableaux offrent une vue d'ensemble des concepts, avantages, et inconvénients de chaque approche, permettant une évaluation claire et comparative pour le développement d'un système de contrôle qualité pour l'inspection thermique des scellages.

3.3 Conclusion

Ce chapitre a présenté le développement de notre système d'inspection de la qualité des scellages des sachets de détergent. Deux approches principales ont été explorées : la détection d'objets et la segmentation d'instances avec analyse des lignes de scellage. Les résultats montrent que YOLOv8 est le modèle le plus performant, offrant à la fois rapidité et précision. L'analyse des performances indique une amélioration significative des métriques de précision et de rappel au fil des itérations, confirmant l'efficacité des modèles utilisés. En combinant

ces approches, nous avons développé un système robuste et polyvalent, capable de détecter et d'évaluer avec précision les défauts de scellage.

Chapitre 4

Analyse de Rentabilité pour le Projet de Système d'Inspection de Qualité

Le projet vise à réduire le taux de réclamations des clients et à augmenter le taux de détection des défauts au niveau de l'usine en implémentant un système d'inspection de qualité utilisant des caméras thermiques et monochromes. L'objectif est d'améliorer la qualité des produits en identifiant et en corrigeant les défauts dès le stade de la production, réduisant ainsi les coûts liés aux retours de produits défectueux.

4.1 Investissement Initial

L'investissement initial pour le projet comprend l'achat de 10 caméras thermiques et 10 caméras monochromes par ligne de production, pour un total de 10 lignes. Le coût total de l'investissement est calculé en utilisant les coûts unitaires de chaque type de caméra et leur nombre, multiplié par le nombre de lignes. Le coût total en euros est ensuite converti en dinars algériens (DA) en utilisant un taux de conversion de 1 euro pour 150 DA.

Soit :

- N le nombre de lignes de production (10),
- C_T le coût unitaire d'une caméra thermique (5800 euros),
- C_M le coût unitaire d'une caméra monochrome (400 euros),
- T le taux de conversion euro en DA (150).

Le coût total en euros (I) est donné par :

$$I = N \times (10 \times C_T + 10 \times C_M)$$

$$I = 10 \times (10 \times 5800 + 10 \times 400) = 10 \times (58000 + 4000) = 10 \times 62000 = 620000 \text{ euros}$$

Le coût total en DA (I_{DA}) est donné par :

$$I_{DA} = I \times T$$

$$I_{DA} = 620000 \times 150 = 93000000 \text{ DA}$$

4.2 Économies Réalisées Grâce à la Réduction des Défauts

Pour estimer les économies réalisées grâce à la réduction des défauts, nous avons utilisé un taux de réduction des défauts de 80%. Ce taux signifie que le nouveau taux de défaut sera de 20% du taux initial.

4.2.1 Détails des Coûts de Retour

Pour chaque taille de produit, les calculs suivants sont appliqués en utilisant les coûts de retour originaux et les prix de vente fournis :

TABLE 4.1 – Détails des Coûts de Retour

Taille	CRO (DA)	PV (DA)	PCR (%)
300g	42.06	50.62	83.08%
750g	93.84	103.27	90.86%
1.5kg	180.42	198.74	90.79%

Note : CRO - Coût de retour original, PV - Prix de vente, PCR - Pourcentage de coût de retour

4.2.2 Calcul des Coûts de Retour Totaux

Pour chaque taille de produit, nous utilisons les coûts de retour totaux pour l'estimation.

Hypothèse : Le taux de production total est de 5,3 tonnes de produits défectueux, ce qui représente 2,3% de la production totale.

Calcul de la Production Totale

Pour estimer la production totale, nous utilisons le taux de défaut global :

$$\text{Production Totale} = \frac{\text{Poids des Défauts Signalés}}{\text{Taux de Défaut}}$$

$$\text{Production Totale} = \frac{5,300 \text{ kg}}{0.023} = 230,434.78 \text{ kg}$$

TABLE 4.2 – Répartition des Défauts par Taille de Produit

Taille du produit	Pourcentage de production totale (%)	Poids des défauts (kg)
300g	55	126,739.13
750g	28	64,522.43
1.5kg	17	39,173.22

TABLE 4.3 – Coûts de Retour Totaux

Taille du produit	Poids des défauts (kg)	Coût de retour original (DA)	Coût de retour réduit (DA)	Économies nettes (DA)
300g	126,739.13	42.06 × 126,739.13 = 5,329,459.88	42.06 × 0.2 × 126,739.13 = 1,065,891.98	5,329,459.88 – 1,065,891.98 = 4,263,567.90
750g	64,522.43	93.84 × 64,522.43 = 6,055,307.52	93.84 × 0.2 × 64,522.43 = 1,211,061.50	6,055,307.52 – 1,211,061.50 = 4,844,246.02
1.5kg	39,173.22	180.42 × 39,173.22 = 7,068,593.91	180.42 × 0.2 × 39,173.22 = 1,413,718.78	7,068,593.91 – 1,413,718.78 = 5,654,875.13

4.3 Calcul du Temps de Récupération de l'Investissement

Le temps de récupération de l'investissement est calculé en divisant le coût total de l'investissement par les économies nettes annuelles.

$$T = \frac{\text{Coût Total de l'Investissement (DA)}}{\text{Économies Nettes Annuelles (DA)}}$$

En utilisant les économies nettes totales pour les trois tailles de produit :

$$\text{Économies Nettes Totales} = 4,263,567.90 \text{ DA} + 4,844,246.02 \text{ DA} + 5,654,875.13 \text{ DA} = 14,762,689.05 \text{ DA}$$

$$T = \frac{93,000,000 \text{ DA}}{14,762,689.05 \text{ DA}} \approx 6.30 \text{ années} \approx 75 \text{ mois}$$

4.4 Conclusion

En conclusion, la mise en œuvre du système d'inspection de qualité permettra de réduire considérablement les coûts de retour des produits défectueux, d'améliorer la qualité des produits et de réaliser des économies substantielles. Toutefois, le temps de récupération de l'investissement initial sera d'environ 6,3 années. Cela démontre que bien que le projet soit bénéfique à long terme, il nécessite une période plus longue pour récupérer les coûts initiaux.

Conclusion générale

L'application des technologies de vision industrielle et d'apprentissage automatique pour l'inspection de la qualité dans le processus de production de détergent en poudre chez Henkel a montré des résultats prometteurs. Le développement d'un système de contrôle qualité automatisé, incluant l'inspection de l'impression des dates et du scellage thermique des sachets, a été détaillé et a démontré des améliorations significatives en termes de précision et d'efficacité.

Cette étude a commencé par la présentation des concepts fondamentaux de la vision industrielle et de l'apprentissage automatique, en mettant l'accent sur les techniques d'apprentissage supervisé, non supervisé, et profond, ainsi que sur les réseaux de neurones convolutifs (CNN) et l'apprentissage par transfert. Ces technologies ont prouvé leur potentiel dans l'automatisation et l'optimisation des processus de contrôle qualité.

En examinant le contexte industriel de Henkel, l'analyse a révélé l'importance cruciale d'un système efficace de détection des défauts pour réduire les coûts de production et améliorer la satisfaction des clients. Le cœur de ce mémoire a porté sur la conception et le développement du système de contrôle qualité, détaillant l'architecture du système, les équipements utilisés, et les méthodologies de collecte et de prétraitement des données. Les modèles d'apprentissage automatique développés pour l'inspection de l'impression des dates et du scellage thermique ont été évalués, démontrant des améliorations significatives.

Les résultats obtenus montrent que l'intégration de techniques avancées de vision industrielle et d'apprentissage automatique permet de détecter les défauts avec une grande précision, contribuant ainsi à l'amélioration globale du processus de production. La réduction des non-conformités détectées par les clients et l'amélioration de la qualité des produits attestent de l'efficacité du système proposé.

De plus, la rentabilité du projet a été analysée, montrant des économies substantielles grâce à la réduction des défauts, avec un retour sur investissement prévu en moins de deux ans. Les économies réalisées grâce à la réduction des non-conformités et l'amélioration de l'efficacité de la production illustrent le potentiel économique du système proposé.

En conclusion, ce mémoire a démontré que l'application de technologies de pointe dans le contrôle qualité offre des avantages significatifs pour l'industrie manufacturière. Le système développé pour Henkel constitue une base solide pour des améliorations futures et ouvre des perspectives intéressantes pour l'optimisation continue des processus de production. Les futures recherches pourraient se concentrer sur l'intégration de nouvelles techniques d'intelligence artificielle et l'amélioration continue des modèles pour répondre aux défis évolutifs de l'industrie. L'implémentation de ce système chez Henkel permettra non seulement de renforcer la qualité des produits mais aussi de réaliser des économies substantielles et d'augmenter la satisfaction des clients, contribuant ainsi à l'atteinte des objectifs de l'entreprise en matière de zéro perte et de qualité optimale.

Bibliographie

- [1] FasterCapital. The impact of defects and returns on customer satisfaction, 2024. Accessed : 2024-07-05.
- [2] EasyODM. Machine vision and industry 4.0, 2023. Accessed : 2024-07-05.
- [3] Lattepanda. Blog, 2024. Accessed : 2024-07-05.
- [4] Philippe Jean-Baptiste. Ia : La régression, 2024. Accessed : 2024-07-05.
- [5] Towards Data Science. Machine learning for beginners, 2024. Accessed : 2024-07-05.
- [6] Brightcape. Le clustering : définition et implémentations, 2024. Accessed : 2024-07-05.
- [7] Wikipedia. Règle d'association, 2024. Accessed : 2024-07-05.
- [8] Scribbr. Reinforcement learning, 2024. Accessed : 2024-07-05.
- [9] Datacamp. What is a confusion matrix in machine learning?, 2024. Accessed : 2024-07-05.
- [10] El Houcine. Rapport master el houcine - version finale, 2020. Accessed : 2024-07-05.
- [11] Inside Machine Learning. Cnn : Couche de convolution, 2024. Accessed : 2024-07-05.
- [12] Imaios. Classification des images médicales : comprendre le réseau de neurones convolutifs (cnn), 2024. Accessed : 2024-07-05.
- [13] Biorxiv. Article title. 2024. Accessed : 2024-07-05.
- [14] Data Science Today. Les réseaux de neurones convolutifs, 2024. Accessed : 2024-07-05.
- [15] Imaios. Classification des images médicales : comprendre le réseau de neurones convolutifs (cnn), 2024. Accessed : 2024-07-05.
- [16] Datascientest. Transfer learning, 2024. Accessed : 2024-07-05.
- [17] Na Liu, Lihong Wan, Qiao Huang, and Yunfeng Ji. Multi-view deep representations with cross-dataset transfer for remote sensing image retrieval and classification. *Multimedia Tools and Applications*, 80, 06 2021.
- [18] Henkel. Henkel's global supply chain, 2024.
- [19] Henkel. Henkel's supplier management, 2024.
- [20] Henkel. Henkel's manufacturing facilities, 2024.
- [21] Henkel. Continuous improvement initiatives at henkel, 2024.
- [22] Henkel. Advanced planning and optimization tools, 2024.
- [23] Henkel. Sustainability at henkel, 2024.
- [24] Nom de l'auteur. fr_ohp_casestudytemplate_packaging failures, 2024. PDF Document.
- [25] Pouria Aryan, Santhakumar Sampath, and Hoon Sohn. An overview of non-destructive testing methods for integrated circuit packaging inspection. *Sensors*, 18(7) :1981, 2018. Accessed : 2024-07-05.
- [26] Daniel Chauveau. Nouvelles possibilités offertes par la technique fmc/tfm embarquée dans un équipement multiéléments portable pour l'aide à la caractérisation des défauts de soudures. *Institut de Soudure Industrie*, 2016. Article.
- [27] J. Poguet, L. Chupin, P. Gendreu, S. Mahaut, and O. Roy. Contrôle ultrasonore au moyen de traducteurs multi-éléments : design et applications. *The e-Journal of Nondestructive Testing & Ultrasonics*, 6(8), 2001.

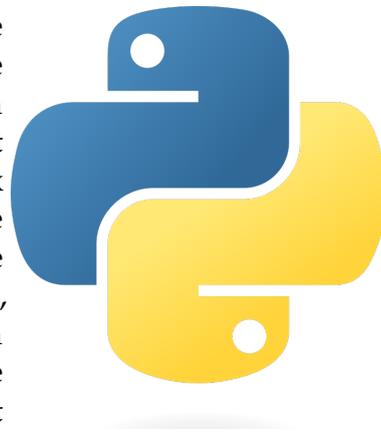
-
- [28] Nicolas Vandenbroucke. Système de vision industrielle, March 10 2015. TECHNIQUES DE L'INGÉNIEUR.
- [29] Abdelmalik Taleb-Ahmed and Luc Duvieubourg. Cours de vision industrielle, 2005. Laboratoire d'Analyse des Systèmes du Littoral, Université du Littoral Côte d'Opale, Calais Codex.
- [30] Richard Szeliski. *Computer Vision : Algorithms and Applications*. Springer, 2010.
- [31] Ramesh Jain, Rangachar Kasturi, and Brian G. Schunck. *MACHINE VISION*. McGraw-Hill, 1995.
- [32] Stuart J. Russell and Peter Norvig. *Artificial Intelligence : A Modern Approach, 3rd ed.* Pearson, 2011.
- [33] T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
- [34] Michael I. Jordan and Tom M. Mitchell. Machine learning : Trends, perspectives, and prospects. *Science*, 349, 2015.
- [35] et al. Markus Löning. sktime : A unified interface for machine learning with time series, 2019. arXiv :1909.07872.
- [36] Sotiris B. Kotsiantis. Supervised machine learning : A review of classification techniques. *Informatica*, 31 :249–268, 2007.
- [37] Olivier Chapelle, Bernhard Scholkopf, and Alexander Zien. *Semi-Supervised Learning*. MIT Press, 2006.
- [38] Colin Raffel Ekin D. Cubuk Ian Goodfellow Alex Oliver, Augustus Odena. Realistic evaluation of deep semi-supervised learning algorithms. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.
- [39] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning : An Introduction, 2nd edition*. MIT Press, 2018.
- [40] David Silver et al. Volodymyr Mnih, Koray Kavukcuoglu. Human-level control through deep reinforcement learning. *Nature*, 518 :529–533, 2015.
- [41] Mohamed Lamine Derouaz. Conception et réalisation d'un système de traitement des données issues des médias sociaux, 2020.
- [42] Ghazal M. Al et al. Cloud-based monitoring of thermal anomalies in industrial environments using ai and the internet of robotic things. *Sensors (Basel)*, 20, 2020.
- [43] Tamoghna Ghosh, Raghav Bali, and Dipanjan Sarkar. *Hands On Transfer Learning with Python*. 2018.
- [44] Grega Vrbančič and Vili Podgorelec. Transfer learning with adaptive fine-tuning. *IEEE Access*, 8 :196197–196211, 2020.
- [45] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once : Unified, real-time object detection. *arXiv preprint*, 2015.
- [46] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn : Towards real-time object detection with region proposal networks. In *Conference on Neural Information Processing Systems (NIPS)*, 2015.
- [47] Al et al. Ghazal M. Comparison analysis of yolo, resnet, and efficientnet for contagious skin diseases. *Sensors (Basel)*, 20(21), 2020.

Annexe

Présentation des outils utilisés

Python

est un langage de programmation interprété, orienté objet, de haut niveau et doté d'une sémantique dynamique. Ses structures de données intégrées de haut niveau, combinées au typage dynamique et à la liaison dynamique, le rendent très attrayant et polyvalent qui peut être utilisé dans de nombreux domaines, tels que le développement Web, l'analyse de données, l'apprentissage profond et de l'intelligence artificielle. Python supporte les modules et les packages, ce qui encourage la modularité des programmes et la réutilisation du code. L'interpréteur Python et sa vaste bibliothèque standard sont disponibles gratuitement pour toutes les principales plates-formes et peuvent être distribués librement.



Numpy

qui signifie "Numerical Python", est une bibliothèque open-source pour le langage de programmation Python. Elle fournit un support pour des tableaux multidimensionnels, des matrices de grande taille, ainsi que pour un large éventail de fonctions mathématiques de haut niveau pour opérer sur ces tableaux. Essentiellement utilisée pour la manipulation numérique des données. Elle offre des opérations de calcul de base telles que l'addition, la soustraction, la multiplication et la division, ainsi que des opérations plus complexes comme les transformations de Fourier, l'algèbre linéaire, et la génération de nombres aléatoires. Cette bibliothèque est à la base de nombreuses autres bibliothèques de data science et d'apprentissage automatique, comme Pandas pour la manipulation des données et Scikit-Learn pour l'apprentissage automatique. NumPy est souvent utilisée dans les domaines de l'apprentissage automatique, de la science des données, de l'analyse des données, et de la visualisation des données.



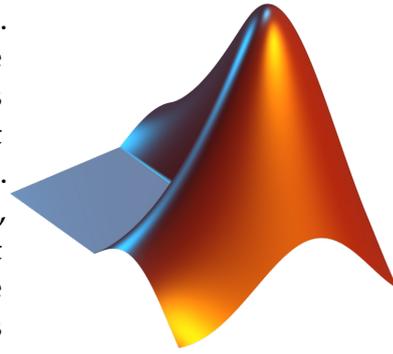
Scikit-learn

est une bibliothèque open source pour le langage de programmation Python qui fournit une gamme d'outils d'apprentissage automatique supervisé et non supervisé. Elle est construite sur deux autres bibliothèques Python, NumPy et SciPy, et s'intègre bien avec d'autres bibliothèques de la pile scientifique Python, comme Pandas et Matplotlib. Scikit-learn offre une interface uniforme pour de nombreux algorithmes d'apprentissage automatique, ce qui facilite leur utilisation et leur comparaison. Les types d'algorithmes incluent la classification, la régression, le clustering, la réduction de dimensionnalité, l'estimation de densité, et bien d'autres. De plus, Scikit-learn comprend des outils pour le prétraitement des données, la sélection et l'évaluation des modèles, et le tuning des hyperparamètres. La bibliothèque est largement utilisée en science des données et en apprentissage automatique, en raison de sa flexibilité, de sa facilité d'utilisation, et du fait qu'elle est soutenue par une vaste communauté de développeurs et d'utilisateurs.



MATLAB

est un environnement de calcul numérique et un langage de programmation développé par MathWorks. MATLAB permet de manipuler des matrices, de visualiser des données, de mettre en œuvre des algorithmes, de créer des interfaces utilisateur et d'interfacer avec d'autres langages de programmation. Il est largement utilisé dans les domaines de l'ingénierie, des sciences et de l'économie pour l'analyse et la simulation des systèmes dynamiques. La vaste bibliothèque de fonctions de MATLAB et ses outils de visualisation intégrés facilitent le traitement des données et le développement rapide de solutions.

**TensorFlow**

est une bibliothèque logicielle open-source développée par Google pour le calcul numérique et l'apprentissage automatique. Elle est particulièrement adaptée au développement et à l'entraînement de modèles d'apprentissage profond, mais peut également être utilisée pour d'autres types d'apprentissage automatique. TensorFlow permet de déployer facilement des modèles sur divers systèmes, allant des appareils mobiles aux clusters de serveurs. Grâce à son architecture flexible, TensorFlow supporte les calculs sur CPU et GPU, ce qui accélère considérablement le traitement des données et l'entraînement des modèles complexes.

**Roboflow**

est une plateforme complète pour la gestion des données d'entraînement, l'annotation et le déploiement de modèles de vision par ordinateur. Roboflow simplifie le flux de travail des développeurs en fournissant des outils pour annoter des images, générer des ensembles de données augmentés et entraîner des modèles de détection d'objets, de segmentation et de classification. Grâce à ses intégrations avec diverses bibliothèques d'apprentissage automatique et ses capacités de déploiement facile, Roboflow permet de créer des applications de vision par ordinateur robustes et performantes. La plateforme supporte également le partage et la collaboration sur les ensembles de données, ce qui facilite le travail en équipe et l'amélioration continue des modèles.



SPINVIEWER

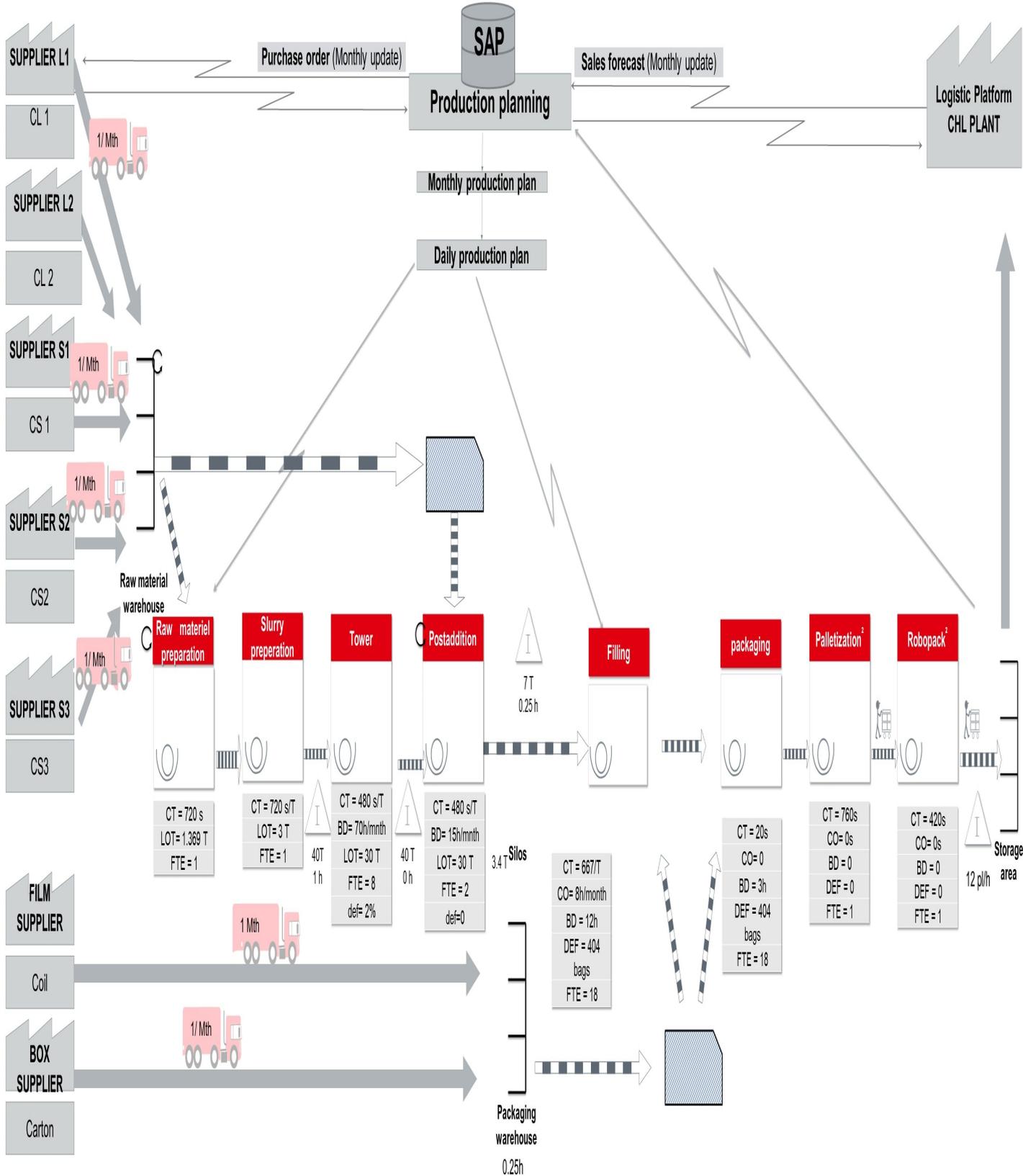
L'architecture Spinnaker est dotée de nouvelles fonctionnalités qui maintiennent zéro erreur à grande vitesse pour la fiabilité GigE maintenant et à l'avenir. Un test de stress de trois jours utilisant un système de 40 caméras Blackfly S 1GigE à 90



DAHENG GALAXY VIEWER

Le programme Daheng Galaxy Viewer est l'outil principal pour tester et programmer les caméras. Avec Galaxy Viewer, il est possible de modifier et d'enregistrer les paramètres de la caméra ainsi que de capturer et de stocker des images. Il dispose de plusieurs plugins pour des fonctionnalités supplémentaires telles que le réticule, la grille et la LUT. Nous allons vous montrer les possibilités et comment utiliser le Daheng Galaxy Viewer.

Value Stream Worksheet	Site: CHL Plant	Area of investigation: U400 & Filling Area 320grs	Team: HPS
	Area: Production	Product: ISIS Multi usage	Date: 2023



Legend: CT = cycle time; LOT = Lot Size; OEE = Overall Equipment Effectiveness; FTE = Full Time Equivalents ; ; crf = Carton; pl = Palette; Mth = Month; Wk = Week; D = Day. 1crf= 25 cons



Manual information flow

Electronic information flow

Note: 1. Based on 5 filling lines running; 2. Shared resource between 5 lines.

