

Développement et Optimisation de modèles de calibration pour distorsion fisheye

Rapport de stage

soutenu le

par

OULAHCENE Anissa

Encadrant universitaire : Dr.CHETOUANI Aladine

Encadrant entreprise : Dr.GARBOUGE Hadhami Et Dr.KAHOUADJI Mouad

Caplogy Innovation



Remerciements

En premier lieu, je tiens à remercier mes responsables de stage , Un grand merci pour leur accueil au sein de l'entreprise, ainsi que pour leur confiance lors de cette aventure dans le monde professionnel et leurs précieux conseils. Je les remercie aussi pour leur disponibilité et la qualité de leur encadrement en entreprise.je tiens a remercier yedir .

Je remercie également toute l'équipe R&D à *Caplogy* pour leur accueil, leur esprit d'équipe et leurs conseils qui m'ont aidé à réussir ce travail.

Je souhaite aussi remercier particulièrement mes responsables du Master Université Sorbonne Paris-Nord, qui m'ont fourni les outils théoriques et pratiques nécessaires au bon déroulement de ce stage .

J'adresse mes profonds remerciements au personnel du Master *en Ingénierie et Innovation en Réseaux et image parcours image de l'Université de Sorbonne Paris-Nord* pour leurs soutiens et la qualité de l'enseignement offert.

Enfin, j'exprime mes vifs remerciements aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir avec leurs remarques.

Ce stage m'a permis d'affiner certaines pistes pour bâtir mon projet d'orientation professionnelle et signer l'aboutissement de mon diplôme universitaire.

Dédicaces

*Je dédie ce travail à
Mon père et ma mère,
Mon frère et ma soeur,
Toute ma famille,
Ainsi qu'à mes amis et collègues,
Qui m'ont soutenu et encouragé tout au long de mon stage.*

Table des matières

Table des figures	v
Liste des tableaux	vii
Liste des acronymes	viii
Glossaire	ix
Résumé	1
Abstract	2
Présentation de l'entreprise	3
Introduction générale	4
1 Introduction	4
2 Contexte général	4
3 Motivation et Objectifs	5
4 Organisation du rapport	5
État de l'art	7
1 Introduction	7
2 Présentation des Caméras Fisheye	7
3 Modèles spécifiques aux caméras fisheye	8
4 Distorsions dans les caméras fisheye	9
5 Définition et objectifs de la calibration	11
5.1 Calibration avec cible	11
6 Calibration sans cible	12
7 Travaux connexes	12

8	Conclusion	16
Contribution		17
1	Introduction	17
2	Collecte et Préparation du Dataset Caplogy	18
2.1	Création d'un dataset Privé Caplogy	18
3	Approche classique de calibration	20
3.1	Pipeline de calibration avec la toolbox OCamCalib	20
3.2	Processus de calibration selon Scaramuzza	20
3.3	Redressement d'image fisheye basée sur le modèle de calibration	21
3.4	Méthodes de rectification pour un champ de vision maximal	21
4	Approche basée sur le deep learning	23
4.1	Préparation et traitement des données d'entraînement	23
4.2	Modèle de rectification basé sur U-Net	24
4.3	Modèle de rectification basé sur un GAN conditionnel (cGAN)	26
5	Conclusion	29
Expérimentations et résultats		30
1	Introduction	30
2	Résultats Expérimentations-Méthode Classique	30
2.1	Évaluation de la méthode classique (OCamCalib)	30
2.2	Précision de la calibration	34
2.3	Résultats visuels de redressement	35
2.4	Méthodes de rectification (FOV)	36
3	Expérimentations et Résultats – Modèles Deep Learning	37
3.1	Modèle 1 : Modèle de rectification basé sur U-Net	37
3.2	Modèle 2 : Modèle de rectification basé sur un GAN conditionnel	42
3.3	Analyse Comparative des méthodes de calibration	49
4	Conclusion	50
Conclusion générale		51
Bibliographie		54

Table des figures

1	Comparaison entre une caméra standard et une caméra fisheye	8
2	Types de distorsions.	11
1	Echantillons Dataset de calibration	18
2	Echantillons Dataset d'inférence	19
3	Architecture du modèle U-Net pour la rectification d'images fisheye	25
4	Architecture du modèle GAN pour la rectification d'images fisheye	27
1	Extraction automatique des coins du damier.	31
2	Estimation initiale des paramètres.	31
3	Superposition des projections sur les coins détectés.	32
4	Analyse des erreurs de reprojection par image.	33
5	Nuage de points Erreurs.	33
6	Comparaison du champ latéral et de la profondeur.	34
7	Comparaison quantitative entre le champ de vision calibré et le champ perdu.	35
8	Comparaison entre une image fisheye originale et l'image calibré	35
9	Comparaison entre l'image fisheye originale l'image rectifiée par méthode classique et l'image après la méthode de rectification.	36
10	Comparaison visuelle entre image fisheye, correction U-Net et vérité terrain.	38
11	Évolution du Peak Signal-to-Noise Ratio durant l'entraînement.	39
12	Évolution du Structural Similarity Index durant l'entraînement.	39
13	Évolution de l'erreur absolue moyenne durant l'entraînement.	40
14	Évolution des losses d'entraînement et de validation.	41
15	Comparaison visuelle entre image fisheye, correction U-Net et vérité terrain.	43
16	Évolution du Peak Signal-to-Noise Ratio durant l'entraînement.	44
17	Évolution du Structural Similarity Index durant l'entraînement.	45
18	Évolution de l'erreur absolue moyenne durant l'entraînement.	46

Liste des tableaux

1	Comparaison des principaux modèles de projection pour caméras fisheye	9
2	Analyse comparative des approches de calibration pour caméras fisheye	15
3	Résumé des deux datasets privés Caplogy	19
4	Comparatif des méthodes de rectification pour un champ de vision maximal	21
5	Ensembles de données réelles	24
6	Ensemble de données synthétique	24
7	Résumé de l'architecture du modèle U-Net utilisé	25
8	Résumé des métriques utilisées pour l'évaluation – U-Net	26
9	Résumé de l'architecture GAN pour la calibration fisheye	27
10	Résumé des métriques utilisées pour l'évaluation	28
1	Résumé de l'environnement d'entraînement – U-Net	37
2	Analyse Globale des Performances du modèle	42
3	Résumé de l'environnement d'entraînement	43
4	Analyse synthétique des performances du modèle GAN	48
5	Comparaison des Méthodes de Correction de Distorsion Fisheye	49

Liste des acronymes

- AI** Artificial Intelligence (Intelligence Artificielle)
- BA** Bundle Adjustment (Ajustement global des paramètres de calibration)
- BDD** Base de Données
- CNN** Convolutional Neural Network (Réseau de neurones convolutif)
- CV** Computer Vision (Vision par ordinateur)
- DL** Deep Learning (Apprentissage profond)
- FOV** Field Of View (Champ de vision)
- FORTRAJ** Forecasting Trajectories (Modèle de prédiction de trajectoires)
- GAN** Generative Adversarial Network (Réseau antagoniste génératif)
- GT** Ground Truth (Vérité terrain)
- IoU** Intersection over Union (Métrique de recouvrement)
- ML** Machine Learning (Apprentissage automatique)
- MSE** Mean Squared Error (Erreur quadratique moyenne)
- OCamCalib** Omnidirectional Camera Calibration Toolbox
- OpenCV** Open Source Computer Vision Library
- RGB** Red Green Blue (Espace de couleur)
- ROI** Region Of Interest (Région d'intérêt)
- SVD** Singular Value Decomposition (Décomposition en valeurs singulières)
- UAV** Unmanned Aerial Vehicle (Drone autonome)
- U-Net** U-shaped Convolutional Network

Glossaire

- Calibration** Processus visant à déterminer les paramètres intrinsèques et extrinsèques d'une caméra pour corriger les distorsions optiques.
- Dataset** Ensemble de données utilisées pour entraîner les modèles, composées ici d'images fisheye brutes et de leurs versions rectifiées.
- U-Net** Architecture CNN en encodeur-décodeur avec connexions de saut.
- Encodeur** Partie du réseau qui extrait des caractéristiques profondes de l'image.
- ResNet34** Modèle CNN préentraîné sur ImageNet utilisé comme encodeur.
- Décodeur** Partie du réseau qui reconstruit l'image à partir des caractéristiques extraites.
- Skip connections** Connexions entre couches de l'encodeur et du décodeur pour garder les détails fins.
- Activation ReLU** Fonction d'activation non linéaire des couches intermédiaires.
- GAN conditionnel (cGAN)** GAN où le générateur est conditionné par une image.
- Pix2Pix** Modèle cGAN populaire pour les transformations image-à-image.
- Discriminateur PatchGAN** Discriminateur basé sur des patches pour juger du réalisme local.
- Réseau de calibration CNN** Réseau convolutionnel pour régresser les paramètres de calibration.
- Optimiseur Adam** Algorithme d'optimisation avec taux d'apprentissage adaptatif.
- Early stopping** Arrêt anticipé de l'entraînement pour éviter le surapprentissage.
- Fonction de perte BCEWithLogitsLoss** Perte binaire combinant sigmoid et entropie croisée.
- L1 Loss** Perte basée sur la somme des différences absolues.

Résumé

Ce stage s'inscrit dans le cadre d'un projet visant à renforcer les capacités d'un drone autonome évoluant dans des environnements complexes, notamment pour des tâches de prédiction de trajectoires. Les caméras fisheye, couramment utilisées dans les systèmes de vision embarquée en raison de leur large champ de vision hémisphérique, induisent cependant d'importantes distorsions optiques. Ces déformations altèrent les mesures géométriques et compromettent la précision des algorithmes de vision par ordinateur, en particulier ceux dédiés à l'analyse du mouvement et à la prédiction de trajectoires.

Le travail mené dans ce projet consiste à corriger ces distorsions tout en préservant un maximum d'informations utiles. Pour ce faire, une méthode de calibration spécifique a été développée, mise en œuvre et évaluée expérimentalement. Cette approche permet de modéliser avec précision les déformations induites par l'optique fisheye, puis de les compenser afin d'obtenir des images rectifiées et exploitables. L'objectif principal est d'améliorer la qualité des données visuelles en vue d'optimiser les performances du modèle de prédiction embarqué sur le drone.

Les résultats obtenus démontrent que la correction des distorsions contribue significativement à l'amélioration des données d'entrée des modules de traitement, renforçant ainsi la fiabilité globale du système autonome.

Mots-clés :Caméra fisheye, calibration, rectification d'image, drone autonome,la distorsion radiale,OCamCalib Toolbox,Approche de Scaramuzza.

Abstract

This internship is part of a project aimed at enhancing the capabilities of an autonomous drone operating in complex environments, particularly for trajectory prediction tasks. Fisheye cameras, commonly used in embedded vision systems due to their wide hemispherical field of view, however introduce significant optical distortions. These deformations affect geometric measurements and compromise the accuracy of computer vision algorithms, particularly those dedicated to motion analysis and trajectory prediction.

The work carried out in this project focuses on correcting these distortions while preserving as much useful information as possible. To achieve this, a specific calibration method was developed, implemented, and experimentally evaluated. This approach allows for precise modeling of the distortions induced by the fisheye lens, and compensates for them to produce rectified and usable images. The main objective is to improve the quality of visual data in order to optimize the performance of the prediction model embedded on the drone.

The results obtained demonstrate that distortion correction significantly contributes to improving the input data for processing modules, thus enhancing the overall reliability of the autonomous system.

Keywords :Fisheye camera, calibration, image rectification, autonomous drone, radial distortion, OCamCalib Toolbox, Scaramuzza's approach.

Présentation de l'entreprise

Caplogy est un acteur en transformation digitale et en prestations dans le secteur de l'éducation et de l'enseignement supérieur. L'entreprise accompagne ses collaborateurs en France et à l'international dans les domaines Digital, IoT, Data/IA, Cybersécurité, DevOps/Cloud. Outre que les prestations d'enseignement, Caplogy propose un pôle R&D dans le but de développer en interne plusieurs projets d'innovation lui permettant de valoriser les compétences de ses collaborateurs, de développer leurs savoir-faire ainsi que d'être à la pointe de l'innovation dans des domaines cibles.

Les activités de recherche de l'équipe R&D de Caplogy sont multidisciplinaires et visent principalement à comprendre le comportement des personnes ; à traiter et analyser les données collectées de leur vie privée (par exemple, la localisation des individus, les signes vitaux des patients) ; à sécuriser ces données et à optimiser leur partage. Les données privées sont de nature spatio-temporelle et peuvent conduire à la divulgation d'une grande variété d'informations sensibles en matière de vie privée concernant les individus, telles que les routines et les habitudes, l'état de santé, les affiliations politiques/religieuses, les préférences et les activités effectuées.

Dans ce cadre, l'équipe R&D développe de nouvelles approches permettant d'améliorer la modélisation des données de la vie privée, optimiser le partage et le stockage d'une manière sécurisée et de prédire les cyberattaques d'une manière rapide et autonome. Les applications possibles incluent : la snaté, les villes intelligentes, les maisons intelligentes, l'e-commerce et la finance.

Des projets de recherche divers sont en cours de développement dans le pôle R&D de Caplogy dont PICO-DEV, KenetIQ, NAUTILUS et ICAPS auquel ce stage s'inscrit.

Introduction générale

1 Introduction

L'usage grandissant des drones autonomes dans des cadres complexes, comme les zones urbaines, les sites industriels ou les espaces réduits, pose d'importants défis en termes de détection et d'interprétation de la scène. Les caméras fisheye, qui font partie des capteurs intégrés, se démarquent par leur faculté à saisir un très grand champ de vision, souvent proche de l'hémisphère. De ce fait, elles sont privilégiées pour une utilisation en vision embarquée. Cependant, cette particularité optique cause des distorsions d'image majeures qui influencent l'exactitude des algorithmes de traitement visuel, notamment ceux associés à la détection de mouvement ou à la prédiction de trajectoires.

Ce stage, effectué dans le contexte d'un projet de recherche appliquée, a pour but majeur d'améliorer la qualité des images prises par une caméra fisheye installée sur un drone autonome. Cela se fait en élaborant une technique de calibration sophistiquée visant à rectifier ces distorsions. Ce projet a pour but d'améliorer les performances d'un modèle de prévision de trajectoires(tache complémentaire), en assurant une plus grande cohérence géométrique des images employées.

Cette introduction se développe autour de trois points essentiels : la motivation et les buts du stage, le contexte dans lequel il s'intègre, et pour finir, une présentation succincte de la structure du rapport.

2 Contexte général

Les systèmes de vision intégrée sont indispensables aux robots autonomes, en particulier aux drones, pour des tâches telles que le suivi de trajectoires, la navigation, l'identification d'obstacles, l'analyse de scènes. Les caméras fisheye sont privilégiées grâce à leur large champ de vision hémisphérique, mais elles introduisent des distorsions géométriques

qui compliquent l'utilisation directe des images. Ces déformations affectent la fiabilité des algorithmes de vision artificielle, notamment pour l'analyse du mouvement et la prévision de trajectoires. La calibration apparaît donc comme une étape essentielle pour corriger ces distorsions et fournir des données visuelles fiables, contribuant ainsi à améliorer les performances des systèmes autonomes en conditions réelles.

3 Motivation et Objectifs

La perception visuelle intégrée est essentielle pour les drones autonomes, notamment pour la navigation et le suivi de trajectoires. Les caméras fisheye offrent un large champ de vision mais génèrent des distorsions optiques importantes, rendant les mesures géométriques moins fiables.

Dans ce cadre, l'objectif principal de ce stage est de rectifier les déformations provoquées par les caméras fisheye pour optimiser la qualité des images employées dans le processus prédictif et fournir des données visuelles corrigées, utilisables par des algorithmes avancés de vision par ordinateur.

Dans le contexte de ce stage, les objectifs visés sont les suivants :

- Élaboration d'une technique de calibration avancée, conçue spécifiquement pour les caméras fisheye, permettant de modéliser et rectifier les distorsions avec précision.
- Évaluation expérimentale en comparaison avec des techniques existantes, pour juger les performances de la méthode proposée en matière de précision et de robustesse.

4 Organisation du rapport

Ce rapport est structuré en quatre sections principales, chacune détaillant les différentes avancées et réalisations du stage.

Introduction Générale : Ce chapitre présente le contexte dans lequel s'inscrit le projet, à savoir le développement d'un système de suivi de personnes par drone autonome équipé d'une caméra fisheye. Il expose la problématique centrale liée aux distorsions optiques importantes de ce type de capteur, ainsi que les objectifs visés par le stage.

État de l'Art : Ce chapitre offre un aperçu des techniques de calibration pour les caméras fisheye. Il détaille les méthodes classique de calibration, y compris les modèles géométriques (Scaramuzza) et polynomiaux, en mettant en évidence leurs contraintes face à de grandes déformations et aux conditions expérimentales rigoureuses. Il souligne aussi les opportunités offertes par les méthodes d'apprentissage profond, avant de présenter un examen approfondi du modèle de Scaramuzza, employé en tant que référence théorique.

Modélisation : Ce chapitre présente les approches de modélisation proposées dans le cadre du stage pour répondre à la problématique de calibration de caméras fisheye : une approche géométrique classique, et une approche moderne par Deep Learning visant à estimer automatiquement les paramètres de calibration à partir d'images fisheye. La première repose sur des techniques bien établies, efficaces en conditions contrôlées, mais limitées dans les environnements dynamiques ou non préparés. La seconde exploite des architectures de réseaux de neurones, offrant davantage de souplesse et de potentiel de généralisation. Parallèlement à ces travaux, un dataset privé a été constitué en interne chez Caplogy.

Expérimentations et Résultats : Dans ce chapitre, les différentes expérimentations menées sont décrites en détail. Les performances des modèles sont analysées et discutées, mettant en lumière les points forts et les limitations de chaque approche.

Conclusion et Perspectives Ce chapitre se termine par une conclusion générale et des perspectives pour des travaux futurs. Il met lumière les résultats les plus significatifs et en suggérant des pistes pour approfondir les recherches entreprises durant le stage.

État de l'art

1 Introduction

La précision de l'étalonnage est déterminante pour les systèmes de vision par ordinateur, en particulier avec les caméras fisheye qui, malgré leur large champ de vision, introduisent de fortes distorsions non linéaires. Les méthodes classiques de calibration se révèlent alors insuffisantes, d'où la nécessité de modèles adaptés. Ce chapitre présente les fondements de la calibration des caméras fisheye, l'analyse des modèles géométriques et un état de l'art des approches existantes, avec un focus sur la méthode de Scaramuzza (OCamCalib), qui constitue la référence de cette étude.

2 Présentation des Caméras Fisheye

Les caméras fisheye sont des équipements optiques spécifiquement élaborés pour enregistrer des images sur une très large étendue, pouvant atteindre jusqu'à 180°, voire 220°. À l'opposé des caméras traditionnelles à projection perspective qui reproduisent la scène de façon linéaire, les caméras fisheye exploitent une projection non linéaire afin d'optimiser le champ visuel dans une seule image. Cela rend cet outil particulièrement approprié pour les applications comme la robotique mobile, la surveillance à 360 degrés ou la réalité virtuelle.

Dans une caméra classique, l'angle de vision est souvent restreint à une fourchette de 60° à 90°, et les lignes de la scène demeurent droites dans l'image. Cependant, l'utilisation d'une caméra fisheye permet de capturer une vue hémisphérique, ce qui entraîne une distorsion considérable : les lignes droites de la réalité apparaissent courbées sur l'image.

Cette distorsion n'est pas un défaut, mais une conséquence délibérée de la

conception optique, autorisant l'appareil photo à courber l'espace visuel pour saisir une plus grande partie de la scène. Toutefois, cela nécessite une rectification de ces distorsions et tirer le meilleur parti des données contenues dans les images enregistrées.

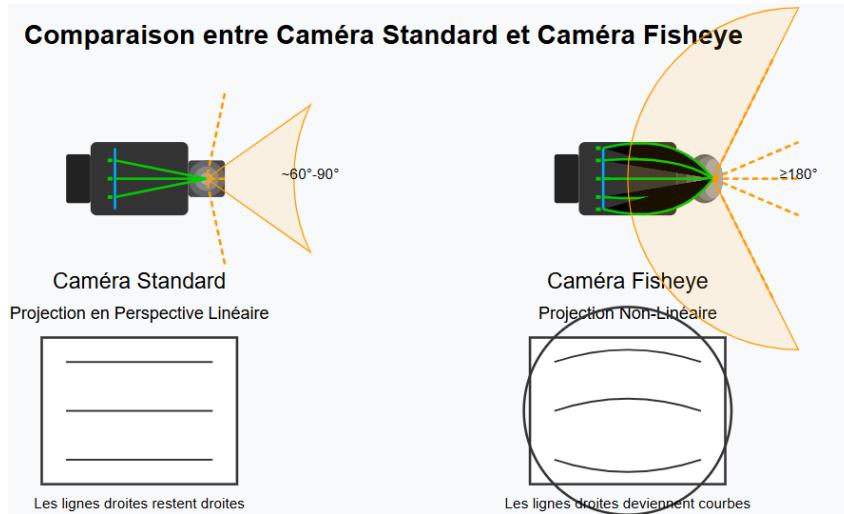


FIGURE 1 – Comparaison entre une caméra standard et une caméra fisheye

La caméra standard utilise une projection perspective linéaire avec un champ de vision restreint, tandis que la caméra fisheye adopte une projection non linéaire pour couvrir un champ beaucoup plus large, au prix d'une distorsion des lignes droites.

3 Modèles spécifiques aux caméras fisheye

Les caméras fisheye utilisent des projections non linéaires pour maximiser le champ de vision. Voici les principales projections utilisées :

3.0.1 Projection équidistante

Formule :

$$r = f \cdot \theta$$

Principe : L'écart entre le centre de l'image et le point projeté est en relation proportionnelle avec l'angle d'incidence.

Usage : Elle offre une représentation angulaire uniforme, ce qui la rend pertinente pour des domaines tels que la robotique mobile et la vision panoramique.

3.0.2 Projection stéréographique

Formule :

$$r = 2f \cdot \tan\left(\frac{\theta}{2}\right)$$

Principe : C'est une projection conforme : elle préserve les angles, ce qui est crucial pour certaines applications, bien qu'elle ne maintienne ni les longueurs ni les surfaces.

Usage : Employée quand la conservation locale de la forme des objets revêt une grande importance.

3.0.3 Projection orthographique

Formule :

$$r = f \cdot \sin(\theta)$$

Principe : L'endroit où l'image est projetée correspond à la projection orthogonale du faisceau incident sur le plan de l'image.

Effet : Génère une image visuellement aplatie ; peu couramment utilisée seule, mais parfois employée comme approximation dans certains modèles.

Modèle de projection	Formule	Avantages	Limites d'utilisation
Perspective (pinhole)	$r = f \cdot \tan(\theta)$	Simple et répandu	$\text{FOV} \leq 90^\circ$
Équidistante	$r = f \cdot \theta$	Facile à modéliser	FOV jusqu'à 180°
Stéréographique	$r = 2f \cdot \tan\left(\frac{\theta}{2}\right)$	Préserve les angles (conforme)	Idéal pour vision panoramique
Orthographique	$r = f \cdot \sin(\theta)$	Bonne approx. au centre	FOV modéré recommandé

TABLE 1 – Comparaison des principaux modèles de projection pour caméras fisheye

4 Distorsions dans les caméras fisheye

Les objectifs fisheye proposent un angle de vue étendu, cependant, ils engendrent d'importantes distorsions optiques en raison de leurs lentilles. Ces dis-

torsions modifient la géométrie des images, exigeant l'utilisation de modèles de rectification.

4.0.1 Distorsion radiale

La distorsion radiale, qui est dominante en fisheye, entraîne un déplacement grandissant des points lorsqu'ils s'éloignent de l'axe optique, ce qui produit une impression de gonflement ou de compression circulaire. On a tendance à la modéliser via une série polynomiale en fonction de la distance radiale r :

$$r_{\text{corrigé}} = r(1 + k_1r^2 + k_2r^4 + k_3r^6 + \dots)$$

où k_1, k_2, k_3 :les coefficients de distorsion radiale à estimer lors de la calibration.

Il est crucial de procéder à ce genre de correction pour assurer une représentation fidèle de la géométrie de la scène.

4.0.2 Distorsion tangentielle

L'asymétrie des lignes droites, particulièrement en périphérie de l'image, est le résultat d'un décalage entre la lentille et le capteur qui provoque une déviation tangentielle des points.

Elle est souvent modélisée par les expressions suivantes :

$$\begin{cases} x_{\text{corr}} = x + [2p_1xy + p_2(r^2 + 2x^2)] \\ y_{\text{corr}} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \end{cases}$$

où p_1 et p_2 :les coefficients de distorsion tangentielle $r = \sqrt{x^2 + y^2}$: la distance au centre optique; habituellement insignifiant en fisheye, peut devenir significatif si des défauts mécaniques ou d'assemblage optique se présentent.

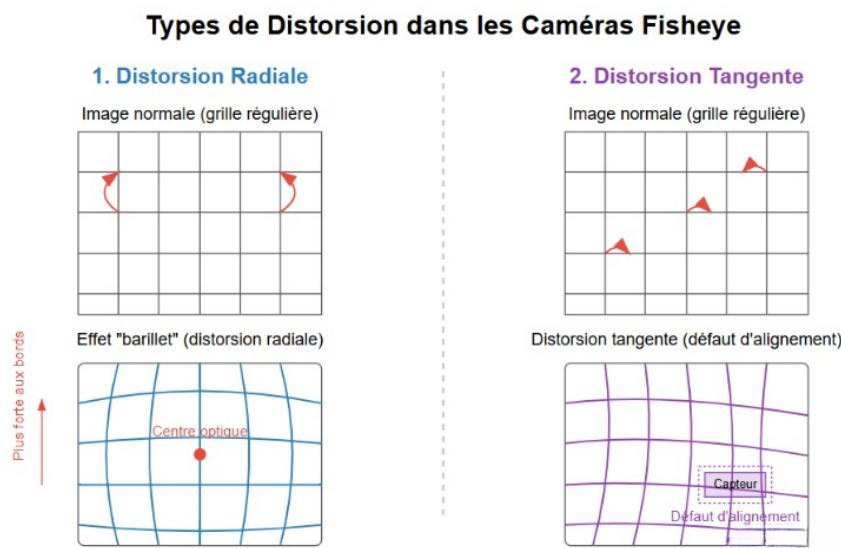


FIGURE 2 – Types de distorsions.

5 Définition et objectifs de la calibration

L’objectif de la calibration d’une caméra est de déterminer ses paramètres intrinsèques (propriétés internes) et extrinsèques (position et orientation), ce qui permet de faire le lien entre les points 3D du monde réel et leur représentation en 2D. Pour les caméras à objectif fisheye, elle est indispensable pour :

- rectifier les importantes déformations géométriques.
- faciliter une reconstruction en trois dimensions.
- Optimiser l’efficacité des algorithmes de vision (detection, suivi, etc.).

5.1 Calibration avec cible

La calibration avec cible implique l’utilisation d’un motif connu, comme un damier, examiné sous divers angles pour faire correspondre les points 2D identifiés sur l’image à leurs coordonnées 3D réelles.

5.1.1 Méthode de Zhang

La technique de Zhang (2000) est une référence établie pour la calibration des caméras en perspective. Elle se base sur diverses perspectives d’un damier à plat pour évaluer les paramètres intrinsèques et extrinsèques grâce à l’optimisation en

moindres carrés. Elle est simple et précise pour les caméras traditionnelles, mais reste inappropriée pour les objectifs à distorsion élevée tels que les fisheye.

5.1.2 Méthode de Scaramuzza (*OCamCalib*)

La technique de Scaramuzza et al. (2006) est élaborée pour les caméras omnidirectionnelles et fisheye. Elle représente la projection à l'aide d'un polynôme ajusté basé sur les images d'un damier, garantissant une calibration précise même pour des angles de vue supérieurs à 180°. *OCamCalib* Toolbox simplifie cette calibration dans MATLAB.

6 Calibration sans cible

Avec l'avènement de l'apprentissage profond, de nouvelles approches de calibration dites "sans cible" ont émergé, s'affranchissant de la nécessité d'utiliser des motifs géométriques imprimés.

Les modèles de deep learning pour la calibration reposent généralement sur des architectures convolutives entraînées sur des jeux de données fisheye. L'idée est d'apprendre une fonction de régression entre les caractéristiques visuelles de l'image (textures, lignes, symétries) et les paramètres intrinsèques de la caméra.

Cette approche présente plusieurs avantages :

- Elle permet une calibration rapide sans besoin de scène contrôlée
- Elle est adaptée aux environnements non coopératifs

Dans ce travail, l'approche sans cible a été envisagée comme piste complémentaire, notamment pour évaluer la faisabilité d'une calibration automatisée dans des environnements dynamiques, tels que ceux rencontrés en suivi de trajectoires par drone.

7 Travaux connexes

L'étalonnage des caméras fisheye constitue un champ de recherche dynamique, compte tenu des importantes distorsions optiques qu'elles provoquent et de leur utilisation grandissante dans les domaines de la robotique, des systèmes intégrés et de la vision panoramique. Les recherches existantes englobent aussi bien les techniques traditionnelles basées sur des objectifs connus et des schémas géométriques

explicites, que les démarches récentes utilisant l'apprentissage profond pour déduire automatiquement les paramètres intrinsèques à partir d'images naturelles. Ces techniques permettent d'aborder la calibration dans des environnements contrôlés ou en conditions réelles, fournissant ainsi un panorama des solutions existantes pour les caméras fisheye.

Ces études présentent les deux paradigmes majeurs de la calibration des caméras fisheye, avec ou sans référence, et proposent un aperçu des solutions actuelles utilisables tant dans des environnements maîtrisés que dans des conditions pratiques.

Travaux connexes pour calibration de caméras fisheye

L'article [1] présente une technique d'ajustement fondée sur l'examen d'un damier sous diverses positions, associant une estimation linéaire des paramètres intrinsèques et extrinsèques à un perfectionnement non linéaire. Ce modèle, initialement conçu pour les caméras en perspective, a servi de référence pour des techniques ultérieures y compris celles destinées aux caméras fisheye. Il représente également une base théorique pour les méthodes axées sur la cible.

L'article [2] suggère un modèle intégré pour les caméras traditionnelles, grand angle et fisheye, qui emploie des fonctions radiales polynomiales pour illustrer les distorsions extrêmes, tout en garantissant une calibration géométriquement cohérente. Ce modèle est conçu pour les applications qui exigent un vaste champ de vision.

L'article [3] propose une méthode de calibration pour caméras fisheye basée sur un polynôme reliant le rayon image à l'angle d'incidence, utilisant des images de damier et un raffinement non linéaire. Cette méthode est implémentée dans **OCamCalib**, employée dans notre projet pour calibrer la caméra fisheye du drone.

L'article [4] élargit la calibration des caméras omnidirectionnelles aux situations sans cible en fusionnant cette méthode avec la reconstruction 3D par Structure-from-Motion (SfM). Ceci permet une calibration adaptable dans des contextes naturels ou non contrôlés, et est particulièrement adapté aux systèmes intégrés.

L’article [5] propose une technique de calibration automatique pour les caméras fisheye basée sur une unique image, sans l’utilisation de cible, en se servant de l’hypothèse du monde de Manhattan. Elle se sert de cartes thermiques pour identifier les directions prédominantes et en inférer les paramètres intrinsèques, grâce à un réseau de neurones profonds formé sur des scènes synthétiques. Cette méthode facilite la calibration d’images naturelles dans des contextes réels ou mobiles.

L’article [6] décrit une approche solide pour la détection automatique de damiers dans des images fisheye altérées ou indistinctes, en faisant appel à des méthodes multi-niveaux et des normes géométriques pour déterminer avec précision les points de repère et simplifier l’étape préliminaire de calibration.

L’article [7] suggère une technique de rectification d’images en fisheye fondée sur une structure à double diffusion, qui permet de rétablir des images en perspective tout en conservant les détails géométriques et l’exactitude des textures. Cela s’inscrit dans les démarches de calibration ou de correction sans cible pour des environnements concrets.

L’article [8] propose une technique de calibration automatique des caméras fisheye basée sur une seule image, en faisant appel à des réseaux de neurones convolutifs pour anticiper les orientations dominantes de la scène (Manhattan World) et ajuster les paramètres intrinsèques, sans avoir besoin d’une cible matérielle. L’article [9] introduit *FishEyeRecNet*, un réseau de neurones profonds dédié à la correction d’images fisheye, capable d’apprendre à rectifier automatiquement les distorsions en fusionnant des caractéristiques locales et globales, sans avoir besoin d’un modèle géométrique explicite ou d’une cible physique.

L’article [10] propose une technique d’ajustement des distorsions fisheye grâce à l’apprentissage profond, basée sur la mise en correspondance des lignes droites et l’analyse multi-niveaux, qui permet d’évaluer les paramètres de distorsion sans avoir recours à une cible matérielle.

	Titre / Référence	Contexte	Objectif	Méthode utilisée	Données	Limites / Points positifs
1	OCamCalib Toolbox – Scaramuzza et al. (2006)	Calibration caméras omnidirectionnelles pour robotique	Méthode flexible et efficace sans modèle projectif	Modèle polynomial $f(\rho)$, ajustement linéaire + raffinement non linéaire	Images damier, acquisition contrôlée	+ Robuste, largement utilisée; - Nécessite damier et conditions contrôlées
2	Zhang(A Flexible New Technique for Camera Calibration)	Calibration projective classique (pinhole)	Estimer paramètres intrinsèques et extrinsèques via damier	Méthode par homographies, minimisation erreur reprojection	Images damier multiples poses	+ Fondement méthodes modernes; - Non adaptée fisheye
3	Kannala & Brandt	Calibration caméras fisheye (modèle générique)	Modèle générique avec fonctions radiales	Modèle radial 4/6 paramètres, moindres carrés	Données damier ou lignes droites	+ Applicable nombreuses lentilles; - Ajustement complexe
4	Rufli et al(Automatic Checkerboard Detection)	Prétraitement pour calibration automatique	Détection automatique damiers flous ou déformés	Détection via gradients et structure adaptative	Images très déformées	+ Utile en pipeline; - Ne fait pas calibration
5	Wakai et al.(Heatmap Regression under Manhattan World)	Calibration automatique sans cible, environnement urbain	Estimer paramètres d'une image fisheye seule	Deep Learning avec heatmaps orientées Manhattan	Images synthétiques et réelles	+ Sans damier; - Dépend alignement scène et pré-entraînement
6	Yang et al.(Dual Diffusion Architecture)	Correction distortion fisheye post-traitement	Corriger images fisheye pour vision classique	Dual-Diffusion Network avec attention spatiale	Bases de données réalistes	+ Très bonne qualité visuelle; - Pas de paramètres calibration
7	Sun et al.(Precise and Robust Calibration)	Calibration robuste pour fisheye mobiles	Calibration précise avec bundle adjustment	Optimisation géométrique complète	Images damier, poses multiples	+ Robuste aux mouvements; - Implémentation complexe
8	Wang et al.(Self-supervised Fisheye Calibration)	Calibration sans supervision, véhicules autonomes	Estimer paramètres à partir paires images sans GT	CNN + contraintes géométriques auto-supervisées	KITTI, images réelles	+ Pas besoin ground truth; - Moins précis
9	Yin et al.(Fisheye Geometry-Aware Network)	Vision ordinateur sur fisheye (détection, segmentation)	Corriger distorsions pour détection fiable	Réseau géométrique + attention	Données urbaines annotées	+ Couplé à détection; - Pas calibration explicite
10	Li et al. (2020)	Calibration par points de fuite sans damier	Estimer calibration via points de fuite	Méthode analytique + ajustement lignes parallèles	Images naturelles, pas de cibles	+ Sans cible; - Dépend détection lignes fiable

TABLE 2 – Analyse comparative des approches de calibration pour caméras fisheye

8 Conclusion

Les caméras fisheye offrent un large champ de vision, mais génèrent des distorsions optiques importantes qui affectent la qualité géométrique des images. La calibration est une étape clé pour estimer les paramètres de la caméra et corriger ces distorsions, ce qui permet d'obtenir des images géométriquement fiables.

Les approches classiques de calibration reposent principalement sur l'utilisation de cibles connues, comme des damiers, et sur des modèles mathématiques paramétriques. Ces méthodes sont robustes et bien établies, mais nécessitent souvent une acquisition soignée des images de calibration.

Plus récemment, des méthodes basées sur le deep learning et les réseaux de neurones ont émergé. Ces approches apprennent directement les corrections de distorsion à partir de grandes bases de données d'images, ce qui peut réduire la dépendance aux cibles physiques et améliorer la capacité d'adaptation à des conditions variées. Cependant, elles demandent un entraînement important et peuvent parfois manquer de transparence dans leur fonctionnement.

La correction des images fisheye, qu'elle soit réalisée par des méthodes classiques ou par deep learning, est essentielle pour garantir la fiabilité des traitements d'images en vision par ordinateur, robotique, et autres domaines liés.

Contribution

1 Introduction

Ce chapitre présente les principales contributions de ce travail, organisé autour de deux approches complémentaires pour la calibration de caméras fisheye : une méthode classique fondée sur un modèle géométrique explicite (OCamCalib), et une approche basée sur l'apprentissage profond.

Afin de répondre aux besoins spécifiques de ces deux méthodes, Un dataset privé a été constitué en interne. Il regroupe des images avec et sans damier, capturées dans divers environnements.

Dans un premier temps, la méthode de calibration géométrique a été mise en œuvre à l'aide de la toolbox MATLAB **OCamCalib**, afin d'obtenir un modèle précis de la distorsion optique. Cette calibration a permis de redresser les images fisheye et d'explorer différentes techniques de rectification optimisant la restitution du champ de vision.

Dans un second temps, des modèles d'apprentissage profond ont été développés pour estimer les images redressées (via U-Net et GAN), en s'affranchissant totalement des cibles classiques. Cette démarche vise à évaluer le potentiel du deep learning pour la calibration automatique dans des environnements non contraints.

Les sections suivantes détaillent la création du dataset, la mise en œuvre des deux approches et les choix méthodologiques effectués.

2 Collecte et Préparation du Dataset Caplogy

2.1 Crédation d'un dataset Privé Caplogy

Dans le cadre de ce projet, un dataset personnalisé a été élaboré pour les deux méthodes examinées :

La Collecte a été réalisé en interne chez Caplogy en utilisant une caméra fisheye Reolink avec une résolution de 2560 pixels, fournissant une qualité visuelle adéquate pour des missions de calibration minutieuses et l'étude de scènes sophistiquées. Les images ont été recueillies dans différents environnements concrets, tels que des bureaux, des couloirs et des zones communes, dans des conditions d'éclairage naturel et artificiel. On a constitué deux ensembles de données séparés.

— Dataset de calibration (images avec damier)

Ce premier dataset est destiné à l'application de méthodes traditionnelles de calibration. Il se compose d'environ 20 images comportant un damier de calibration utilisé comme référence. Les images ont été capturées sous différents angles, orientations et distances, en accord avec les critères des méthodes géométriques, qui préconisent généralement entre 10 et 20 images. Le motif en damier permet une identification automatique des coins, et la résolution élevée des images augmente la précision dans la localisation des points de repère.

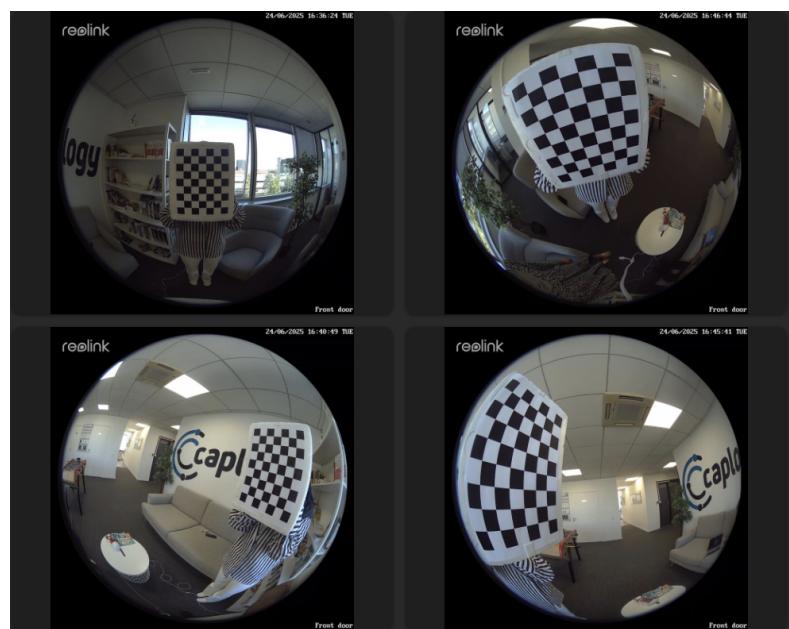


FIGURE 1 – Echantillons Dataset de calibration

— Dataset d'évaluation (scènes réelles sans damier)

Le second ensemble de données, élaboré pour des inférences en conditions réelles, compile plus de cent images fisheye illustrant une diversité de scènes naturelles. Il offre la possibilité d'analyser de façon qualitative et quantitative les performances des techniques de correction et des modèles d'apprentissage profond sans faire appel à des motifs de calibration.



FIGURE 2 – Echantillons Dataset d'inférence

Jeu de données	Nombre d'images	Caméra	Modèle	Résolution
Dataset de calibration	20	Reolink fisheye	FE-W	2560px
Dataset d'inférence (scènes réelles)	103	Reolink fisheye	FE-W	2560px

TABLE 3 – Résumé des deux datasets privés Caplogy

3 Approche classique de calibration

3.1 Pipeline de calibration avec la toolbox `OCamCalib`

La première méthode envisagée pour le projet se base sur la technique traditionnelle de calibration de Scaramuzza à l'aide de la boîte à outils MATLAB `OCamCalib`. Cela permet d'évaluer les paramètres intrinsèques et extrinsèques des caméras fisheye en utilisant des images de damiers.

Le processus de calibration suit les étapes suivantes :

3.2 Processus de calibration selon Scaramuzza

La calibration de la caméra fisheye, mise en œuvre via MATLAB `OCamCalib`, suit les étapes principales suivantes :

- **Acquisition d'images** : capture de plusieurs images d'un damier sous différents angles et positions.
- **Modélisation du capteur** : approximation de la distorsion radiale par une série de Taylor sans terme linéaire pour modéliser la fonction $g(\rho)$.
- **Estimation des paramètres extrinsèques** : projection des coins du damier et résolution linéaire pour obtenir la matrice de rotation \mathbf{R} et le vecteur de translation \mathbf{t} via SVD.
- **Estimation des paramètres intrinsèques** : résolution linéaire des coefficients du polynôme de distorsion et autres paramètres intrinsèques à partir des paramètres extrinsèques.
- **Raffinement linéaire** : correction des estimations initiales par minimisation des moindres carrés pour améliorer la précision.
- **Recherche du centre image** : optimisation du centre de l'image pour minimiser la somme des erreurs de reprojection.
- **Raffinement non linéaire** : optimisation globale des paramètres intrinsèques et extrinsèques via l'algorithme de Levenberg–Marquardt.
- **Validation de la calibration** : évaluation de la qualité par projection des rayons sur les images et mesure de l'erreur de reprojection.

3.3 Redressement d'image fisheye basée sur le modèle de calibration

une fois le modèle calibré obtenu, une procédure externe à la toolbox permet de redresser les images fisheye à l'aide de la fonction `undistort`. Cette fonction utilise les paramètres du modèle `ocam_model` pour produire une version rectifiée de l'image, en corrigeant la distorsion radiale.

Cette étape permet de visualiser l'impact réel de la calibration sur les images fisheye.

3.4 Méthodes de rectification pour un champ de vision maximal

Une fois les images fisheye calibrées à l'aide d'`OCamCalib`, il est possible de générer des images redressées fidèles à la géométrie du système optique. À partir de ces images rectifiées, plusieurs méthodes de post-traitement peuvent être appliquées pour optimiser l'exploitation du champ de vision.

Ces techniques ont pour objectif de transformer les images rectifiées en représentations plus adaptées à des besoins spécifiques. Elles conservent au maximum le champ angulaire capté par la caméra, tout en adaptant la géométrie de la projection. Le tableau ci-dessous présente un comparatif des principales méthodes utilisées dans ce cadre.

Méthode	Objectif	Principe	Limites
Vue perspective paramétrée via <code>OCamCalib</code>	Couvrir un large champ avec plusieurs vues orientées	Caméras virtuelles dans différentes directions ; champ paramétré (90° – 150°) ; vues fusionnables	Champ global reconstruit partiellement ; discontinuités entre vues ; assemblage soigné nécessaire
Projection cylindrique	Créer une vue panoramique horizontale	Projection sur cylindre virtuel ; déroulement horizontal avec lignes verticales linéaires	Perte d'information verticale ; courbure si cylindre mal aligné
Vues sphériques multiples	Couvrir toutes les directions autour de la caméra	Vues selon 6 axes (avant/arrière, gauche/droite, haut/bas) ; reconstitution sphérique	Traitements lourds ; navigation complexe ; redondances visuelles
Projection équirectangulaire	Vue panoramique lisible sur 360°	Conversion fisheye → coordonnées sphériques (azimut/élévation) → re-projection cartésienne	Introduit encore quelques déformations sur les bords extrêmes de l'image.

TABLE 4 – Comparatif des méthodes de rectification pour un champ de vision maximal

Les méthodes présentées constituent des traitements de post-rectification appliqués sur les images redressées à l'aide du modèle de calibration `OCamCalib`. Parmi

elles, la **projection équirectangulaire** se distingue comme étant la plus pertinente dans un contexte de restitution du champ de vision complet. Elle permet une représentation continue de la scène, tout en assurant une bonne linéarité des structures verticales et horizontales au centre de l'image. Ce type de projection facilite ainsi la lecture globale de la scène et s'avère particulièrement adapté à des applications de visualisation immersive ou d'analyse globale de l'environnement.

Toutefois, cette méthode présente certaines *limitations*, notamment des déformations géométriques aux bords extrêmes de l'image, résultant de la conversion sphérique–cartésienne.

4 Approche basée sur le deep learning

Dans la partie précédente, nous avons étudié en profondeur la méthode de calibration géométrique de Scaramuzza, fondée sur une modélisation analytique précise de la distorsion. Bien que cette approche soit robuste et largement utilisée, elle présente certaines limites structurelles qui freinent son extension à des contextes plus complexes ou moins bien contrôlés.

C'est dans ce contexte que se situe l'intérêt de la transition vers l'**apprentissage profond**, une option plus adaptable, apte à acquérir des représentations complexes directement à partir des données. Effectivement, grâce à l'avancement d'IA, en particulier des réseaux de neurones profonds, de nouvelles perspectives se dessinent pour optimiser la calibration, surtout face à des distorsions importantes, des scènes naturelles non structurées ou en manque de cibles de calibration explicites.

Dans cette perspective, nous suggérons l'élaboration et de deux modèles de calibration basés sur l'apprentissage profond, spécialement conçus pour le traitement des images fisheye. Les sous-sections suivantes présenteront ces méthodes en détail.

4.1 Préparation et traitement des données d'entraînement

L'efficacité des modèles d'apprentissage profond dépend principalement de la qualité et de la pertinence des données utilisées pour l'entraînement. Dans ce projet, nous employons des ensembles de données constitués d'images fisheye, prises dans divers environnements, pour former un modèle capable de calibrer des images fisheye, sans faire appel à des repères géométriques explicites. Donc La sélection, l'organisation et le traitement préalable des données (normalisation, mise en forme, ajustement de la résolution) sont donc des étapes cruciales de notre processus.

4.1.1 Datasets réels

Dès le début du projet, une difficulté majeure a été rencontrée : l'absence de jeux de données open source fournissant simultanément des images fisheye brutes (inputs) et leurs versions rectifiées (outputs) obtenues par une calibration précise. Après une recherche approfondie, deux ensembles de données exploitables ont été identifiés. Le tableau ci-dessous en présente les principales caractéristiques :

Nom	Nombre d'images	Type de scènes	Résolution
Corridor-Out	457	Scènes extérieures	3264 × 3264 px
Night-Out	485	Scènes extérieures	3264 × 3264 px
Corridor+Night	942	Fusion des deux	3264 × 3264 px

TABLE 5 – Ensembles de données réelles

L’association des deux ensembles de données a favorisé l’augmentation du nombre d’exemples tout en diversifiant les situations visuelles (éclairage, angle de vue.), ce qui améliore la capacité du modèle à se généraliser.

4.1.2 Dataset synthétique *Multi-FoV*

Afin de pallier le manque de données dans les images fisheye réelles, un jeu de données synthétique a été utilisé afin d’augmenter la diversité des scènes d’entraînement. Il s’agit du dataset *Multi-FoV*, généré à l’aide du moteur 3D *Blender*, en utilisant un modèle de caméra omnidirectionnelle.

Dans cette étude, l’environnement urbain ”City Environment” a été utilisé. une scène citadine, où Cette scène recrée le mouvement d’un véhicule dans une ville, avec des textures authentiques et des structures architecturales qui facilitent l’analyse géométrique.

Dataset	Type de Scène	Nb images
Urban Canyon	Véhicule en mouvement dans un canyon urbain	2500 Inputs / 2500 Outputs

TABLE 6 – Ensemble de données synthétique

4.2 Modèle de rectification basé sur U-Net

4.2.1 Présentation du modèle

Le modèle conçu dans le contexte de ce projet est basé sur l’architecture U-Net, célèbre pour son efficacité dans les tâches de conversion d’image en image (image-to-image). Dans ce contexte, le U-Net est ajusté pour la rectification d’images fisheye. Il apprend à convertir une image déformée optiquement en une version géométriquement corrigée, sans nécessiter de phase explicite de calibration.

4.2.2 Objectif du modèle

L'objectif principal du modèle est de fournir une méthode de redressement automatique d'images fisheye, sans recourir à une calibration explicite basée sur des paramètres intrinsèques. En apprenant directement à partir de paires d'images fisheye / rectifiées, le modèle vise à généraliser cette transformation sur des scènes diverses, tout en réduisant les artefacts géométriques.

4.2.3 Architecture du modèle

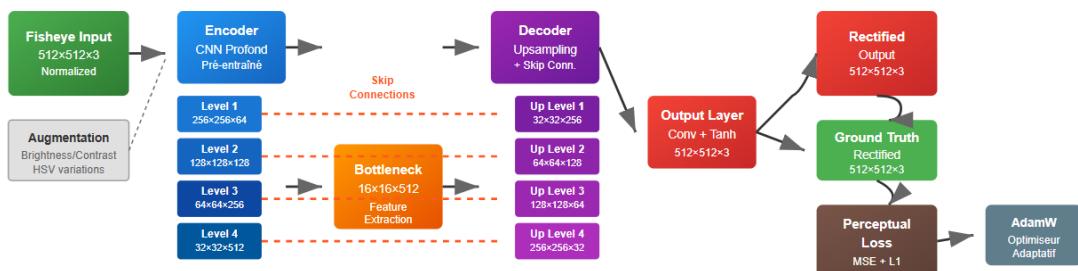


FIGURE 3 – Architecture du modèle U-Net pour la rectification d'images fisheye

Composant	Description
Type de modèle	U-Net : architecture encodeur-décodeur conçue pour les tâches image-to-image.
Encodeur	ResNet34 préentraîné sur ImageNet, utilisé pour extraire des caractéristiques hiérarchiques et profondes.
Décodeur	Bloc de déconvolution (upsampling + convolutions) pour reconstruire l'image dans l'espace d'origine.
Skip connections	Connexions entre couches symétriques du réseau pour préserver les détails fins et la continuité spatiale.

TABLE 7 – Résumé de l'architecture du modèle U-Net utilisé

4.2.4 Choix de conception

Plusieurs choix de conception ont été guidés par la nature du problème :

- L'architecture U-Net est naturellement adaptée aux tâches de transformation géométrique grâce à sa symétrie et à ses connexions de saut, essentielles pour la conservation des structures fines dans les images rectifiées.
- L'utilisation de ResNet34 comme encodeur permet de bénéficier d'un modèle robuste et bien optimisé, capable de généraliser efficacement même avec un nombre limité de données.

- La structure convolutionnelle permet de capturer localement les distorsions induites par les lentilles fisheye, tandis que le décodage progressif reconstruit une image cohérente.

4.2.5 Métriques d'évaluation

Métrique	Description
PSNR	Peak Signal-to-Noise Ratio, mesure la qualité de reconstruction en comparant l'image prédite et la vérité terrain (en dB, plus élevé = meilleur).
SSIM	Structural Similarity Index Measure, évalue la similarité structurelle et perceptuelle entre deux images (valeur entre 0 et 1).
MAE	Erreur absolue moyenne (Mean Absolute Error) entre les pixels des deux images.
Loss train/ Loss val	fonctions de coût calculées respectivement sur les données d'entraînement et de validation.

TABLE 8 – Résumé des métriques utilisées pour l'évaluation – U-Net

4.3 Modèle de rectification basé sur un GAN conditionnel (cGAN)

4.3.1 Présentation du modèle

Ce second modèle repose sur un GAN conditionnel (cGAN) de type Pix2Pix pour rectifier les images fisheye. Il comprend **un générateur** U-Net, chargé de produire des images rectifiées à partir des images déformées, et **un discriminateur** PatchGAN, qui évalue le réalisme des images générées par rapport aux images rectifiées réelles. Un **réseau CNN de régression** est également intégré pour estimer les paramètres de calibration optique($f_x, f_y, c_x, c_y, k_1, k_2, k_3, k_4$). Cette architecture permet d'apprendre simultanément la correction géométrique et le réalisme visuel.

4.3.2 Objectif du modèle

L'objectif de ce deuxième modèle est double :

- Générer une image rectifiée visuellement plausible à partir d'une image fisheye.

- Estimer simultanément les paramètres intrinsèques de la caméra fisheye, ce qui peut être utile pour des tâches de modélisation 3D ou de reconstruction.

Cette approche unifie en un seul modèle les aspects visuels et numériques de la calibration, tout en bénéficiant des capacités génératives et adaptatives du GAN.

4.3.3 Architecture du modèle

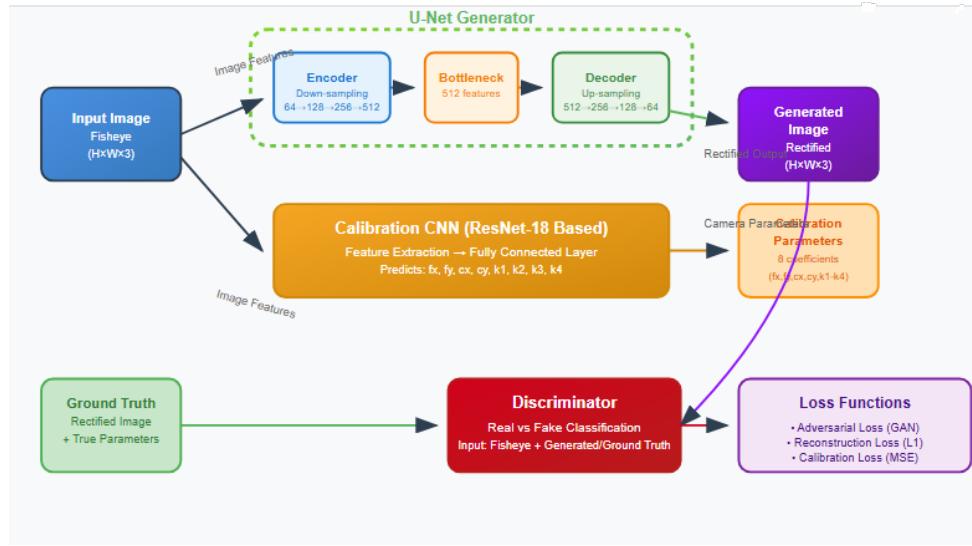


FIGURE 4 – Architecture du modèle GAN pour la rectification d’images fisheye

Le tableau ci-dessous résume les principaux composants de l’architecture GAN utilisée :

Composant	Description
Type de modèle	GAN conditionnel de type Pix2Pix pour la transformation image-to-image
Générateur	Architecture U-Net entièrement convolutionnelle, 4 blocs d’encodage + 4 blocs de décodage avec connexions de saut
Discriminateur	Discriminateur de type PatchGAN (convolutions locales) qui juge la qualité de patchs image
Réseau de calibration	CNN basé sur ResNet18 modifié pour régression directe de 8 paramètres : fx, fy, cx, cy, k1–k4
Fonctions de perte	Le GAN utilise <code>BCEWithLogitsLoss</code> + <code>L1Loss</code> (pondérée à $\lambda = 100$). Le réseau de calibration est entraîné séparément avec <code>MSELoss</code> .

TABLE 9 – Résumé de l’architecture GAN pour la calibration fisheye

4.3.4 Choix de conception

Les choix suivants ont guidé la conception du modèle :

- L'utilisation d'un GAN permet une génération d'image redressée réaliste, en intégrant des contraintes visuelles par l'intermédiaire du discriminateur.
- Le générateur U-Net permet d'exploiter les bénéfices du premier modèle, notamment les connexions de saut utiles à la précision géométrique.
- Le PatchGAN discrimine à l'échelle locale, ce qui favorise une qualité visuelle homogène sur l'image redressée.
- La régression directe des paramètres optiques offre un complément numérique à la transformation visuelle.

4.3.5 Métriques d'évaluation

Les métriques utilisées pour évaluer le modèle sont résumées dans le tableau 10.

Métrique	Description
PSNR	Peak Signal-to-Noise Ratio, mesure la qualité de reconstruction entre l'image prédite et la vérité terrain .
SSIM	Structural Similarity Index Measure, évalue la similarité structurelle et perceptuelle entre les images.
MAE paramètres	Erreur absolue moyenne entre les paramètres prédits (f_x , f_y , c_x , c_y , k_1 – k_4) et les vrais paramètres de calibration.
GAN Losses	permettant d'évaluer la stabilité de l'entraînement(G ;D).

TABLE 10 – Résumé des métriques utilisées pour l'évaluation

5 Conclusion

Les contributions de ce travail se situent à la croisée des approches classiques de calibration géométrique et des méthodes émergentes basées sur l'apprentissage profond.

La méthode de Scaramuzza, implémentée via la toolbox `OCamCalib`, a constitué la base de départ de ce travail. Nous l'avons exploitée comme plateforme de référence, en y apportant une contribution sous forme de post-traitement visant à préserver le champ de vision (FOV) après redressement. Bien que cette approche permette une calibration précise, elle reste contrainte par plusieurs limitations : dépendance à des conditions expérimentales rigides, nécessité de cibles structurées, temps de traitement important et difficulté d'adaptation à des scénarios dynamiques ou en environnement réel.

À l'inverse, les approches par deep learning — notamment via des architectures U-Net et GAN offrent une alternative prometteuse. Cette capacité à fonctionner dans des contextes visuels complexes ouvre des perspectives vers une calibration plus autonome, adaptative et déployable sur des systèmes embarqués.

Ce travail jette ainsi les bases d'un changement de paradigme : passer d'une calibration contrainte et ponctuelle à une calibration continue et contextuelle, portée par le deep learning.

Expérimentations et résultats

1 Introduction

Ce chapitre met en lumière les résultats majeurs obtenus lors de l'évaluation des deux approches mises en place pour la rectification de distorsion fisheye : la méthode classique via OCamCalib, ainsi que les modèles d'apprentissage automatique (U-Net et GAN).

L'objectif consiste à mesurer l'efficacité de chaque technique en termes de précision, qualité d'ajustement, robustesse et rapidité, dans le cadre d'images prises par une caméra Fisheye.

Cette analyse s'appuie sur des critères quantitatifs et qualitatifs, pour souligner les avantages et les contraintes de chaque méthode en fonction des besoins applicatifs.

2 Résultats Expérimentations-Méthode Classique

2.1 Évaluation de la méthode classique (OCamCalib)

Ce passage détaille les étapes clés de la boîte à outils OCamCalib, en mettant l'accent sur celles qui influencent directement la qualité de calibration.

— Extract grid corners:

Détection automatique des points d'intersection du damier sur chaque image. La précision de cette étape conditionne directement la qualité de la calibration.

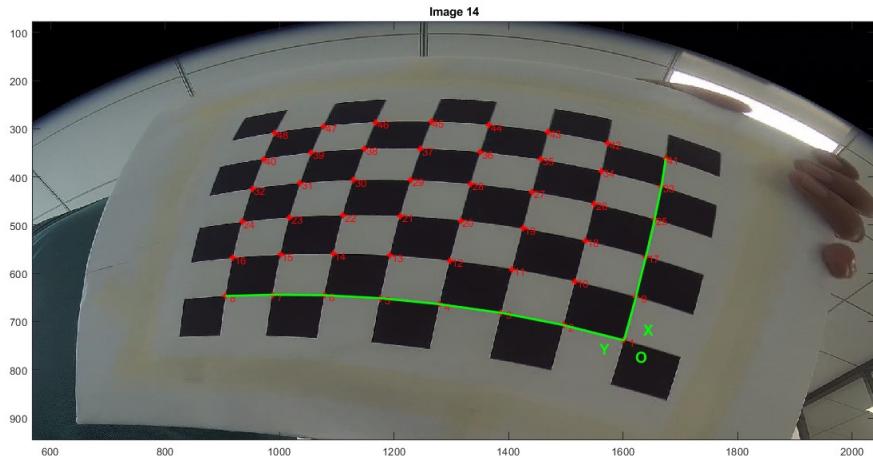


FIGURE 1 – Extraction automatique des coins du damier.

Cette figure illustre l'étape de détection automatique des coins du damier de calibration. L'algorithme a identifié avec succès tous les points d'intersection du motif en échiquier et les a numérotés pour le traitement.

- **Calibration:** Calcul initial des paramètres intrinsèques et extrinsèques du modèle polynomial.

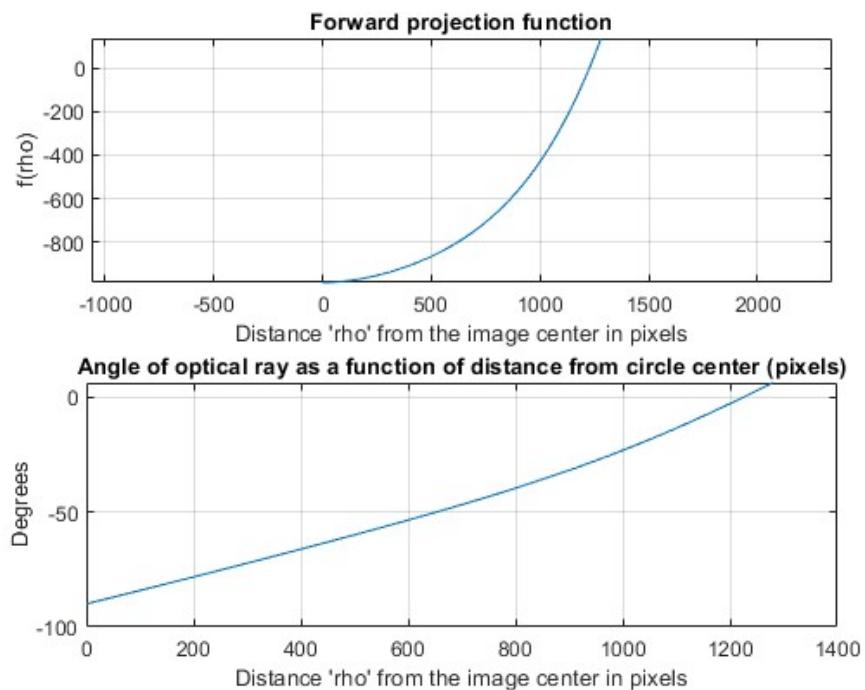


FIGURE 2 – Estimation initiale des paramètres.

Ces graphiques présentent les caractéristiques du modèle fisheye après calibration. Le premier graphique montre la fonction de projection qui décrit la relation entre distance radiale et projection géométrique. Le second établit la correspondance entre position dans l'image et angle optique, révélant un champ de vision d'environ 60 degrés pour cette caméra fisheye.

— **Reprojection images:**

Superposition des points projetés par le modèle sur les points détectés dans les images originales, permettant de vérifier la justesse du modèle.

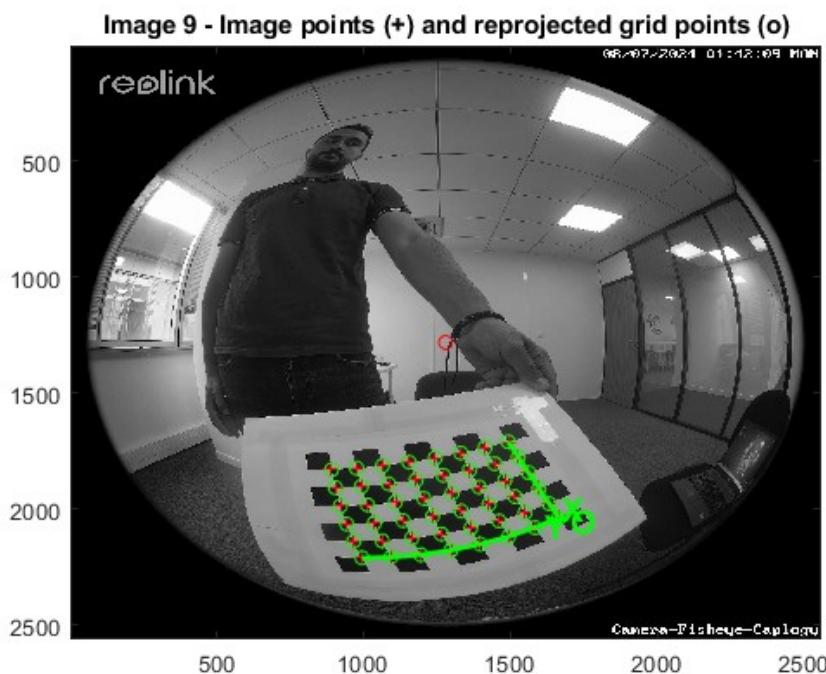


FIGURE 3 – Superposition des projections sur les coins détectés.

Cette image montre la validation du modèle de calibration obtenu. Les croix vertes représentent les points reprojetés par notre modèle calibré, superposés à l'image réelle du damier. L'alignement précis entre ces points théoriques et les positions réelles confirme que la calibration a été réalisée avec succès. **Find center Et Calibration Refinement:** Ces deux étapes d'affinement visent à réduire l'erreur de reprojection en ajustant le centre optique et en optimisant l'ensemble des paramètres de calibration.

— **Analyse error:** Mesure de l'erreur de reprojection pour chaque image et calcul de l'erreur moyenne.

```
Average reprojection error computed for each chessboard [pixels] :

0.71 ± 0.39
0.65 ± 0.34
0.71 ± 0.31
0.99 ± 0.53
0.97 ± 0.55
0.88 ± 0.61
1.16 ± 0.59
0.53 ± 0.34
0.75 ± 0.41
0.60 ± 0.39
1.15 ± 0.74
1.04 ± 0.63
0.74 ± 0.41
0.93 ± 0.56

Average error [pixels]

0.842269
```

FIGURE 4 – Analyse des erreurs de reprojection par image.

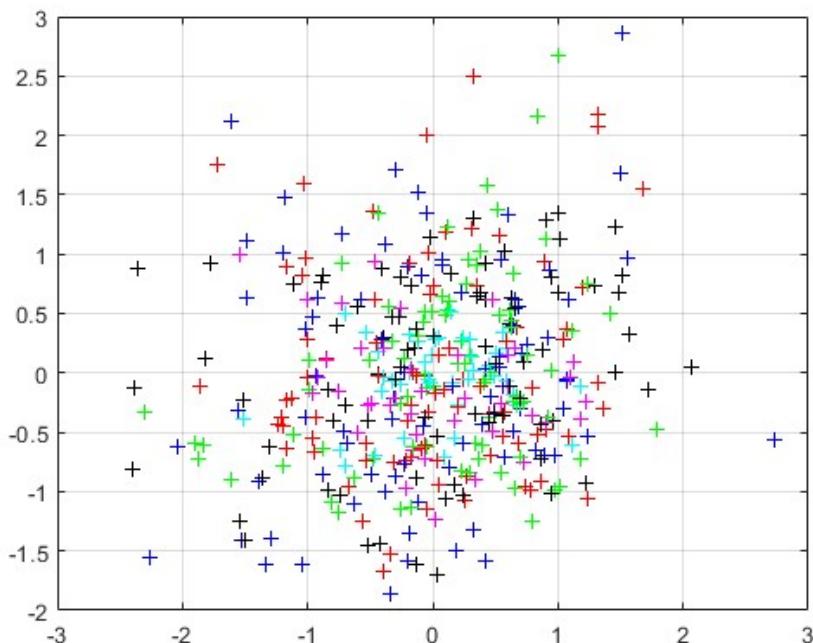


FIGURE 5 – Nuage de points Erreurs.

Ce graphique analyse la précision de notre calibration en visualisant les erreurs de reprojection. Chaque point représente l'écart entre une position détectée et sa position calculée par le modèle. La concentration des erreurs autour de zéro avec une dispersion

inférieure à un pixel démontre la qualité de la calibration obtenue.

2.2 Précision de la calibration

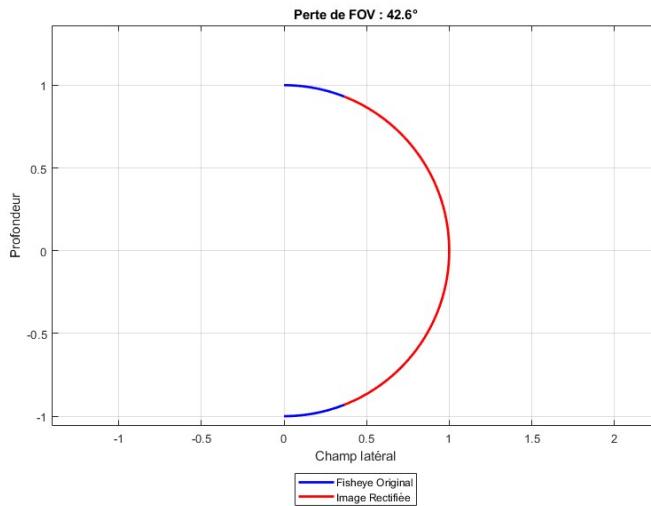


FIGURE 6 – Comparaison du champ latéral et de la profondeur.

La Figure 6 illustre la différence entre le champ couvert par l'image fisheye originale (courbe bleue) et celui de l'image rectifiée (courbe rouge). On observe une perte angulaire d'environ **42.6°**, due à la calibration pour obtenir une projection perspective géométriquement correcte.

Pour quantifier cette perte, la Figure 7 compare directement le FOV calibré et le FOV perdu. Le FOV initial dépassait les **100°**, tandis que le FOV exploitable après calibration est réduit à environ **78°**.

Ce résultat confirme que la méthode de calibration est efficace pour corriger la distorsion, mais au prix d'une réduction notable de la couverture angulaire.

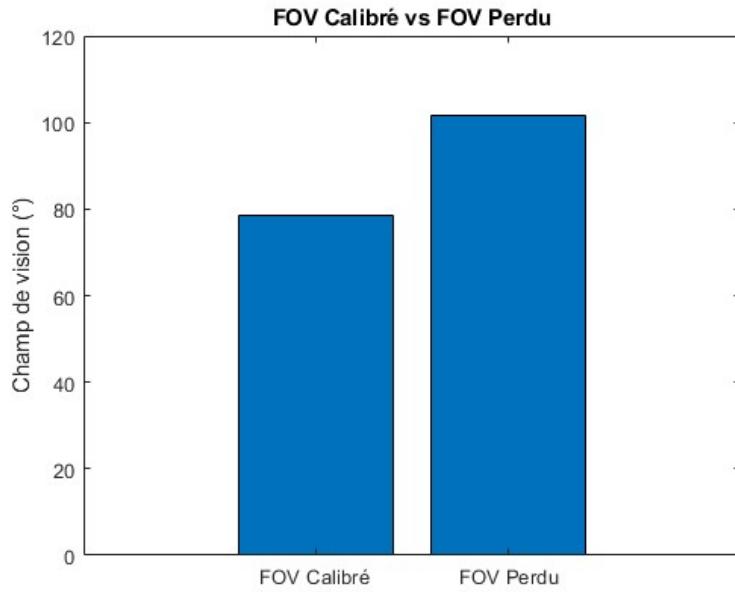


FIGURE 7 – Comparaison quantitative entre le champ de vision calibré et le champ perdu.

2.3 Résultats visuels de redressement

Après estimation des paramètres de calibration via la toolbox OCamCalib, nous avons procédé au redressement des images fisheye.

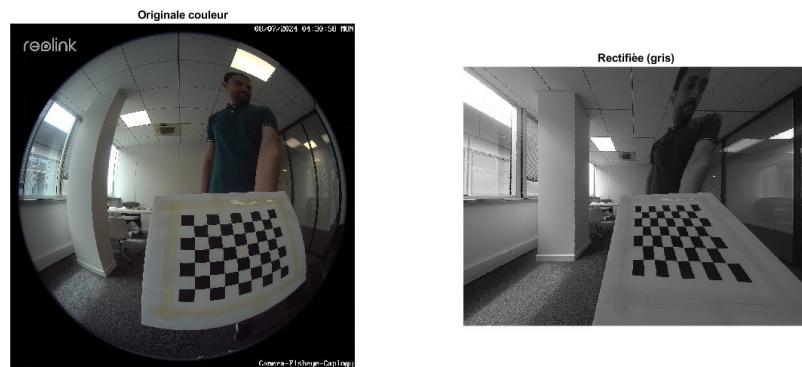


FIGURE 8 – Comparaison entre une image fisheye originale et l'image calibré .

La Figure 8 illustre un exemple représentatif, montrant côté à côté l'image fisheye originale et l'image corrigée. On observe une correction satisfaisante de la distorsion, les

lignes droites étant restaurées dans l'image rectifiée, confirmant la bonne prise en compte de la distorsion radiale par le modèle.

Cependant, comme observé lors de l'analyse du champ de vision ([figure 6](#)), la correction entraîne une réduction perceptible et significative de la zone couverte, notamment dans les bords de l'image. Cette limitation, nous a conduit à envisager l'application de méthodes spécifiques de restitution du champ de vision.

2.4 Méthodes de rectification (FOV)

Afin de maximiser le champ de vision exploitable après calibration via `OCamCalib`, plusieurs approches de post-rectification ont été testées ;parmi ces méthodes, seule la projection équirectangulaire a produit un rendu visuellement satisfaisant et directement exploitable pour la restitution complète du champ de vision. Les autres, bien que techniquement valides, présentaient trop de limitations pour l'application visée et ne sont donc pas illustrées ici.

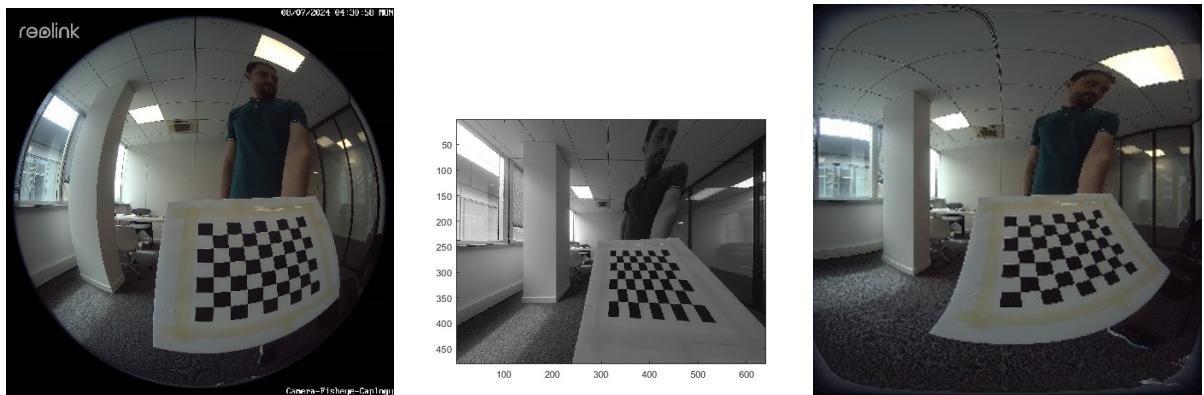


FIGURE 9 – Comparaison entre l'image fisheye originale l'image rectifiée par méthode classique et l'image après la méthode de rectification.

La projection équirectangulaire permet de restituer l'ensemble du champ de vision avec un bon redressement des lignes droites, améliorant nettement la lisibilité de la scène. Cependant, elle introduit encore quelques déformations géométriques aux bords extrêmes de l'image.

3 Expérimentations et Résultats – Modèles Deep Learning

3.1 Modèle 1 : Modèle de rectification basé sur U-Net

3.1.1 Environnement d’entraînement

Le modèle a été entraîné sur un jeu de données de paires images fisheye / rectifiées. Les paramètres d’entraînement utilisés sont résumés dans le tableau suivant :

Élément	Valeur
Plateforme	Kaggle
GPU	CUDA (GPU Kaggle)
Frameworks et Librairies	PyTorch, segmentation-models-pytorch, Albumentations, OpenCV, NumPy, Matplotlib, scikit-learn, tqdm
Taille dataset total	2500input/2500 output images
Split dataset	80% pour l’entraînement, 20% pour la validation
Taille entrée réseau	(Images originales (3264×3264) redimensionnées à (512×512))
Batch size	8
Nombre d’époques	100
Optimizer	AdamW (Optimiseur Adam avec diminution du taux d’apprentissage selon la stagnation de la validation)
Patience Early Stopping	Activé

TABLE 1 – Résumé de l’environnement d’entraînement – U-Net

3.1.2 Évaluation visuelle de la correction fisheye

Cette partie vise à exposer qualitativement les résultats générés par le modèle U-Net. Nous mettons en place une comparaison visuelle entre les images fisheye originales, celles qui ont été rectifiées par le modèle, et leur référence de ground truth.



FIGURE 10 – Comparaison visuelle entre image fisheye, correction U-Net et vérité terrain.

Cette comparaison visuelle démontre l’efficacité du modèle pour corriger la distorsion fisheye sur une scène urbaine. L’image originale présente une courbure caractéristique très prononcée.

La rectification permet une correction efficace de la distorsion fisheye, restituant une perspective cohérente de la scène urbaine et préservant la géométrie architecturale.

On peut néanmoins observer une légère perte de netteté dans l’image rectifiée comparativement à la référence, ainsi qu’un léger flou dans les zones périphériques, ce qui reste acceptable compte tenu de la complexité de la correction appliquée à cette distorsion fisheye importante.

3.1.3 Graphiques des métriques

Afin d’évaluer la qualité et la stabilité de l’entraînement, nous présentons dans cette section les métriques quantitatives, incluant les courbes de PSNR,SSIM,MAE ainsi que les fonctions de perte.

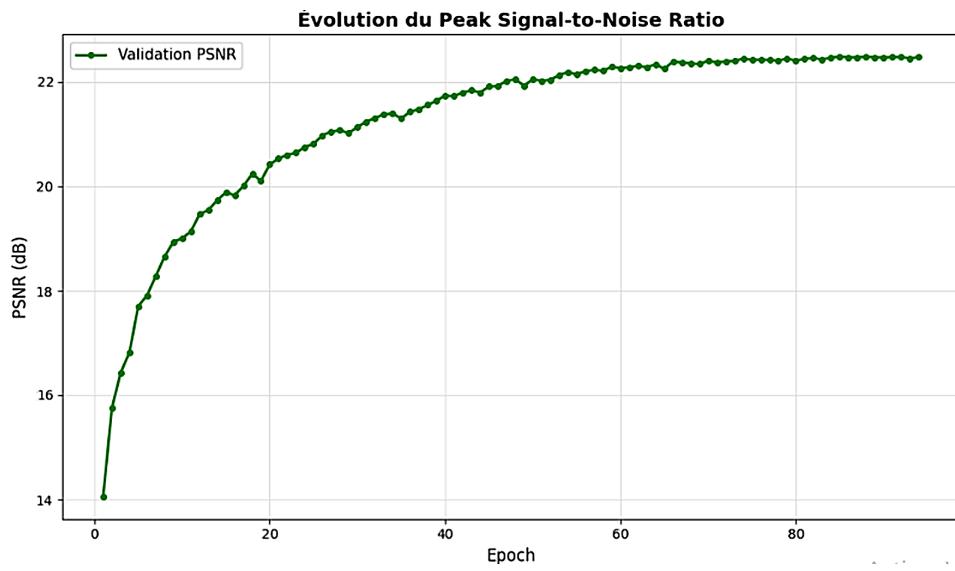


FIGURE 11 – Évolution du Peak Signal-to-Noise Ratio durant l’entraînement.

Cette courbe illustre l’amélioration progressive de la qualité de reconstruction du modèle. Le PSNR démarre à environ 14 dB et connaît une croissance rapide durant les 20 premières époques, atteignant 17 dB. L’amélioration se poursuit de manière plus graduelle jusqu’à stabilisation autour de 23 dB après 70 époques. Cette évolution témoigne d’un apprentissage efficace avec une convergence stable vers des performances optimales de reconstruction d’image.

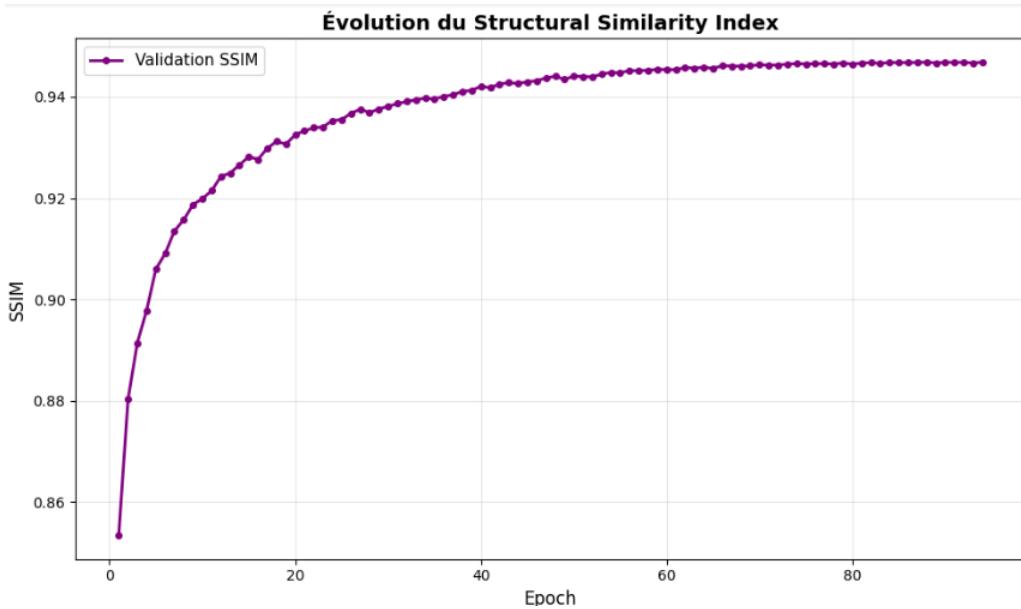


FIGURE 12 – Évolution du Structural Similarity Index durant l’entraînement.

L'indice SSIM suit une progression similaire au PSNR, partant d'une valeur initiale de 0,85 pour atteindre rapidement 0,93 dès les premières époques. La métrique continue de s'améliorer progressivement jusqu'à se stabiliser à environ 0,947 après 60 époques. Cette évolution démontre que le modèle apprend efficacement à préserver les structures et textures importantes de l'image lors de la correction de distorsion, maintenant une très bonne similarité structurelle avec les images de référence.

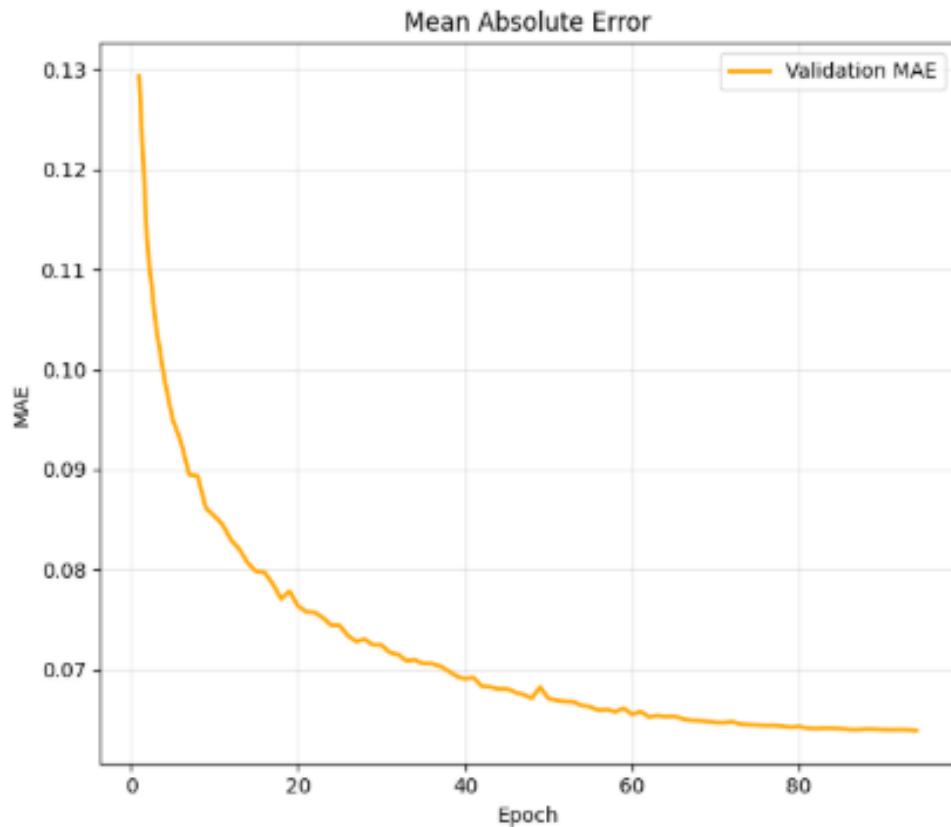


FIGURE 13 – Évolution de l'erreur absolue moyenne durant l'entraînement.

La courbe du MAE présente une décroissance rapide et marquée, passant de 0,13 à environ 0,07 en seulement 20 époques. L'erreur continue de diminuer de façon plus modérée pour atteindre approximativement 0,065 en fin d'entraînement. Cette diminution constante de l'erreur pixel par pixel confirme la capacité du modèle à produire des prédictions de plus en plus précises au fil de l'entraînement, avec une convergence satisfaisante vers des valeurs d'erreur faibles.

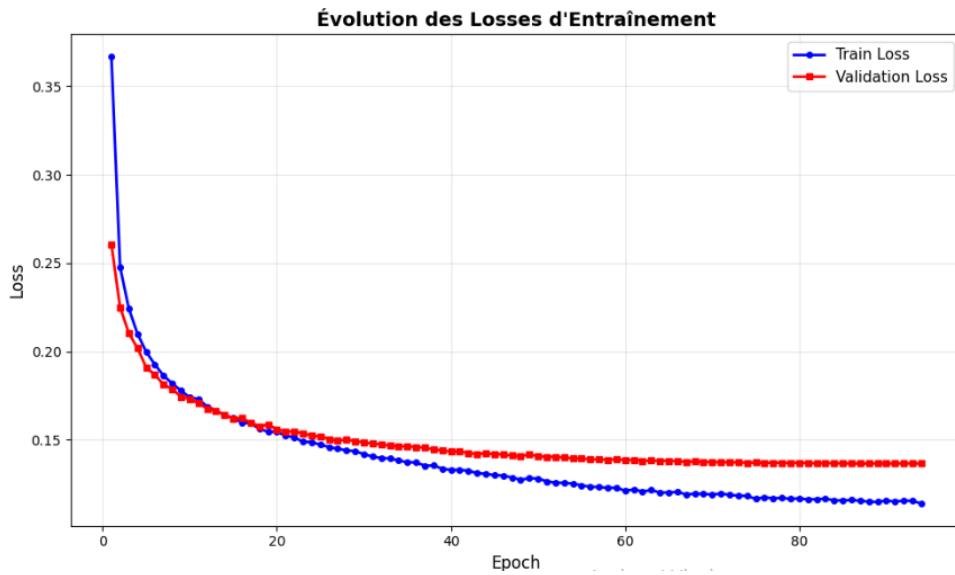


FIGURE 14 – Évolution des losses d’entraînement et de validation.

Cette courbe démontre un apprentissage sain du modèle avec une décroissance rapide et simultanée des losses d’entraînement (bleu) and de validation (rouge). Les deux courbes partent de valeurs élevées (0,37 pour l’entraînement, 0,26 pour la validation) et chutent rapidement durant les 20 premières époques pour se stabiliser respectivement autour de 0,12 et 0,14. La proximité constante entre les deux courbes et l’absence d’écart croissant indiquent que le modèle généralise bien sans surapprentissage significatif.

3.1.4 Analyse Globale des performances

Cette section interprète les résultats présentés ci-dessus.

Aspect Évalué	Critères d'Analyse	Résultats Observés
Qualité Visuelle	Correction des distorsions fisheye	Rectification efficace des courbures, géométrie architecturale bien restituée
	Préservation des détails	Légère perte de netteté en périphérie, flou résiduel aux zones extrêmes
	Cohérence perspective	Perspective urbaine cohérente, lignes droites correctement alignées
Performances Quantitatives	Qualité de reconstruction (PSNR)	Amélioration constante avec convergence stable vers des valeurs élevées
	Préservation structurelle (SSIM)	Excellente conservation des structures et textures importantes
	Précision pixel (MAE)	Diminution progressive de l'erreur avec convergence satisfaisante
Stabilité d'Apprentissage	Convergence du modèle	Apprentissage stable sans oscillations
	Généralisation	Performances équilibrées entre entraînement et validation

TABLE 2 – Analyse Globale des Performances du modèle

3.2 Modèle 2 : Modèle de rectification basé sur un GAN conditionnel

3.2.1 Environnement d'entraînement

Le tableau suivant présente un résumé de l'environnement d'entraînement du modèle

Élément	Valeur
Frameworks et Librairies	PyTorch, Torchvision, scikit-image, scikit-learn, tqdm, OpenCV, NumPy, Matplotlib
Taille dataset total	2500 images (fisheye + calibrées, générées avec labels.csv)
Split dataset	90% train / 10% test
Taille entrée réseau	(256 × 256 × 3)
Batch size	16
Nombre d'époques max	120
Optimiseur	Adam pour réseaux G, D et CalibrationCNN
Loss Générateur	BCE adversarial + 100× L1 (fidélité visuelle)
Loss Discriminateur	BCE (vrai/faux)
Loss CalibrationCNN	MSE (régression sur paramètres {fx, fy, cx, cy, k1–k4})
Transformations data	Resize , Normalisation [-1, 1], ToTensor

TABLE 3 – Résumé de l'environnement d'entraînement

3.2.2 Évaluation visuelle de la correction fisheye



FIGURE 15 – Comparaison visuelle entre image fisheye, correction U-Net et vérité terrain.

Les deux comparaisons démontrent l'efficacité remarquable du modèle GAN dans la correction des distorsions fisheye caractéristiques. Dans les deux cas, les images fisheye

originales présentent des courbures prononcées, particulièrement visibles sur les lignes droites des bâtiments, des marquages routiers et des structures architecturales. Les images rectifiées par le GAN révèlent une correction substantielle de ces aberrations géométriques, avec un redressement notable.

La restitution de la géométrie perspective naturelle est globalement excellente et montre une forte concordance visuelle avec les images de référence calibrées.

L’alignement entre les images rectifiées par le GAN et les images de référence est particulièrement frappant : les proportions géométriques, l’orientation des structures et la disposition spatiale des éléments urbains correspondent de manière très satisfaisante. Cette forte similarité valide l’efficacité du modèle à généraliser sa correction sur diverses configurations urbaines, maintenant une cohérence spatiale remarquable.

On peut noter une légère atténuation de certains détails fins dans les images rectifiées, mais cela n’altère en rien la qualité globale de la correction géométrique obtenue.

3.2.3 Graphiques des métriques

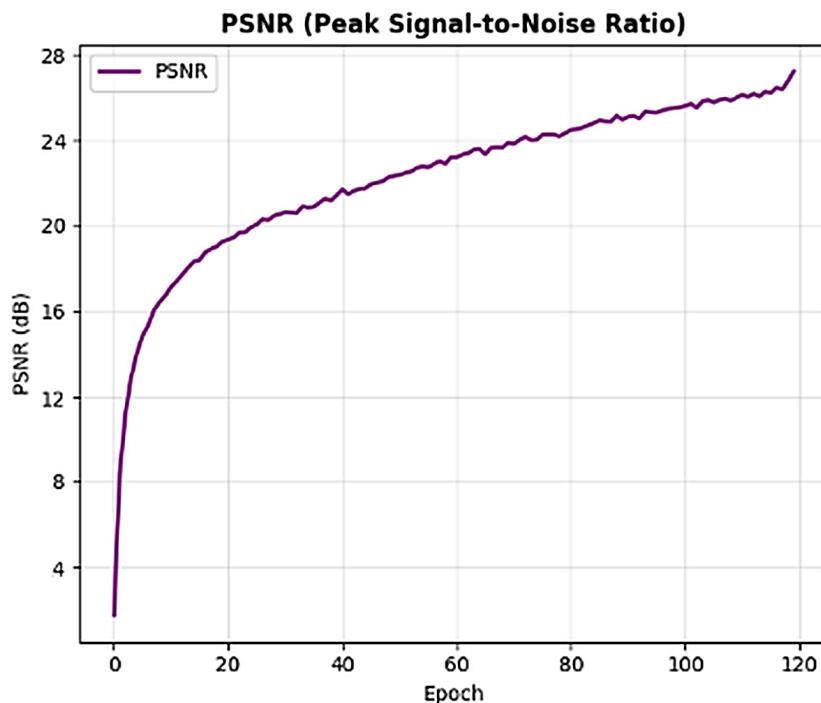


FIGURE 16 – Évolution du Peak Signal-to-Noise Ratio durant l’entraînement.

L’évolution du PSNR au cours de l’entraînement révèle une amélioration continue et significative de la qualité de reconstruction. Débutant à environ 3 dB, la métrique atteint

progressivement jusqu'à 28 dB à l'époque 120, témoignant d'une réduction substantielle du bruit et d'une meilleure fidélité des images rectifiées par rapport aux références. La croissance quasi-monotone de la courbe indique une convergence stable du modèle vers des reconstructions de haute qualité. Cette évolution présente une progression particulièrement rapide lors des 40 premières époques, suivie d'une stabilisation progressive vers la valeur finale. Cette tendance témoigne d'une capacité croissante du modèle à reconstruire fidèlement les données d'entrée.

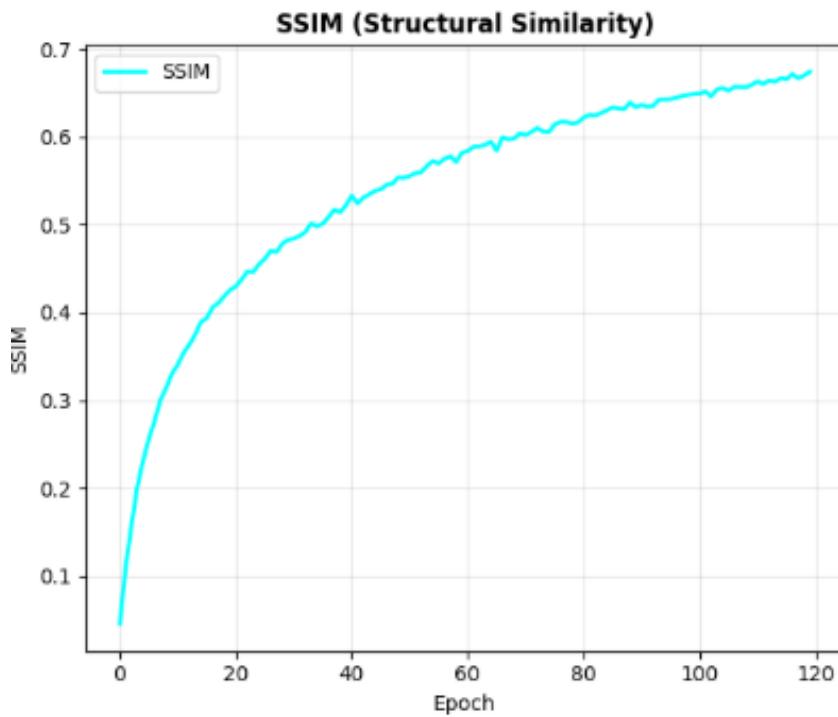


FIGURE 17 – Évolution du Structural Similarity Index durant l'entraînement.

L'indice SSIM montre une amélioration constante de cette métrique, évoluant d'une valeur initiale de 0,08 vers une convergence à 0,70 après 120 époques d'entraînement. La progression suit une courbe sigmoïdale caractéristique avec une accélération notable entre les époques 10 et 60, puis une stabilisation progressive. Une valeur SSIM de 0,68 indique une similarité structurelle satisfaisante, bien que perfectible comparée à la valeur optimale de 1,0. Cette métrique confirme néanmoins que le modèle préserve efficacement les caractéristiques visuelles importantes des données d'origine.

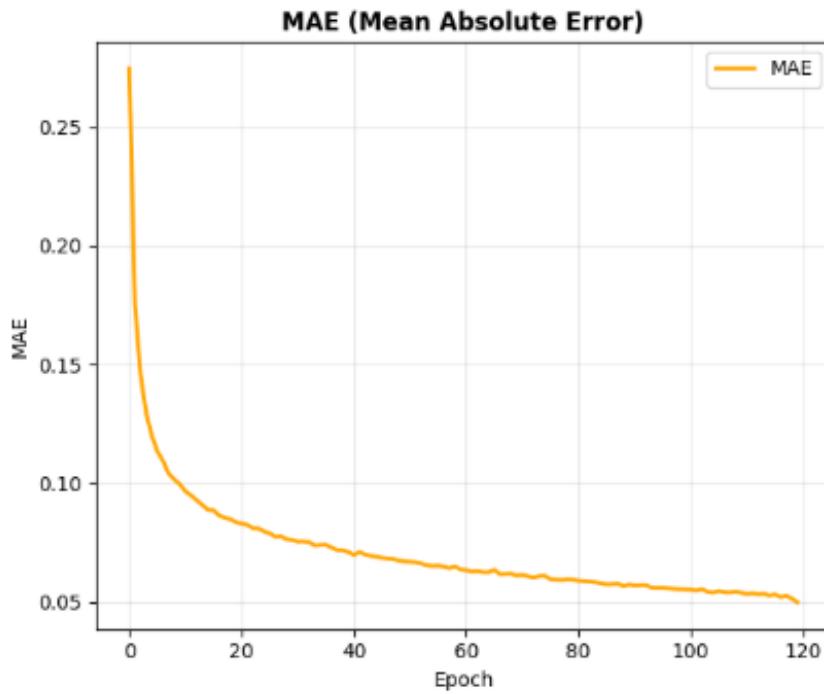


FIGURE 18 – Évolution de l'erreur absolue moyenne durant l'entraînement.

L'analyse de la courbe MAE révèle une diminution exponentielle de l'erreur, partant d'une valeur initiale d'environ 0,28 à l'époque initiale pour converger vers 0,05 à l'époque 120. Cette réduction représente une amélioration de 82 pourcent par rapport à l'erreur initiale. La courbe présente une décroissance particulièrement rapide durant les 20 premières époques, suivie d'une amélioration plus graduelle mais constante. Cette tendance confirme l'efficacité de l'algorithme d'optimisation utilisé et suggère que le modèle atteint progressivement ses limites de performance théoriques.

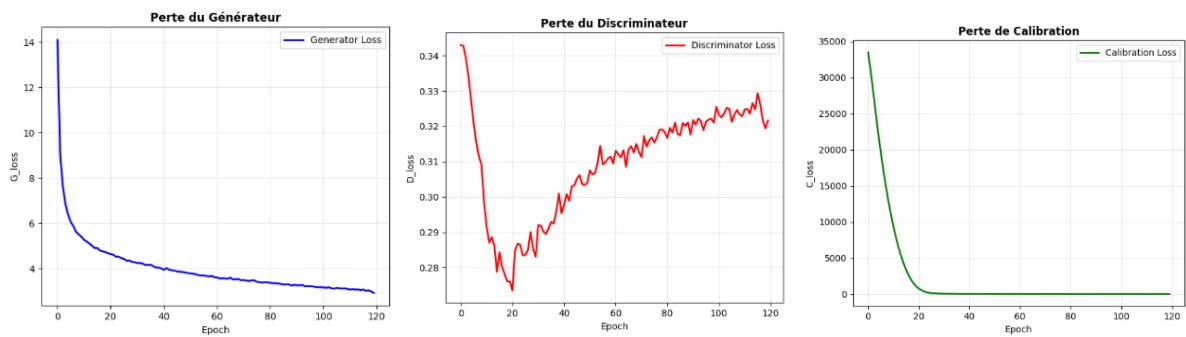


FIGURE 19 – Évolution des Pertes durant l'Entraînement du GAN avec CalibrationCNN.

La Figure ci-dessus révèle la dynamique complexe de l'entraînement adversarial du système hybride GAN-CalibrationCNN sur 120 époques, illustrant parfaitement le "jeu

minimax” entre générateur et discriminateur.

La perte du générateur chute drastiquement de 14 à 3,2, témoignant de son appren-tissage progressif à tromper le discriminateur. Cette décroissance exponentielle reflète l’amélioration de sa capacité à générer des échantillons de plus en plus convaincants et réalistes, forçant le discriminateur à s’adapter constamment.

La perte du discriminateur exprime la nature adversariale de l’entraînement. Sa chute initiale brutale (époque 0-20) révèle sa supériorité temporaire, mais la remontée progres-sive vers 0,32 démontre comment le générateur riposte en devenant plus sophistiqué. Cette “guerre” perpétuelle entre les deux réseaux crée l’oscillation caractéristique.

La perte de calibration présente une convergence spectaculaire, s’effondrant de 35000 vers zéro en seulement 40 époques. Cette chute vertigineuse illustre l’efficacité remarquable du CalibrationCNN à corriger les artefacts générés par le GAN. Cette triple dynamique révèle un entraînement adversarial stable et équilibré, où chaque composante pousse les autres vers l’excellence, créant un système hybride particulièrement robuste.

Aspect Évalué	Critères	Résultats Clés / Observations
Qualité Visuelle	Correction des distorsions fisheye	Courbures redressées, lignes et marquages alignés, structures architecturales correctement rectifiées
	Préservation des détails	Détails fins légèrement atténus, mais qualité globale conservée
	Cohérence perspective	forte concordance visuelle avec les images de référence
Performances Quantitatives	PSNR	Amélioration continue de la qualité de reconstruction et fidélité visuelle élevée
	SSIM	Structures globales bien préservées, similarité structurelle satisfaisante
Stabilité d'Apprentissage	MAE	Réduction significative de l'erreur moyenne, reconstruction précise des pixels
	Convergence du générateur	Apprentissage progressif efficace, pertes stables
	Équilibre adversarial	Oscillations contrôlées, dynamique G-D équilibrée
	CalibrationCNN & Estimation automatique	Convergence rapide et efficace, paramètres de calibration calculés automatiquement

TABLE 4 – Analyse synthétique des performances du modèle GAN

3.3 Analyse Comparative des méthodes de calibration

Critère	Méthode Classique (OCamCalib)	Deep Learning - U-Net	Deep Learning - GAN Conditionnel
Principe de fonctionnement	Calibration géométrique par détection automatique de motifs de damier	Apprentissage supervisé direct de la transformation fisheye → rectifiée	Apprentissage adversarial avec génération conditionnelle haute fidélité
Type de Données	Images de damiers calibrées obligatoires	Images fisheye + rectifiées (paire d'entraînement)	Images fisheye + rectifiées + annotations de calibration
Volume de données	Minimal : quelques dizaines d'images de damier	Modéré : centaines à milliers de paires	Élevé : plusieurs milliers d'images avec annotations
Facilité d'acquisition	Contraintes strictes : conditions contrôlées, damiers parfaits	Flexible : images du monde réel acceptées	Très flexible : toute scène naturelle
Temps de d'entraînement	Rapide : Toolbox MATLAB	Modéré : plusieurs heures avec GPU	Long : heures à jours selon architecture
Qualité de correction	Acceptable mais limitée aux zones calibrées	Bonne correction globale uniforme	Excellente correction optimale complète
Fidélité visuelle	Moyenne avec perte d'information aux bords	Élevée avec structures bien préservées	Très élevée avec détails et textures parfaitement conservés
Précision géométrique	Correcte au centre , distorsions résiduelles en périphérie	Cohérente sur l'ensemble de l'image	Optimale avec perspective naturelle globale
Complexité de déploiement	Simple : algorithme léger	Modérée : modèle de taille moyenne	Complexé : modèle lourd, infrastructure spécialisée
Réaliste du résultat	Aspect "corrigé" facilement identifiable	Rendu naturel globalement convaincant	Photo-réalisme indistinguables de l'image naturelle

TABLE 5 – Comparaison des Méthodes de Correction de Distorsion Fisheye

4 Conclusion

Ce chapitre a exposé une Comparaison expérimentale entre l'approche classique de calibration (OCamCalib) et les approches fondées sur le deep learning(U-Net et cGAN). La méthode classique permet une calibration précise dans une certaine mesure, redressant correctement les structures centrales, mais elle présente des contraintes et limitations importantes face à la forte distorsion de nos images fisheye. En revanche, les modèles de Deep Learning garantissent une correction plus flexible et plus efficace. Le U-Net présente une bonne restauration structurelle, a tandis que le cGAN se distingue en proposant une régression directe des paramètres optiques, constituant ainsi un complément numérique à la rectification d'images et améliorant la qualité perceptuelle des résultats.

En définitive, bien que la calibration classique garde un certain intérêt pratique, les méthodes d'apprentissage profond, notamment le cGAN, apparaissent mieux adaptées aux exigences actuelles de correction d'images fisheye.

Conclusion générale

Ce travail s'inscrit dans la problématique fondamentale de correction des distorsions générées par les caméras fisheye, un défi technique majeur pour la vision par ordinateur moderne et ses applications industrielles émergentes. Nous avons d'abord implémenté la méthode classique de Scaramuzza (OCamCalib) avec un post-traitement pour optimiser les résultats. Cette étape nous a permis d'identifier les limites des approches classiques.

Les résultats obtenus avec nos modèles de deep learning démontrent un saut qualitatif significatif par rapport aux méthodes existantes. L'architecture U-Net développée présente déjà des performances remarquables, avec une correction homogène des distorsions et une préservation optimale des détails visuels. Plus remarquable encore, notre modèle cGAN atteint un niveau de réalisme et de précision géométrique qui approche la qualité photographique naturelle, plus une régression directe des paramètres optiques.

Ces modèles s'adaptent automatiquement à différentes conditions sans recalibration manuelle. Cette flexibilité est un avantage important pour les applications de computer vision.

La constitution de notre dataset privé Caplogy constitue en elle-même une contribution méthodologique importante, fournissant une base d'évaluation rigoureuse qui enrichit considérablement la portée scientifique de nos travaux et ouvre des perspectives pour de futures recherches dans le domaine.

Notre modèle cGAN représente une solution pratique et efficace pour passer d'une calibration manuelle contraignante vers une correction automatique et intelligente. Ces résultats ouvrent des perspectives concrètes pour des systèmes de vision plus performants et autonomes.

Cette étude démontre le potentiel du deep learning pour la correction de distorsions fisheye. Cette approche révolutionnaire ouvre des horizons prometteurs pour le développement de systèmes embarqués de vision plus performants, capables de s'adapter dynamiquement aux exigences opérationnelles sans intervention humaine.

Perspectives pour les Futures Recherches

Optimisation du Modèle cGAN

L'architecture cGAN développée présente un potentiel considérable qui mérite d'être pleinement exploité. Les travaux futurs devraient se concentrer sur l'optimisation et le perfectionnement de ce modèle prometteur, notamment par :

- **Problématique des Datasets**

Le principal obstacle identifié concerne l'accès à des datasets de qualité suffisante. Le manque de données open-source constitue un frein majeur au développement optimal des modèles GAN pour la correction fisheye.

Besoins en Datasets Volumineux et Diversifiés

- **Volume critique** : accès à plusieurs dizaines de milliers de couples d'images fisheye/rectifiées pour permettre un apprentissage robuste.
- **Diversité des conditions** : datasets incluant des scénarios variés.
- **Qualité des couples** : images fisheye parfaitement appariées avec leurs versions rectifiées, conservant un FOV maximal et sans artefacts de calibration.

Enjeux d'Accessibilité aux Données

- **Limitation des datasets publics** : rareté des bases de données open-source spécialisées dans la correction fisheye.
- **Qualité variable** : datasets existants souvent limités en termes de quantité de données, résolution, diversité ou précision de calibration.

— Infrastructure Matérielle et Optimisation GPU

L'entraînement des modèles GAN impose des exigences computationnelles considérables qui nécessitent une infrastructure adaptée et des stratégies d'optimisation avancées.

Équipement GPU Haute Performance

- **Cartes graphiques professionnelles** : GPUs dédiés au deep learning avec mémoire VRAM élevée.
- **Configurations multi-GPU** : systèmes parallélisés permettant de distribuer l'entraînement et réduire significativement les temps de calcul.

- **Mémoire système** : RAM importante (128–256 GB) pour la gestion des datasets volumineux et le preprocessing des images.

Bibliographie

- [1] Aras Jalal Mhamad Karim and Nawzad Muhammed Ahmed. Vector autoregressive integrating moving average (varima) model of covid-19 pandemic and oil price. *International Journal of Professional Business Review : Int. J. Prof. Bus. Rev.*, 8(1) :13, 2023.
- [2] Adedayo Ajayi et al. Energy forecasting model for ground movement operation in green airport. *Energies*, 16(13) :5008, 2023.
- [3] Yimeng Wang. Review of the evolution of statistical models and deep learning models : Based on the perspective of arima and svm. *Highlights in Science, Engineering and Technology*, 70 :339–342, 2023.
- [4] Heng Shi, Minghao Xu, and Ran Li. Deep learning for household load forecasting—a novel pooling deep rnn. *IEEE Transactions on Smart Grid*, 9(5) :5271–5280, 2018.
- [5] Ahmad O. Aseeri. Effective rnn-based forecasting methodology design for improving short-term power load forecasts : Application to large-scale power-grid time series. *Journal of Computational Science*, 68 :101984, 2023.
- [6] Mobarak Abumohsen et al. Electrical load forecasting using lstm, gru, and rnn algorithms. *Energies*, 16(5) :2283, 2023.
- [7] Nishant Saxena et al. Hybrid knn-svm machine learning approach for solar power forecasting. *Environmental Challenges*, 14 :100838, 2024.
- [8] Akshit Kurani et al. A comprehensive comparative study of artificial neural network (ann) and support vector machines (svm) on stock forecasting. *Annals of Data Science*, 10(1) :183–208, 2023.
- [9] Rasoul Ameri et al. Forecasting pm 2.5 concentration based on integrating of ceemdan decomposition method with svm and lstm. *Ecotoxicology and Environmental Safety*, 266 :115572, 2023.
- [10] Antoine Gatera et al. Comparison of random forest and support vector machine regression models for forecasting road accidents. *Scientific African*, 21 :e01739, 2023.