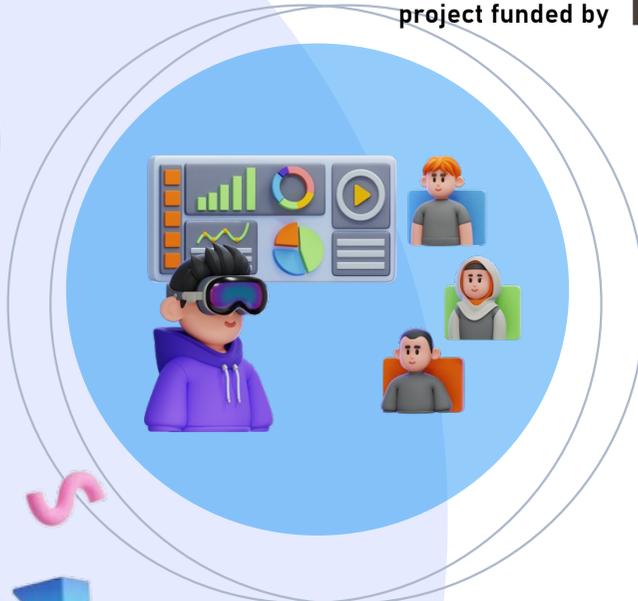


This work was supported by the VISI03D+
project funded by **bpifrance**



Classification au service de la garantie de la
qualité d'expérience pour les utilisateurs
d'un système de diffusion de vidéos à 360°

Par Sidi-Mohammed Senouci



Equipe **SIC** du laboratoire DRIVE

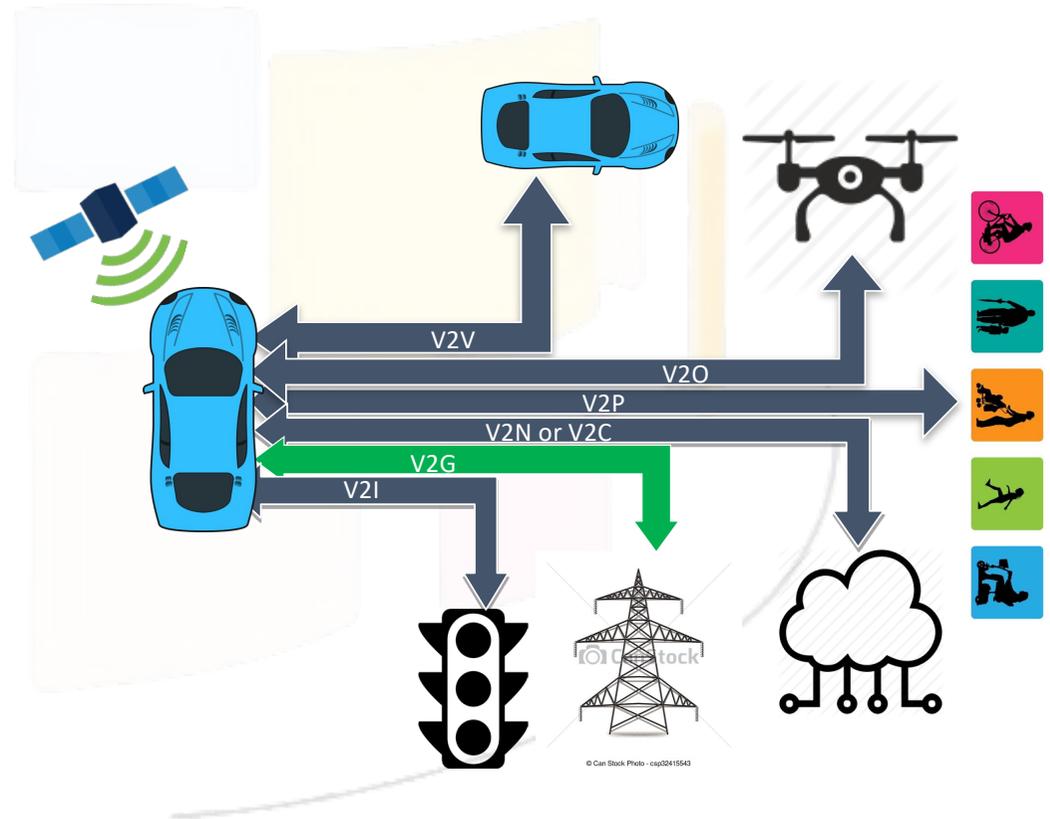


Vision et démarche

L'équipe SIC se concentre sur la **Collecte coopérative**, **Echanges** coopératifs et **sécurisés** et **Traitement** de données **en considérant les ressources limitées** (Réseau, Energie,..) avec des applications comme les **systèmes de transport/mobilité**, les réseaux **SmartGrid** ainsi que les **réseaux de communications sans fil** et la **Cyber sécurité**

Exemple de projets en cours

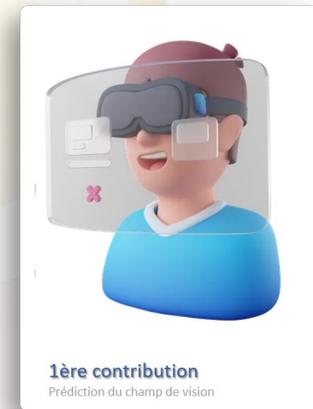
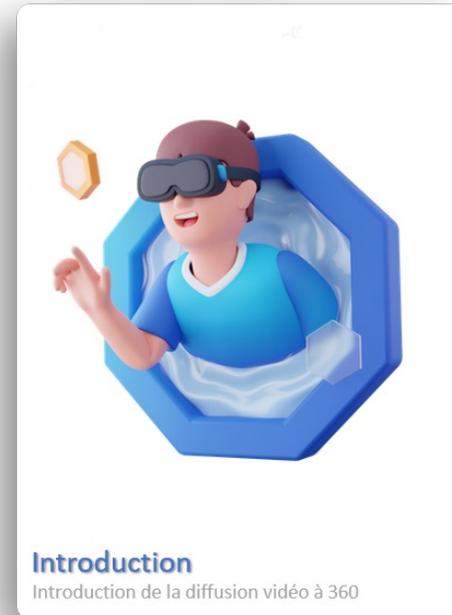
- Horizon Europe 6G-TWIN
- Horizon Europe OPEVA
- ANR-FNR 5G-INSIGHT



Plan de la présentation

Introduction

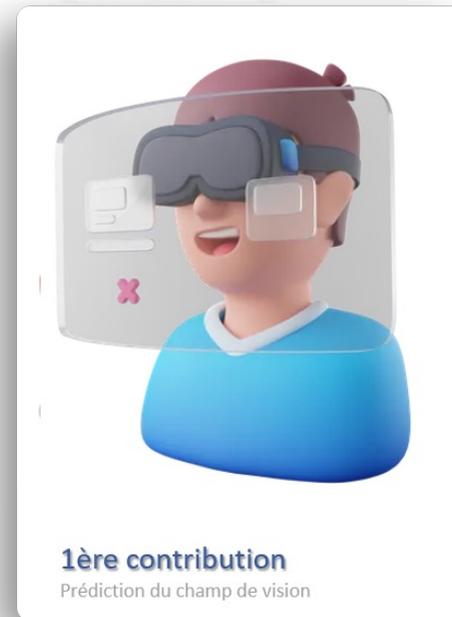
Introduction à la diffusion de la vidéo 360° et aux domaines d'application.



Plan de la présentation

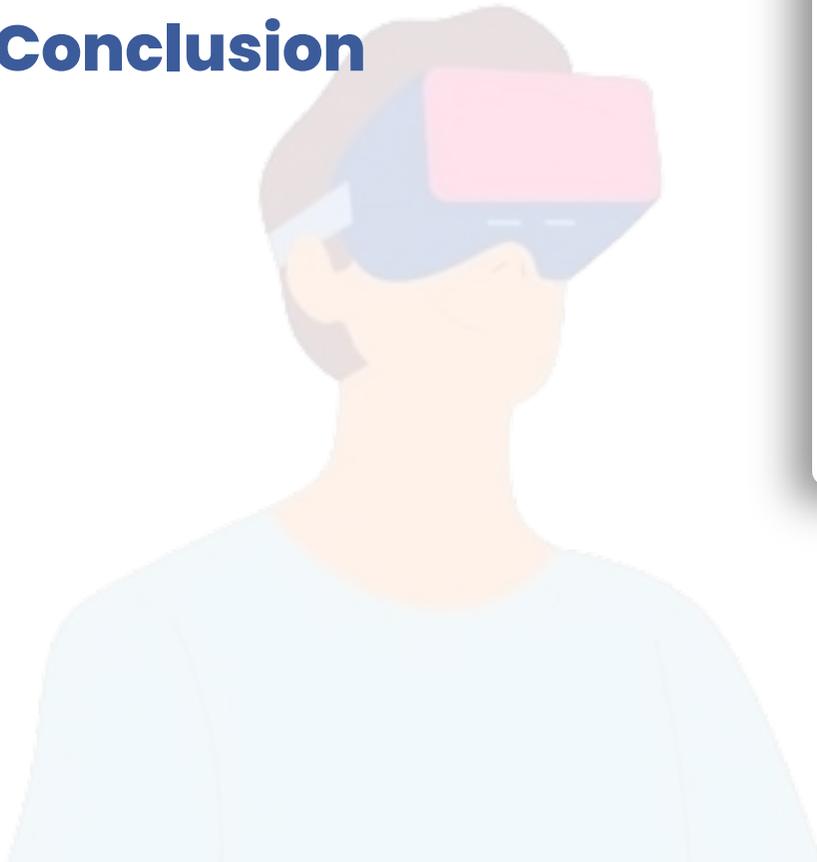
Contribution

Classification au service de La prédiction du champ de vision dans les vidéos à 360°



Plan de la présentation

Conclusion



Conclusion

Conclusion, perspectives et future directions de recherche.



Introduction

Contribution

Conclusion

La vidéo à 360°

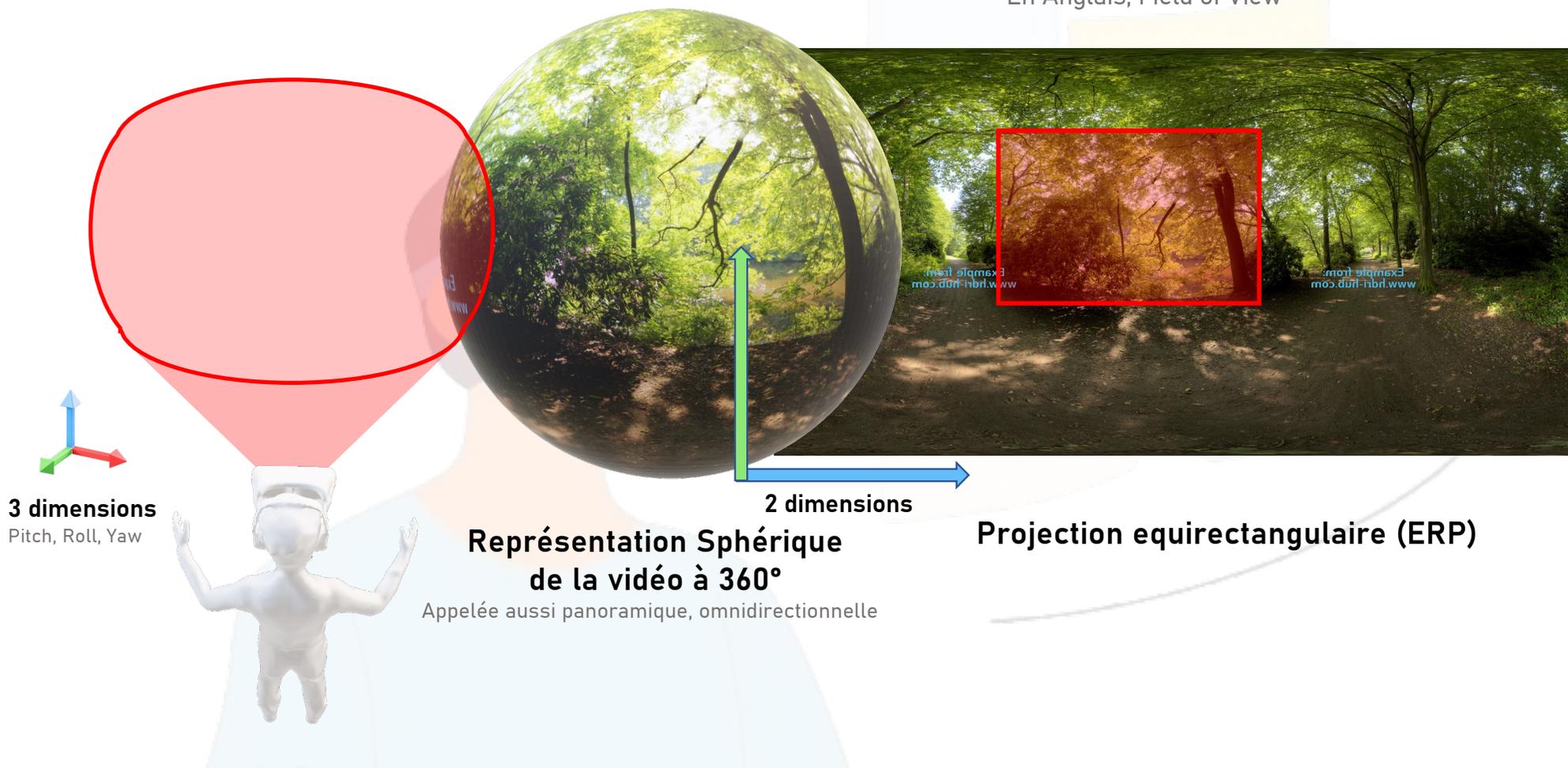
Domaines d'application

Types de diffusion

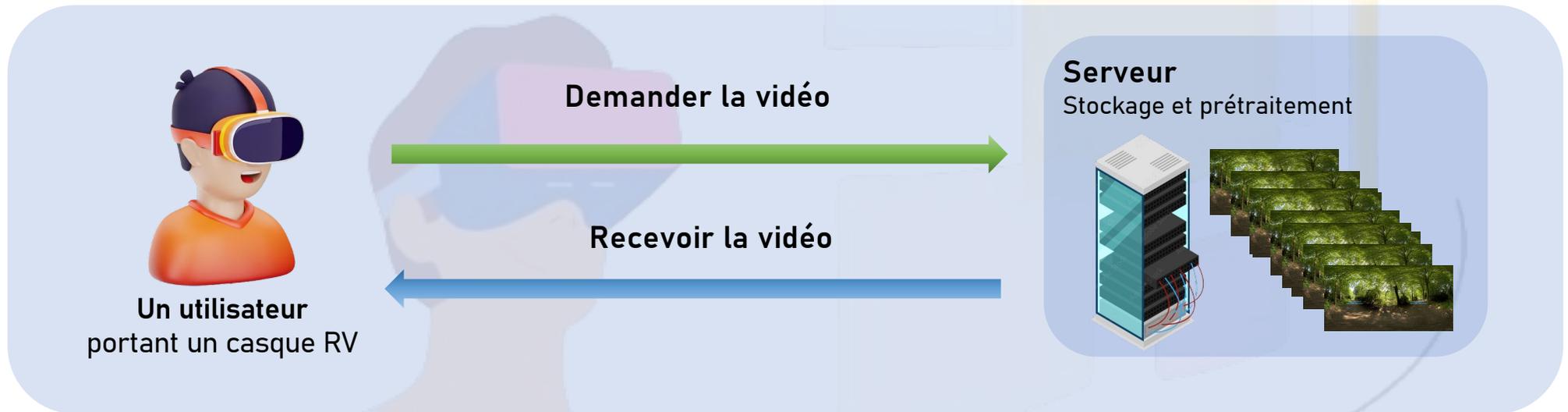
Objectifs et motivations

Champ de vision (FOV)

En Anglais, Field of View



La diffusion de la vidéo 360° (Architecture de base)



Introduction

La vidéo à 360°

Contribution

Domaines d'application

Conclusion

Types de diffusion

Objectifs et motivations

La diffusion de la vidéo 360°

Demander la vidéo

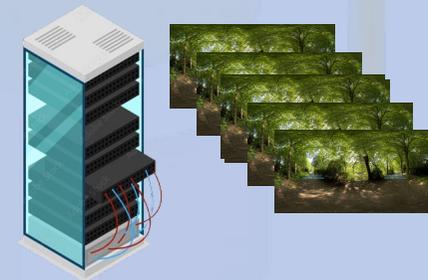
Recevoir la vidéo

Serveur

Stockage et prétraitement



Un utilisateur
portant un casque RV



Introduction

La vidéo à 360°

Contribution

Domaines d'application

Conclusion

Types de diffusion

Objectifs et motivations

La diffusion de la vidéo 360°

Demander la vidéo

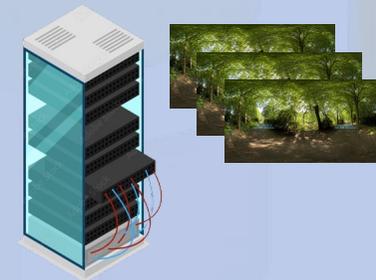
Recevoir la vidéo

Serveur

Stockage et prétraitement



Un utilisateur
portant un casque RV



Introduction

La vidéo à 360°

Contribution

Domaines d'application

Conclusion

Types de diffusion

Objectifs et motivations

La diffusion de la vidéo 360°

Demander la vidéo

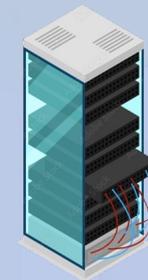
Recevoir la vidéo

Serveur

Stockage et prétraitement



Un utilisateur
portant un casque RV



Immobilier

- Visites virtuelles
 - Visualisation réaliste des propriétés



Événements sportifs

- Assister à des matchs de football comme dans un stade
- période de restrictions de déplacement



Tourisme

- Visites virtuelles
 - Destinations touristiques
 - Musées
- Prévisualisation immersive avant visite.



Marketing et publicité

Des campagnes publicitaires innovantes pour explorer un produit



Éducation

- Visites virtuelles sans quitter la classe
 - Lieux historiques
 - Écosystèmes naturels
- Simulations de formation des professionnels





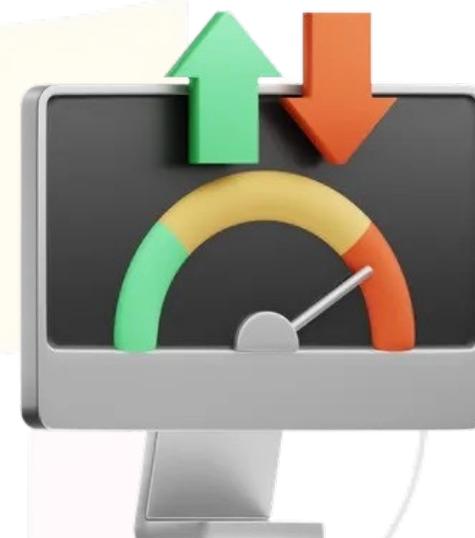
La diffusion en direct

Les utilisateurs visionnent des événements en temps réel, offrant une expérience immersive comme s'ils y assistaient en personne.



La diffusion à la demande

Les utilisateurs peuvent visionner des vidéos 360° à tout moment à partir d'une bibliothèque de vidéos.



Nécessite une **bande passante élevée**



Réduire la taille de la vidéo à transmettre

La prédiction du FOV est essentielle pour optimiser le streaming et améliorer la qualité de l'expérience utilisateur.



Assurer la mise à l'échelle.

Assurer la mise à l'échelle (*scalability*) de la diffusion de la vidéo à 360° dans le contexte d'une diffusion multicast

Introduction

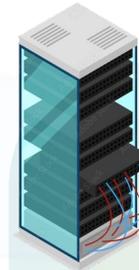
Contribution

Conclusion

Problématique

Prédiction du FOV

Résultats



Haute
Qualité



Faible
Qualité

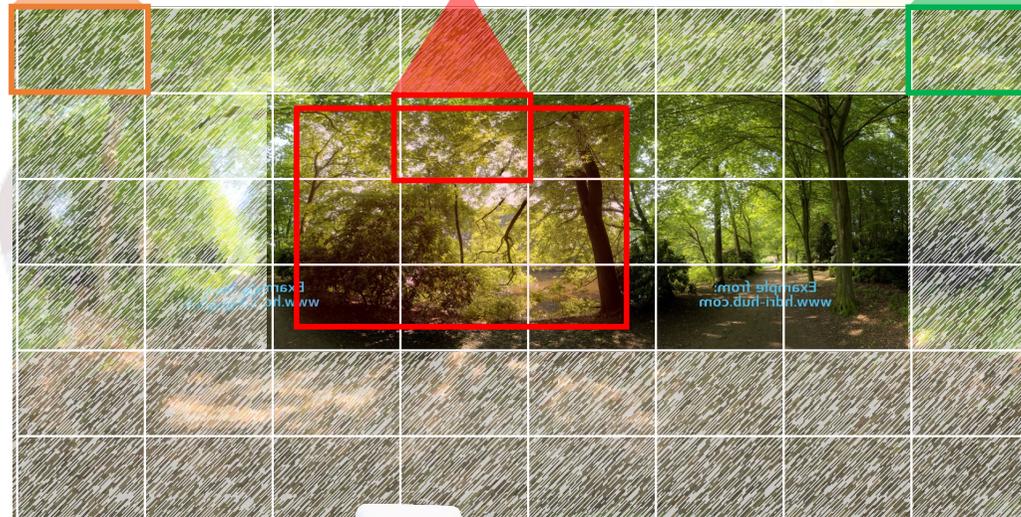


Consommation de bande passante élevée
Qualité plus faible si la bande passante est faible → QoE
Taille d'images plus importantes

Faible qualité
Hors FOV

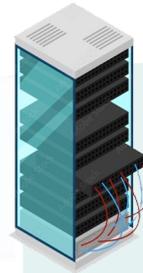
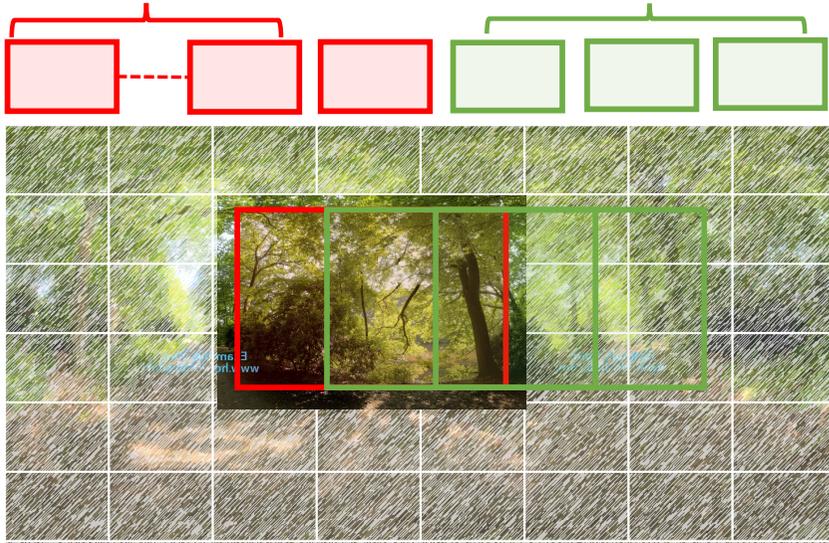
Haute qualité
FOV

Tuile d'image
En Anglais, Tile



FOVs Précédents

Futures FOVs



Consommation de bande passante plus faible
 Bonne qualité du FOV même si la bande passante est faible
 Taille d'images moins importantes



Méthode de prédiction
de champ de vision
FOV
(Avec tuilage)



Réseaux neuronaux

Extrêmement efficaces mais **très lents** en termes de **temps d'apprentissage** et **d'inférence** (prédiction).



Méthodes autorégressive (ex. ARIMA*)

Rapides mais doivent être **régulièrement entraînées** pour mieux comprendre les données, ce qui devient **lent** si la taille de la série temporelle augmente pour un flux en direct.



Prévision de séries temporelles à l'aide de la classification

Rapide à la fois en termes de **temps d'inférences** et **d'entraînements**.

*AutoRegressive Integrated Moving Average

Introduction

Contribution

Conclusion

Problématique

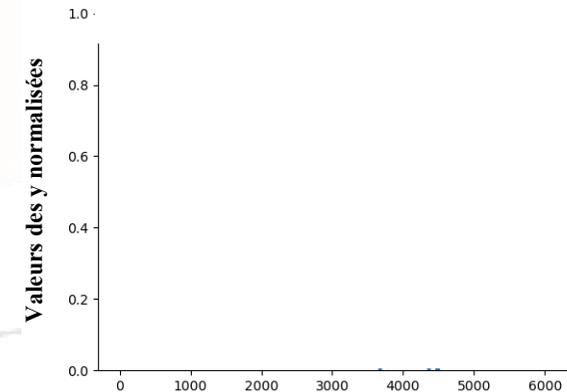
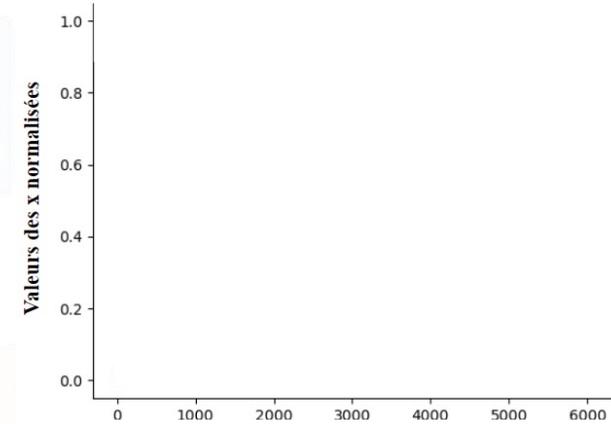
Prédiction du FOV

Résultats

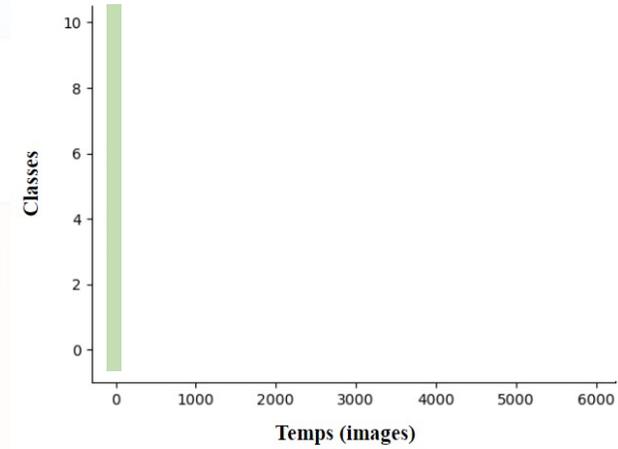
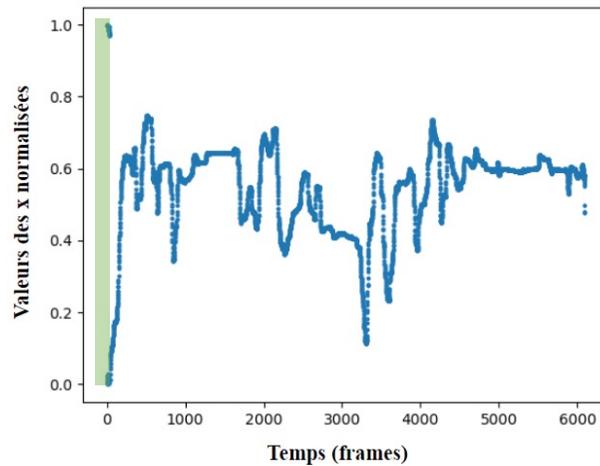


Projection ERP

Projeter les données de mouvements en 2 dimensions



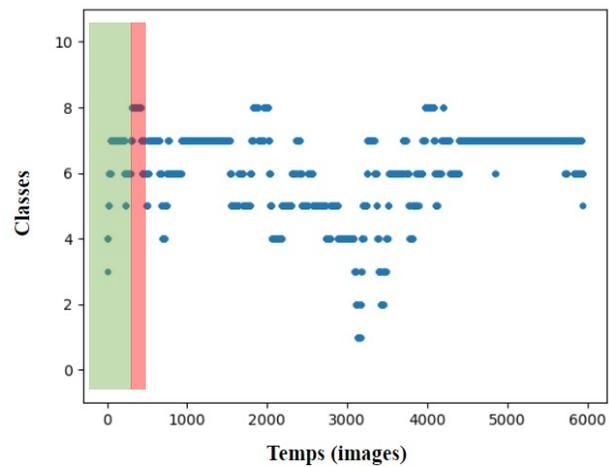
Séries temporelles



Discrétisation

Diviser les valeurs continues en intervalles ou catégories équivalentes.

$$X \rightarrow \{C_1, C_2, \dots, C_k\}$$



Fenêtre de taille W

Cible

t1	t2	t3	t4	t5	Target
6	6	6	5	5	5
6	6	5	5	5	5
		.			
3	3	3	3	2	2
		.			
2	2	3	4	5	6
2	3	4	5	6	6
3	4	5	6	6	7

Construction du jeu de données

En fixant une taille de fenêtre W représentant les W derniers champ de vision (historique) et une cible représentant le prochain champ de vision.

t1	t2	t3	t4	t5	Target
6	6	6	5	5	5
6	6	5	5	5	5
		.			
3	3	3	3	2	2
		.			
2	2	3	4	5	6
2	3	4	5	6	6
3	4	5	6	6	7

Quelle modèle d'apprentissage?

Naïve Bayes

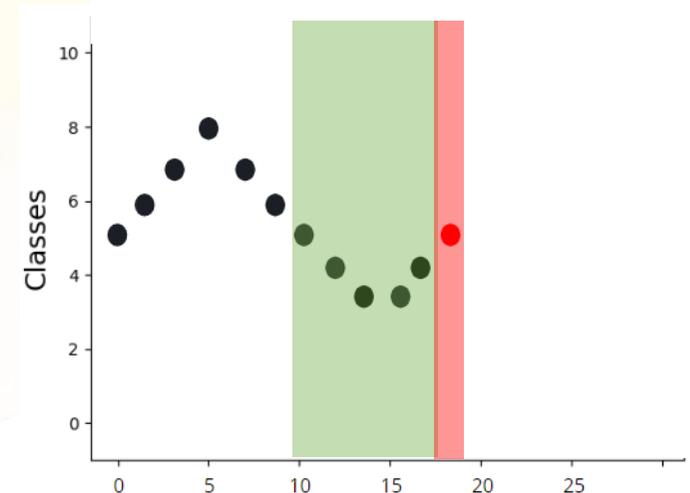
Repose sur le théorème de Bayes, défini comme suit :

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)}$$

$$P(X|C_i) = \prod_{j=1}^n P(x_j|C_i)$$

Entraîner le modèle

Sur le jeu de données



Prédire les futures FOV

En se basant sur les précédentes FOVs

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

Supervisé

Classification

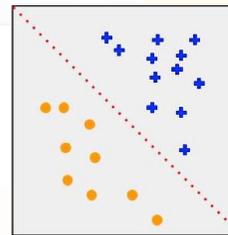
✓ La sortie est **discrète** pour une catégorie ou une classe à partir de données (exp., voiture/camion, filtrage du courrier électronique en spam ou pas)

Non Supervisé

Régression

✓ La sortie de la prediction est une valeur **continue** (exp., prix, vitesse, consommation, etc.)

Par Renforcement

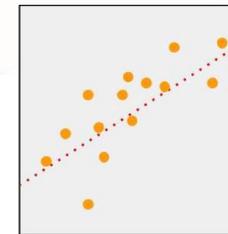


Les classifications

Lorsqu'on cherche à prédire une classe :

Chien Chat Chien Chien Chat

Exemples : un animal, un état de santé, une couleur, un type de véhicule, ...



Les régressions

Lorsqu'on cherche à prédire une quantité :

Exemples : prix d'une maison, durée d'un trajet, température, poids, ...

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

- Méthode de **classification** basée sur le **théorème de Bayes**.

Théorème de Bayes

Permet de calculer la probabilité d'une classe donnée C (par exemple, un spam ou non-spam) en fonction de l'observation de certaines caractéristiques X (comme des mots dans un email) et s'exprime ainsi :

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Où :

- $P(C|X)$: probabilité que la classe C soit vraie, sachant X (les données observées)
- $P(X|C)$: probabilité d'observer X, sachant C
- $P(C)$: probabilité a priori de la classe C
- $P(X)$: probabilité a priori d'observer X

- Appelé "naïf" car il fait l'hypothèse que les caractéristiques utilisées pour la classification sont **indépendantes** les unes des autres (rarement vrai) mais permet de simplifier les calculs

Hypothèse d'indépendance

Toutes les caractéristiques (ou attributs) X_1, X_2, \dots, X_n sont indépendantes les unes des autres :

$$P(X|C) = P(X_1|C) \cdot P(X_2|C) \cdots P(X_n|C)$$

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

- Question 1 : Est-ce qu'on peut utiliser Naïve Bayes avec un jeu de données pour prédire une maladie spécifique (classe) en fonction des attributs suivants?
 1. Tension artérielle (Pression systolique et diastolique)
 2. Taux de cholestérol
 3. Indice de masse corporelle (IMC)
 4. Antécédents familiaux de maladies cardiaques
- Question 2 : Est-ce qu'on peut utiliser Naïve Bayes avec un jeu de données pour véhicule autonome en fonction des attributs suivants?
 1. Vitesse du véhicule
 2. Accélération
 3. Pente de la route (inclinaison)
 4. Régime moteur (tr/min)
 5. Consommation instantanée de carburant (l/100 km)

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

- Exemple : Nous voulons calculer la probabilité qu'une voiture ait un **problème de moteur (A)** sachant qu'elle fait un **bruit étrange (B)**, $P(A|B)$:

- **P(A)** : La probabilité qu'une voiture ait un problème de moteur, disons 10% ($P(A)=0.1$).
- **P(B)** : La probabilité qu'une voiture fasse un bruit étrange en général, disons 5% ($P(B)=0.05$).
- **P(B | A)** : La probabilité qu'une voiture fasse un bruit étrange si elle a un problème de moteur, disons 80% ($P(B|A)=0.8$).

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} = \frac{0.8 \times 0.1}{0.05} = 1.6 = 160\%.$$

Intuition?

- Intuition du théorème
 - Le théorème de Bayes permet d'ajuster une probabilité initiale à la lumière de nouvelles informations.
 - Ici, bien qu'une voiture ait seulement 10% de chances d'avoir un problème de moteur a priori, cette probabilité monte à 160% si elle commence à faire un bruit étrange, car le bruit étrange est plus probable en cas de problème moteur.

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

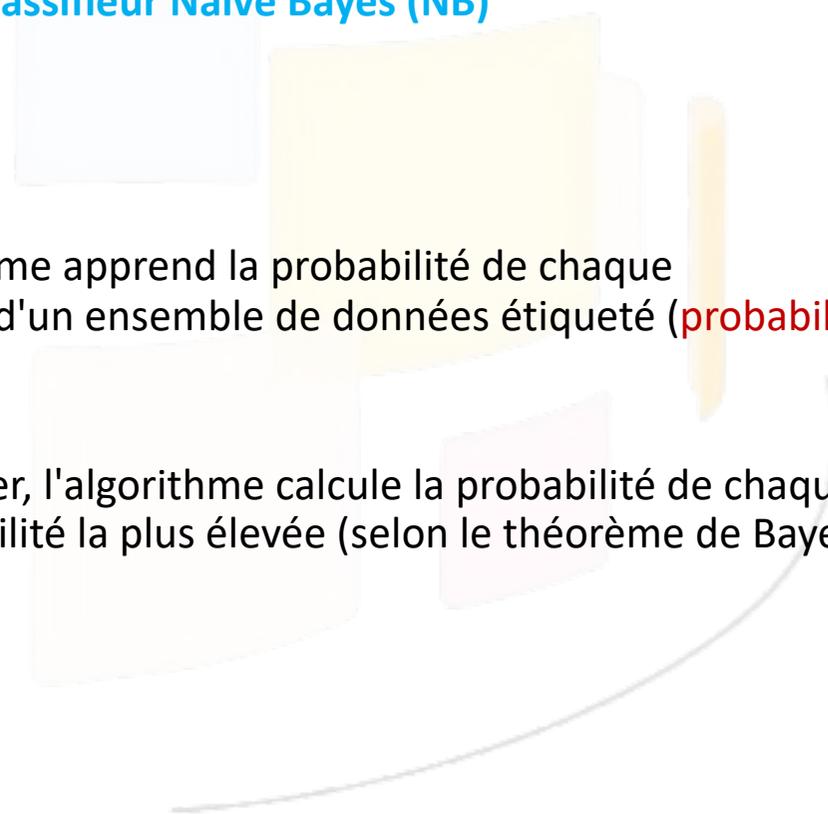
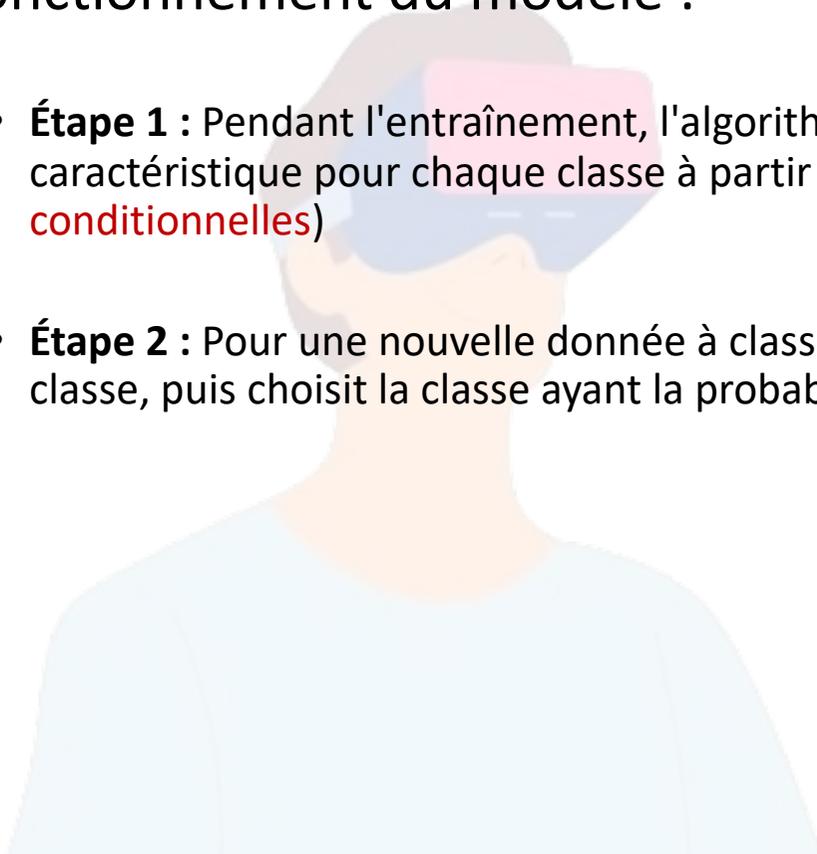
• Domaines d'application

- Analyse des sentiments dans les **avis clients** ou les critiques de films classant les textes comme positifs, négatifs ou neutres.
- **Catégorisation de documents** en fonction de leurs contenus dans des catégories spécifiques (ex : sport, politique, technologie).
- **Filter les Email** (Spam) en analysant le contenu des emails (contient des mots comme "argent", "offre spéciale", "cliquez ici", ...)
- Systèmes de **recommandation** pour recommander des **produits ou du contenu** en fonction des préférences de l'utilisateur (**suggérer des films similaires** à ce que vous regardez - **estimer la probabilité d'appréciation d'un nouveau film**)
-

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

- **Fonctionnement du modèle :**

- **Étape 1 :** Pendant l'entraînement, l'algorithme apprend la probabilité de chaque caractéristique pour chaque classe à partir d'un ensemble de données étiqueté (**probabilités conditionnelles**)
- **Étape 2 :** Pour une nouvelle donnée à classer, l'algorithme calcule la probabilité de chaque classe, puis choisit la classe ayant la probabilité la plus élevée (selon le théorème de Bayes)



Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

- Échantillon de 10 objets observés selon les attributs :
 - Couleur : **Rouge(R)**, **Jaune(J)**
 - Forme : Rond(**Ro**) et Cylindrique(**Cy**)
 - Diamètre : $\geq 10\text{cm}$ (**S10**) et $< 10\text{cm}$ (**I10**)
- La sortie représentée par 2 classes :
 - C_1 : Oui, c'est une pomme
 - C_2 : Non, ce n'est pas une pomme

Objets	Couleur	Forme	Diamètre	Pomme ?
1	R	Ro	$\geq 10\text{cm}$	Oui
2	R	Ro	$\geq 10\text{cm}$	Non
3	R	Ro	$\geq 10\text{cm}$	Oui
4	J	Ro	$\geq 10\text{cm}$	Non
5	J	Ro	$< 10\text{cm}$	Oui
6	J	Cy	$< 10\text{cm}$	Non
7	J	Cy	$< 10\text{cm}$	Oui
8	J	Cy	$\geq 10\text{cm}$	Non
9	R	Cy	$< 10\text{cm}$	Non
10	R	Ro	$< 10\text{cm}$	Oui

On considère un nouvel objet $X = (R, Ro, S10)$, est ce que X est une pomme ?

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Pour y répondre :

- Étape 1** : calcul des probabilités conditionnelles
- Étape 2** : calcul de la probabilité de chaque classe $P(C_1/X)$ et $P(C_2/X)$

Dis-moi un peu plus sur le Classifieur Naïve Bayes (NB)

On considère un nouvel objet $X = (R, Ro, S10)$, est ce que X est une pomme ?

Pour y répondre il faudrait calculer $P(C_1/X)$ et $P(C_2/X)$:

$$P(C_1/X) = (P(C_1) * P(R/C_1) * P(Ro/C_1) * P(S10/C_1)) / P(X)$$

$$P(C_2/X) = (P(C_2) * P(R/C_2) * P(Ro/C_2) * P(S10/C_2)) / P(X)$$

Étape 2 : calcul de la probabilité de chaque classe

$$P(C_1/X) = 60/750$$

$$P(C_2/X) = 30/750$$

on a $P(C_1/X) > P(C_2/X)$ donc X est une pomme.

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Étape 1: apprendre probabilités conditionnelles

$$P(C_1) = 5/10$$

$$P(C_2) = 5/10$$

$$P(R/C_1) = 3/5$$

$$P(J/C_1) = 2/5$$

$$P(R/C_2) = 2/5$$

$$P(J/C_2) = 3/5$$

$$P(Ro/C_1) = 2/3$$

$$P(Cy/C_1) = 1/4$$

$$P(Ro/C_2) = 1/3$$

$$P(Cy/C_2) = 3/4$$

$$P(S10/C_1) = 2/5$$

$$P(I10/C_1) = 3/5$$

$$P(S10/C_2) = 3/5$$

$$P(I10/C_2) = 2/5$$

$$P(X) = P(R) * P(Ro) * P(S10) = 150/1000$$

Objets	Couleur	Forme	Diametre	Pomme ?
1	R	Ro	≥10cm	Oui
2	R	Ro	≥10cm	Non
3	R	Ro	≥10cm	Oui
4	J	Ro	≥10cm	Non
5	J	Ro	<10cm	Oui
6	J	Cy	<10cm	Non
7	J	Cy	<10cm	Oui
8	J	Cy	≥10cm	Non
9	R	Cy	<10cm	Non
10	R	Ro	<10cm	Oui

DataSeT d'enregistrements des mouvements de tête de **48 volontaires** regardant **18 vidéos** 360° différentes.

Vidéos divisées spatialement en **12 × 10 tuiles** et temporellement en séquences d'**une seconde**.



# Vidéo	Durée	Contenu	Catégorie
1	2:44	Conan360 °-Sandwich	Performance
2	2:52	Conan360 °-Weird AI	Performance
3	3:21	Freestyle Skiing	Sport
4	3:25	GoPro VR-Tahiti Surf	Sport
5	2:44	LOSC Football	Sport
6	6:01	Women's basketball match	Sport



Figure 1: Exp. 1 Video 2



Figure 2: Exp. 1 Video 3

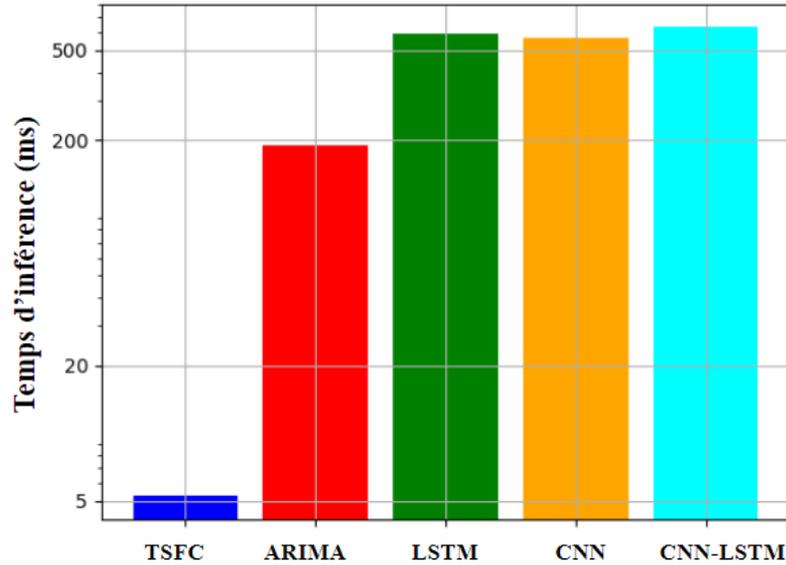


Figure 3: Exp. 2 Video 4

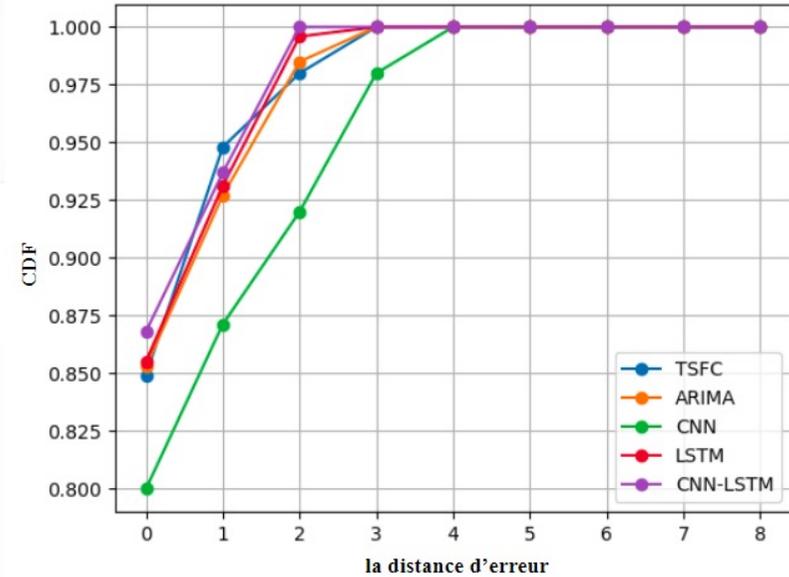


Figure 4: Exp. 2 Video 9

Chenglei Wu et al. "A dataset for exploring user behaviors in VR spherical video streaming". In : Proceedings of the 8th ACM on Multimedia Systems Conference. 2017, p. 193-198.



Comparaison de **temps d'inférence** pour prédire une fenêtre de 30 valeurs (W) pour les différents algorithmes.



Comparaison de la fonction de distribution cumulative (CDF) des prédictions correctes avec une marge d'erreur (distance d'erreur) pour les différents algorithmes.

Model	TSFC	CNN	LSTM	CNN-LSTM
Training	2s132ms	15min49s	46min32s	55min07s

Comparaison du **temps d'entraînement** pour les différents algorithmes (entraînement avec 48 utilisateurs).



Conclusion



Importance de la prédiction du champ de vision (FOV)

La prédiction du champ de vision est cruciale pour améliorer l'expérience de l'utilisateur dans les applications de streaming vidéo à 360°.



Avantages de la classification

Les techniques de classification peuvent améliorer à la fois le **temps d'apprentissage** et le **temps d'inférence** pour les tâches de prédiction du champ de vision.



Classification pour les problèmes de grandes séries temporelles

Dans les scénarios avec de grandes séries de données temporelles, **les algorithmes de classification rapide** peuvent être un choix efficace pour la prédiction de FOV

01 VTC2023

A. Saadallah, I. El-Korbi, S. M. Senouci and P. Brunet, "A New Time Series Forecasting Approach Using Classification: Application to Field of View Prediction in 360° videos," VTC2023-Spring, Florence, Italy, 2023, pp. 1-5, doi: [10.1109/VTC2023-Spring57618.2023.10199671](https://doi.org/10.1109/VTC2023-Spring57618.2023.10199671).

02 GLOBECOM23

A. Saadallah, P. Brunet, I. El-Korbi, S. -M. Senouci and S. Cherkaoui, "Leveraging Graph Theory for Efficient Cache Policy Design in 360° Video Streaming," GLOBECOM 2023, Kuala Lumpur, Malaysia, 2023, pp. 976-981, doi: [10.1109/GLOBECOM54140.2023.10436720](https://doi.org/10.1109/GLOBECOM54140.2023.10436720)

03 GLOBECOM24

A. Saadallah, S. -M. Senouci, I. El-Korbi and P. Brunet , "Dynamic Field-of-View-Based Clustering for Efficient 360-degree Multicast Streaming," GLOBECOM 2024 - 2024 IEEE Global Communications Conference.

04 GLOBECOM24

W. Abdallaoui, A. Saadallah, S. -M. Senouci, I. El-Korbi and P. Brunet , "Modeling Users' Behavior in 360-Degree Videos: A Framework for Synthetic Data Generation" GLOBECOM 2024 - 2024 IEEE Global Communications Conference.

05 Soumission journal

A. Saadallah, S. -M. Senouci, I. El-Korbi and P. Brunet , "A Graph-Based Caching Approach for Optimizing 360° Video Streaming Using Edge Caching and Field-of-View Prediction" (Soumis à un journal)



**Merci Pour
votre attention**

TSFC vs. Autres algo?

- LSTM (*Long Short Term Memory*) : pour la prédiction à plusieurs étapes, un réseau LSTM bidirectionnel prendra en compte à la fois les informations directes et inverses de l'entrée.
- CNN (*Convolutional Neural Network*) : deux couches de convolution séquentielles à une dimension sont utilisées pour filtrer les informations, et une couche de Max pooling est ajoutée par la suite. Cette dernière est aplatie avant d'être reliée par une couche entièrement connectée.
- CNN-LSTM : deux couches CNN et une couche de Max pooling intermédiaire. Après aplatissage des couches CNN, une cellule LSTM est reliée à une couche entièrement connectée.
- ARIMA (AutoRegressive Integrated Moving Average) : modèle de prévision qui combine des composants autorégressifs et de moyenne mobile. Nous avons fixé les paramètres à $(2,1,1)$ pour l'axe des X et $(3,1,0)$ pour l'axe des Y.