

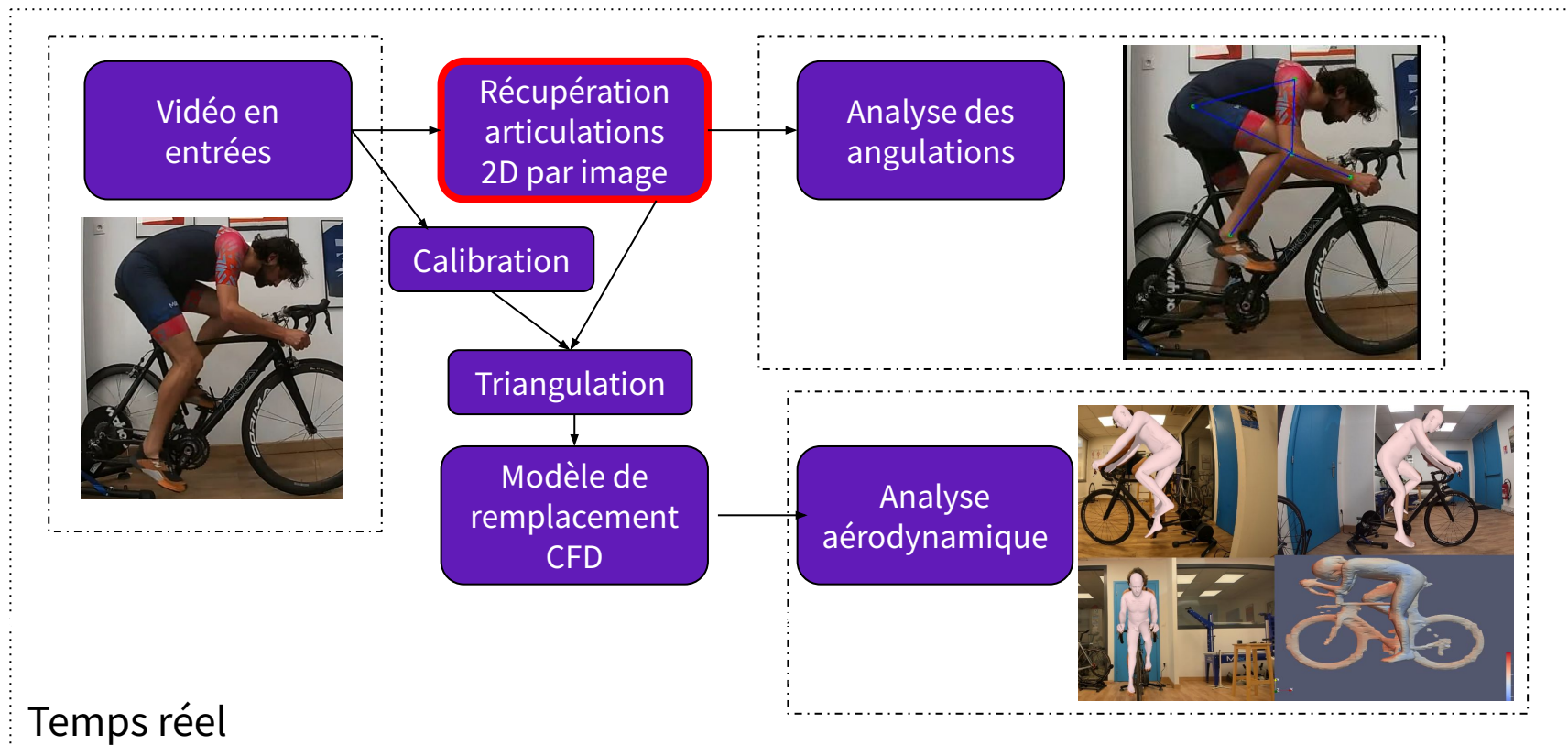
Benchmark temps réel des modèles de reconstruction 2D pour le cyclisme stationnaire

02/07/2026 GTAS 2026

The logo for GTAS, with the letters "G", "T", "A", and "S" in a bold, sans-serif font. The "G" is green, "T" is red, "A" is blue, and "S" is purple.

Pierre Rivière^{1,2}, Florence Zara², Nicolas Pronost²

¹ MorphoLogics SAS, France, ² Université Lyon 1, INSA Lyon, CNRS, LIRIS



Notre besoin ?

Bonne **reconstruction 3D**
=
Très bonne **récupération 2D**

The accuracy of several pose estimation methods for 3D joint centre localisation *oct 2021*

Laurie Needham¹, Murray Evans, Darren P. Cosker, Logan Wade, Polly M. McGuigan, James L. Bilzon & Steffi L. Colyer

Overview of 3D Human Pose Estimation *2023*

Jianchu Lin^{1,2}, Shuang Li³, Hong Qin^{3,4}, Hongchang Wang³, Ning Cui⁶, Qian Jiang⁷, Haifang Jian^{1,4} and Gongming Wang^{5,7}

Pose2Sim: An End-to-End Workflow for 3D Markerless Sports Kinematics—Part 1: Robustness *sept 2021*

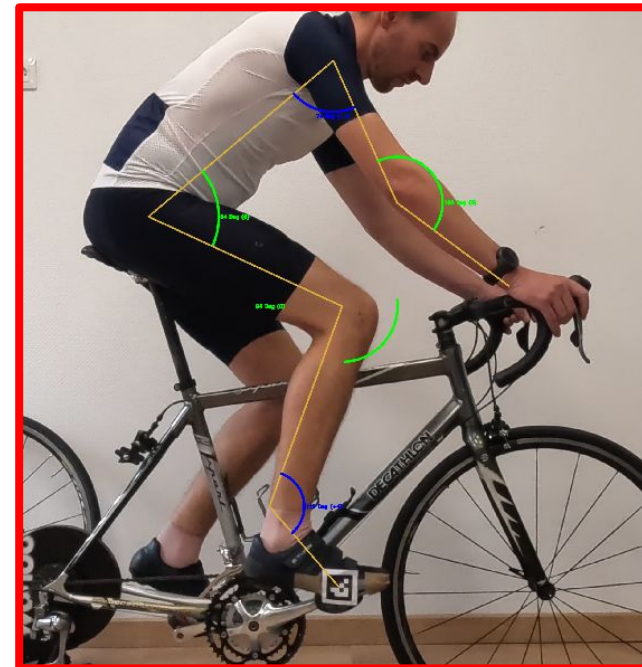
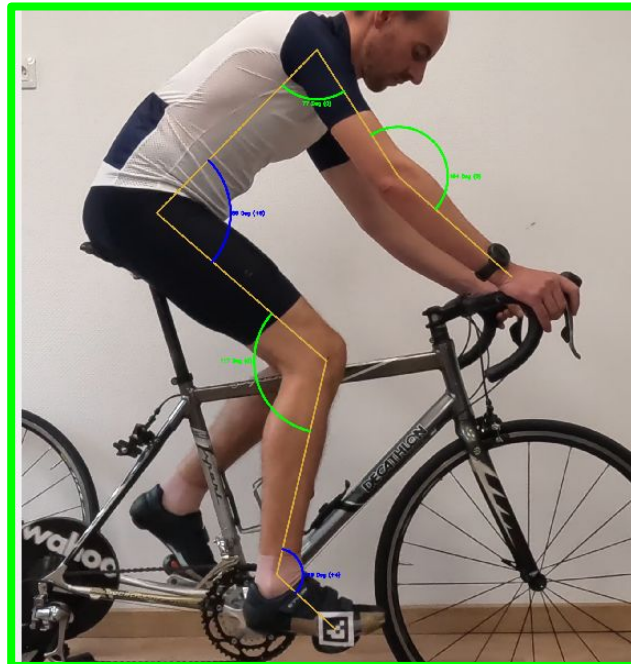
David Pagnon^{1,4}, Mathieu Domalain² and Lionel Reveret^{1,3}

Assessment of monocular human pose estimation models for clinical movement analysis *nov 2025*

David Rode^{1,2}, Annika Dunkel³, Romina Willi³, Peter Wolf³, Michele Xiloyannis² & Robert Riener^{1,4}

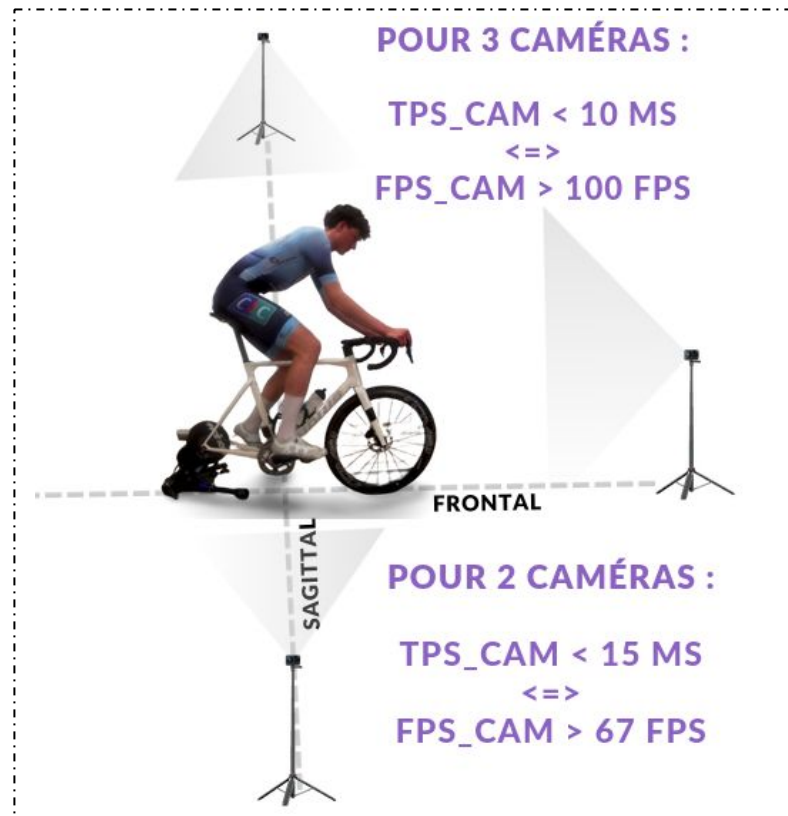
Notre besoin ?

Besoin d'une
récupération **fiable**
et **répétable**
(éviter les
incohérences)



Notre besoin ?

Analyse en **temps réel**



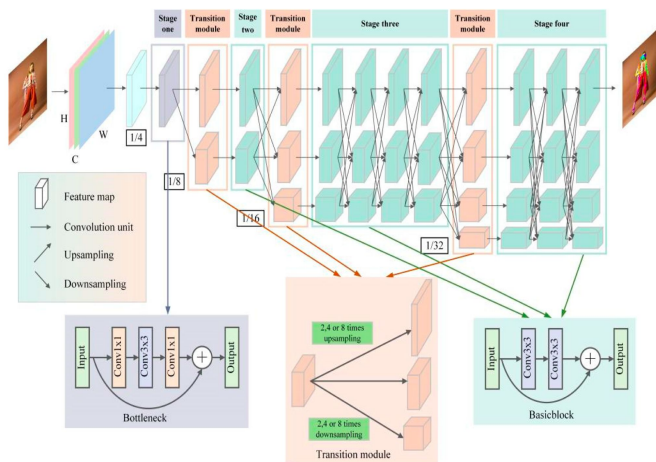
Modèles de reconstruction ?

Pierre Rivière^{1,2}, Florence Zara², Nicolas Pronost²

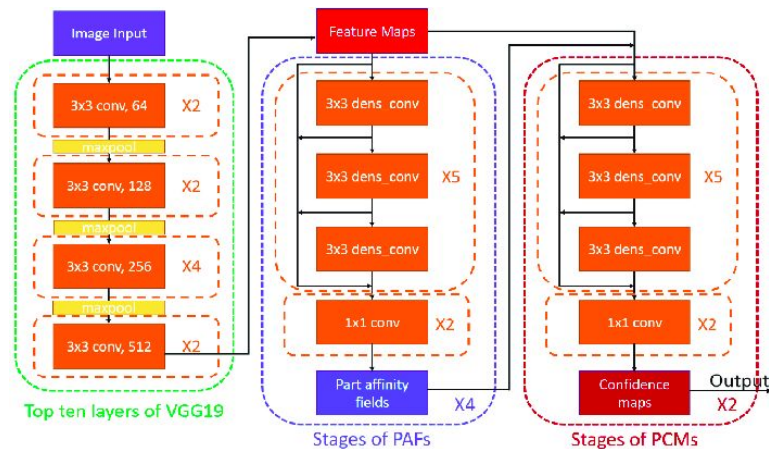
¹ MorphoLogics SAS, France, ² Université Lyon 1, INSA Lyon, CNRS, LIRIS

Familles et architectures de réseaux

Haute résolution maintenue (HRNet) - high-resolution representation

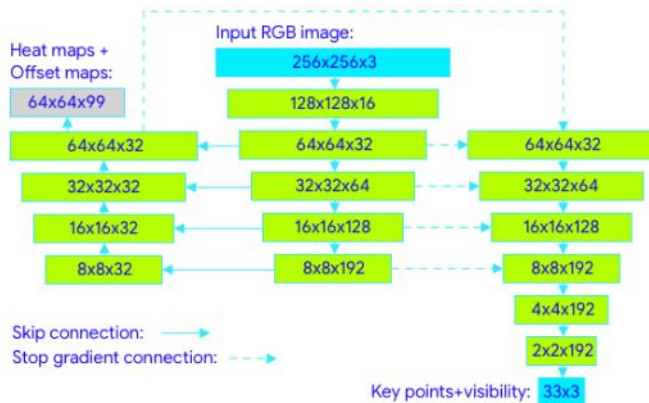


Bottom-up par champs d'affinité (OpenPose) - part affinity fields

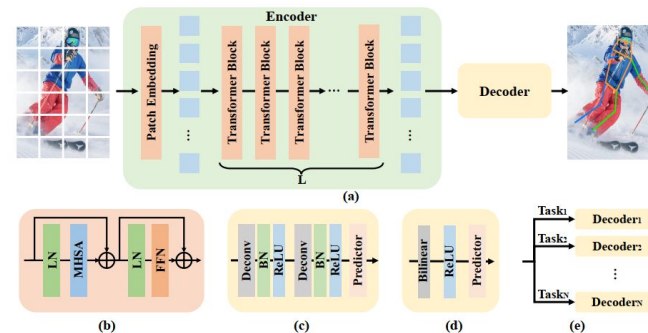


Familles et architectures de réseaux

Régression directe légère / mobile - lightweight direct regression



Transformers pour la pose - vision transformer



BlazePose: On-device Real-time Body Pose tracking

Valentin Bazarevsky Ivan Grishchenko Karthik Raveendran Tyler Zhu Fan Zhang Matthias Grundmann, arXiv 2017 & CVPR 2020

ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation
Yufei Xu, Jing Zhang, Qiming Zhang, Dacheng Tao, arXiv 2022 & NeurIPS 2022

Classiques 2D	HRNet	Léger / mobile	YOLO / one-stage	Transformers	Multi-personn es	3D / vidéo
OpenPose	HRNet	MediaPipe Pose	YOLOv7-Pose	ViTPose	AlphaPose	VideoPose3D
CMU-Pose	HigherHRNet	BlazePose	YOLOv8-Pose	ViTPose+	PolarPose	MotionBERT
DeepPose	Lite-HRNet	MoveNet	YOLOv11-Pose	TokenPose	DWPose	PoseFormer
Conv. Pose Machines	Dite-HRNet	PoseNet	YOLO-NAS Pose	TransPose	MegviiPose	DeepLabCut
Stacked Hourglass	X-HRNet	Lightweight OpenPose	CenterNet Pose	PCT		SLEAP
DeepCut	UDP-Pose	TF-Pose	PifPaf	PETR		Martinez et al. baseline
DeeperCut	MSPN	EfficientPose	OpenPifPaf	ED-Pose		
SimpleBaseline		MobilePose		GroupPose		
Mask R-CNN Keypoint		RTMPose		PRTR		
DensePose		RTMPose-Lite		RLE-Pose		

Modèles inclus dans le benchmark (en jaune)

Métriques choisies pour le benchmark

Rapidité

Ressources
CPU/GPU

Répétabilité
(200 itérations)

Précision

Articulations

épaule

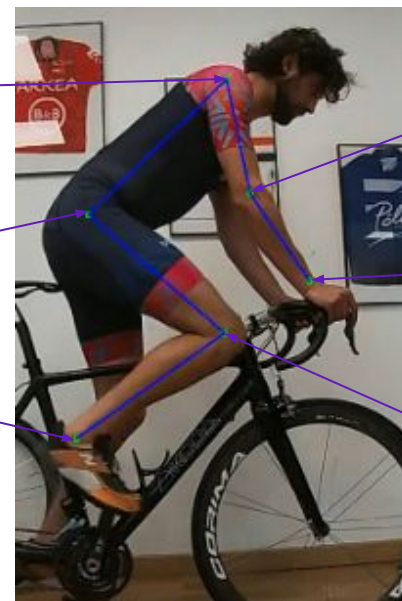
hanche

cheville

coude

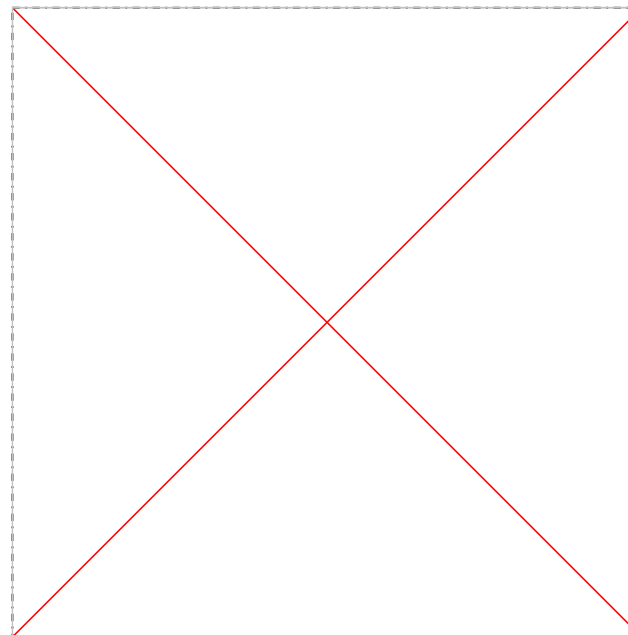
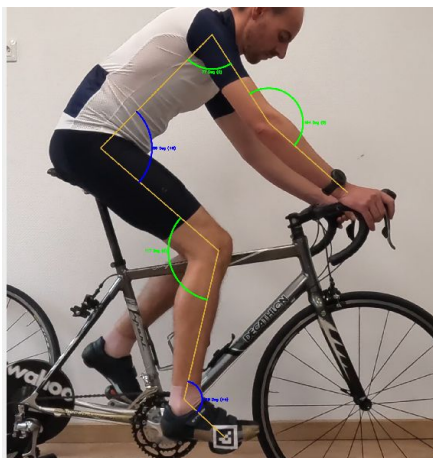
poignet

genou



Ground Truth ?

Comparaison des modèles avec
OpenPose



Résultats : rapidité (temps d'inférence)

model	mean_ms	fps_mean
_rtmpose_aic_coco_t_256 participant_x	11.6847573441886	85.5815803909138
_rtmpose_aic_coco_s_256 participant_x	12.388353972095	80.7209740900621
_rtmpose_aic_coco_t_256 participant_x	13.6624242250089	73.1934526062739
_simplebaseline_coco_aic_50_256 participant_x	14.0189502573133	71.3320171371837
_rtmpose_aic_coco_s_256 participant_x	14.0807961023175	71.0187117783355
_rtmpose_aic_coco_t_256 participant_x	14.260887159515	70.1218647069086
_rtmpose_aic_coco_t_256 participant_x	14.400766404282	69.4407486328403
_rtmpose_aic_coco_t_256 participant_x	14.4481576956155	69.2129765654111
_rtmpose_aic_coco_s_256 participant_x	14.5069330293956	68.9325578310512
_rtmpose_aic_coco_t_256 participant_x	14.509400058705	68.920837247164

Résultats : répétabilité (écart-type intra)

model	global kp_dist	rank_globa	rank_shoulde	rank_elbow	rank_hip	rank_knee	rank_ankle
_vitpose_coco_s_256 participant_x	49.1973068612619	2	194	196	126	7	4
_vitpose_coco_b_256 participant_x	56.4562875229718	4	122	185	76	11	10
_hrnetlite_coco_18_384 participant_x	57.1863495696395	5	217	133	183	4	44
_vitpose_coco_b_256 participant_x	61.5052364755577	6	94	94	64	17	26
_rtmpose_aic_coco_s_256 participant_x	64.076825701775	7	123	246	153	6	30
_rtmpose_aic_coco_s_256 participant_x	65.0144714124102	8	125	207	65	163	2
_simplebaseline_coco_152_256 participant_x	65.5764427634779	9	205	161	220	23	7
_rtmpose_aic_coco_s_256 participant_x	66.7808922724646	10	152	214	253	10	17
_hrnetlite_mpii_18_256 participant_x	67.401159524268	11	278	258	271	5	3
_vitpose_coco_s_256 participant_x	68.3667075045651	12	167	213	209	19	15

Résultats : précision (écart-type à OpenPose)

model	global kp_error	rank_globa	rank_shoulder	rank_elbow	rank_hip	rank_knee	rank_ankle
rtmpose_aic_coco_m_256 [participant_x]	14.933379109834	1	1	15	8	19	14
simplebaseline_coco_152_384 [participant_x]	15.362901601607	2	2	8	10	13	27
vitpose_coco_h*_256 [participant_x]	15.774814029084	3	43	17	3	6	5
rtmpose_aic_coco_l_256 [participant_x]	15.927182067217	4	27	21	9	12	4
rtmpose_cocktail14_x_384 [participant_x]	16.275651714206	5	33	1	13	3	23
vitpose_coco_h_256 [participant_x]	16.466015393985	6	62	16	7	7	3
hrnet_coco_w48_384 [participant_x]	16.538892130113	7	4	45	17	4	20
rtmpose_cocktail14_l_384 [participant_x]	16.539600955729	8	23	10	18	2	29
hrnet_coco_wholebody_w48 [participant_x]	17.185290309656	9	21	13	35	16	7
simplebaseline_coco_50_256 [participant_x]	17.265683171764	10	12	50	25	5	10

Conclusion benchmark

Temps d'inférence des modèles précis trop haut



Précision des modèles rapides trop faible



Réalisation d'un modèle custom ?

Modèle custom

Dataset :

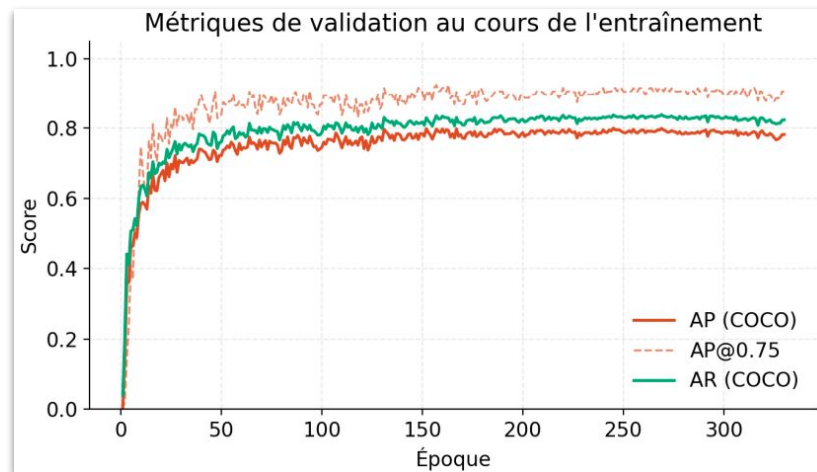
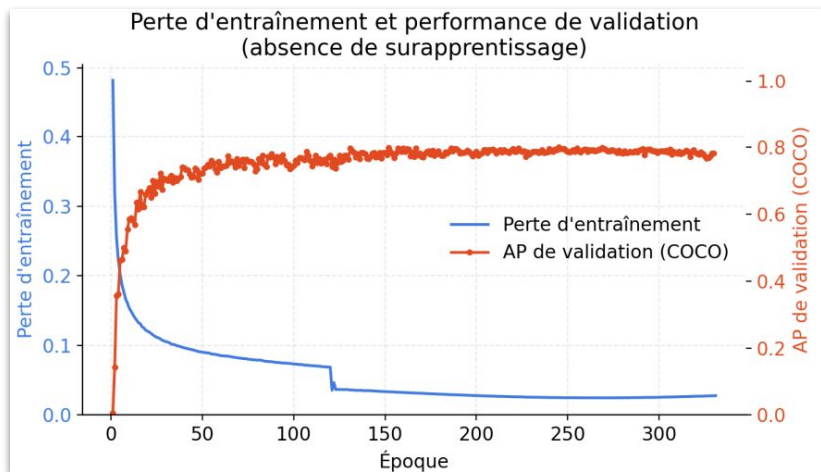
- 10 000 frames en entrées (> 50 sujets)
- Annotation via OpenPose
- Validation visuelle sur vidéo annotée
- Annotation manuelle sur erreur OpenPose (taux erreur < 0.001%)
- Annotation propre convertit au format COCO

Configuration

(inspirée du top-modèle RTMpose) :

- DataAugmentation => échelle, inclinaison, luminosité
- Backbone => CSPNeXt
- Neck => CSPNeXtPAFPN
- Head => RTMWHead avec décodeur SimCC

Analyse de l'apprentissage du modèle



Résultats : rapidité (temps d'inférence)

model	mean_ms	fps_mean
CUSTOM_22_04_26	8.5783812360436	116,5721098754963
rtmpose_aic_coco_t_256 participant_x	11.6847573441886	85.5815803909138
rtmpose_aic_coco_s_256 participant_x	12.388353972095	80.7209740900621
rtmpose_aic_coco_t_256 participant_x	13.6624242250089	73.1934526062739
simplebaseline_coco_aic_50_256 participant_x	14.0189502573133	71.3320171371837
rtmpose_aic_coco_s_256 participant_x	14.0807961023175	71.0187117783355
rtmpose_aic_coco_t_256 participant_x	14.260887159515	70.1218647069086
rtmpose_aic_coco_t_256 participant_x	14.400766404282	69.4407486328403
rtmpose_aic_coco_t_256 participant_x	14.4481576956155	69.2129765654111
rtmpose_aic_coco_s_256 participant_x	14.5069330293956	68.9325578310512

Résultats : répétabilité (écart-type intra)

model	global_kp_dist	rank_globe	rank_shoulder	rank_elbow	rank_hip	rank_knee	rank_ankle
CUSTOM_22_04_26	33.5880555055453	1	93	129	30	13	1
_vitpose_coco_s_256 [participant_x]	49.1973068612619	2	194	196	126	7	4
_vitpose_coco_b_256 [participant_x]	56.4562875229718	4	122	185	76	11	10
_hrnetlite_coco_18_384 [participant_x]	57.1863495696395	5	217	133	183	4	44
_vitpose_coco_b_256 [participant_x]	61.5052364755577	6	94	94	64	17	26
_rtmpose_aic_coco_s_256 [participant_x]	64.076825701775	7	123	246	153	6	30
_rtmpose_aic_coco_s_256 [participant_x]	65.0144714124102	8	125	207	65	163	2
_simplebaseline_coco_152_256 [participant_x]	65.5764427634779	9	205	161	220	23	7
_rtmpose_aic_coco_s_256 [participant_x]	66.7808922724646	10	152	214	253	10	17
_hrnetlite_mpii_18_256 [participant_x]	67.401159524268	11	278	258	271	5	3

Résultats : précision (écart-type à OpenPose)

model	global kp error	rank_globa	rank_shoulder	rank_elbow	rank_hip	rank_knee	rank_ankle
CUSTOM_22_04_26	11.176464820713	12	13	7	1	1	2
rtmpose_aic_coco_m_256 participant_x	14.933379109834	1	1	15	8	19	14
simplebaseline_coco_152_384 participant_x	15.362901601607	2	2	8	10	13	27
vitpose_coco_h*_256 participant_x	15.774814029084	3	43	17	3	6	5
rtmpose_aic_coco_l_256 participant_x	15.927182067217	4	27	21	9	12	4
rtmpose_cocktail14_x_384 participant_x	16.275651714206	5	33	1	13	3	23
vitpose_coco_h_256 participant_x	16.466015393985	6	62	16	7	7	3
hmet_coco_w48_384 participant_x	16.538892130113	7	4	45	17	4	20
rtmpose_cocktail14_l_384 participant_x	16.539600955729	8	23	10	18	2	29
hmet_coco_wholebody_w48_384 participant_x	17.185290309656	9	21	13	35	16	7

Conclusion, travaux en cours et futurs

