

Détection de la Distraction au Volant

Par la géométrie des poses du corps

Abakar Abbas Abakar · Pr. Yacine Yaddaden

Département de mathématiques, informatique et génie — UQAR



1,35M

décès annuels
sur les routes
OMS (Organisation Mondiale de
la santé)

25%

des accidents liés
à l'inattention
NHTSA (National Highway
Traffic Safety Administration)

?

Comment détecter
la distraction avec
un modèle léger ?

Les approches CNN actuelles sont précises mais coûteuses, peu interprétables et nécessitent de grandes quantités de données.

01

Pipeline léger et non intrusif

Basé uniquement sur la géométrie du haut du corps, sans capteur physique ni GPU requis.

02

Classification binaire

Reformulation du problème en Normal vs. Distract, adaptée aux contraintes réelles de sécurité embarquée.

03

Évaluation rigoureuse

Protocole LOSO-CV strict sur deux datasets publics (StateFarm & AUCDD), garantissant la généralisation inter-sujets.

StateFarm

4 850 images

10 classes de comportements (c0–c9)

25 conducteurs

Caméra embarquée en véhicule réel

Classes équilibrées : c0 (normal) = c1–c9 fusionnées (distrain)

2 425 images par classe

AUCDD

6 480 images

10 classes de comportements (c0–c9)

40 conducteurs

Véhicule réel + simulateur

Classes équilibrées : c0 (normal) = c1–c9 fusionnées (distrain)

3 240 images par classe

1. Pré-traitement

HRNet détecte 7 points clés
du haut du corps
(nez, épaules, coudes, poignets)
Seuil de confiance : $ci \geq 0.7$



2. Extraction

Vecteur $f \in \mathbb{R}^{31}$
21 distances euclidiennes
normalisées
10 angles articulaires



3. Classification

Normal vs. Distrait
Random Forest (500 arbres)
LOSO-CV — 5 folds

Protocole LOSO-CV — Leave-One-Subject-Out Cross-Validation

Aucun conducteur commun entre entraînement et test → évaluation véritablement indépendante du sujet

StateFarm

25 conducteurs → 5 groupes de 5

AUCDD

40 conducteurs → 5 groupes de 8

Métriques utilisées :

Exactitude

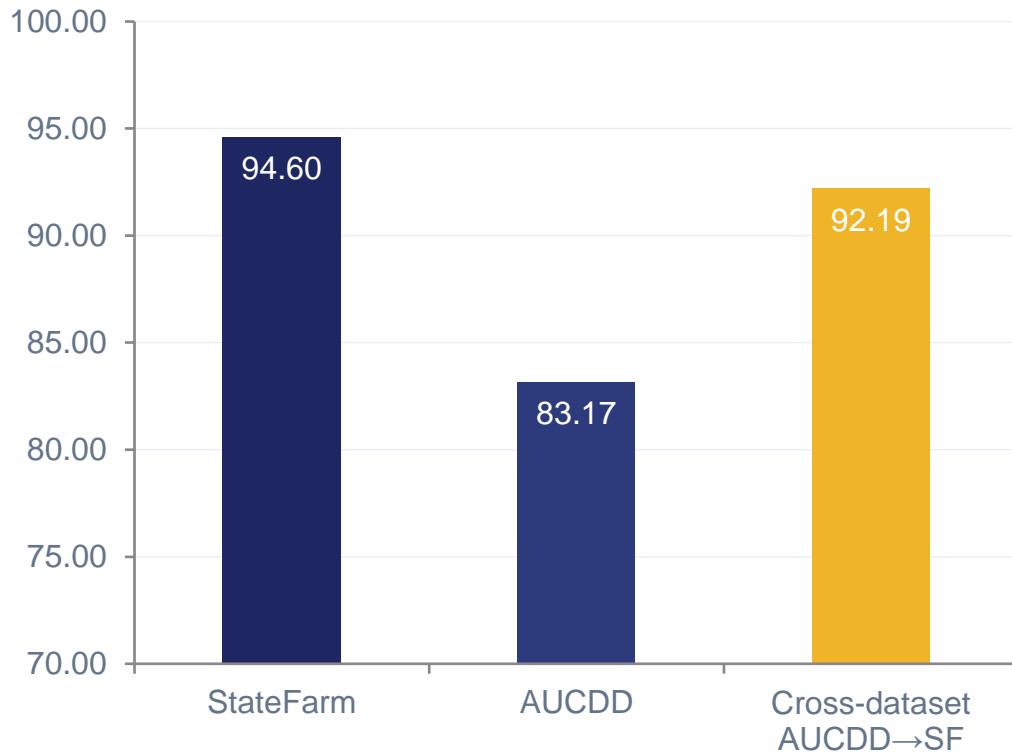
Précision

Rappel (Normal & Distrait)

F1-score

AUC-ROC

Exactitude par dataset (%)



0.9806

AUC-ROC — StateFarm

0.9028

AUC-ROC — AUCDD

92,19 %
généralisation inter-datasets
sans réentraînement

StateFarm

	Prédit Normal	Prédit Distrait
Réel Normal	94,6 %	5,4 %
Réel Distrait	5,4 %	94,6 %

AUCDD

	Prédit Normal	Prédit Distrait
Réel Normal	83,8 %	16,2 %
Réel Distrait	17,5 %	82,5 %

Interprétation

StateFarm : rappel parfaitement équilibré entre les deux classes (94,6 % / 94,6 %) — le modèle ne favorise aucune classe.

AUCDD : légère baisse due à la variabilité inter-caméras, mais reste robuste avec une AUC-ROC de 0,90.

Bilan

- ✓ 94,60 % sur StateFarm — 83,17 % sur AUCDD sous protocole LOSO-CV strict, sans GPU.
- ✓ Généralisation inter-datasets de 92,19 % sans réentraînement — robustesse démontrée.
- ✓ Alternative légère et interprétable aux CNN, adaptée au déploiement embarqué temps réel.

Perspectives

- Ajout d'une dimension temporelle (LSTM / Transformer) pour capturer l'évolution comportementale.
- Intégration de landmarks faciaux pour enrichir la représentation.
- Validation sur le dataset DMD (caméra triaxiale) avec DINOv2.