

Article

AgroPest-Trust Edge : conception et audit d'un système frugal, explicable et conscient de l'incertitude pour la reconnaissance de ravageurs et stress biotiques sur smartphone

Samson Kambale Muviri^{1*}, Nsenge Mpia Héritier²

¹Département d'Informatique de Gestion, Institut Supérieur de Commerce de Butembo, Butembo, RDC

²Faculté des Sciences et Technologie, Université de l'Assomption au Congo, Butembo, RDC

Correspondant : samsommuviri06@gmail.com

Citation : Muviri, S.K. & Mpia, H.N. AgroPest-Trust Edge : conception et audit d'un système frugal, explicable et conscient de l'incertitude pour la reconnaissance de ravageurs et stress biotiques sur smartphone. *Etincelle*, 2026, Vol. 26, no. 2. <https://doi.org/10.61532/rime262136>

Reçu : 10 juin 2026

Accepté : 25 juillet 2026

Publié : 28 juillet 2026

Note de l'éditeur : Ishango-UAC reste neutre en ce qui concerne les revendications juridiques dans les cartes publiées et les affiliations institutionnelles des auteurs.



Copyright : © 2026 par les auteurs. Soumis pour une publication en libre accès selon les termes de la licence Creative Commons Attribution (CC BY 4.0), sous réserve de l'acceptation éditoriale.

Abstract : L'identification précoce des ravageurs agricoles à partir d'images de smartphone demeure difficile lorsque les prises de vue sont floues, mal exposées ou encombrées par des arrière-plans complexes. Cette étude présente AgroPest-Trust Edge, un artefact de science de la conception qui associe classification légère, contrôle de qualité, calibration des probabilités, abstention sélective, explication Grad-CAM et recommandations prudentes de lutte intégrée. Un échantillon stratifié de 1 500 images a été constitué à partir des 1 591 fichiers du Dangerous Farm Insects Dataset, puis réparti en 1 042 images d'entraînement, 218 de validation et 240 de test. L'audit taxonomique montre que les quinze étiquettes ne décrivent pas exclusivement des insectes : treize classes correspondent à des insectes, une aux acariens et une à une maladie végétale. MobileNetV3-Small est retenu comme architecture cible en raison de son orientation mobile ; le protocole prévoit AdamW, mixup, pondération des classes, arrêt anticipé et température scaling. L'évaluation quantitative présentée distingue strictement les résultats de préparation du corpus des performances d'un benchmark ViT public. Ce dernier rapporte 90,90 % d'exactitude et 90,22 % de F1 macro sur 714 images, mais seulement 31,91 % de rappel pour Fall Armyworms. Cette hétérogénéité justifie une évaluation fondée sur les résultats par classe, la calibration, la couverture, le risque sélectif, la latence et la taille du modèle. L'article ne revendique pas de métriques MobileNetV3 non exécutées ; il fournit un cadre reproductible et auditable pour une validation locale future à Butembo.

Mot clés : vision par ordinateur ; ravageurs agricoles ; MobileNetV3 ; calibration ; Grad-CAM ; classification sélective ; edge AI ; Butembo.

1. Introduction

Dans les exploitations familiales, l'observation directe reste la première méthode de surveillance des insectes et des symptômes végétaux. Elle demeure irremplaçable, mais sa fiabilité dépend de l'expérience de l'observateur, de la taille de l'organisme, du stade d'infestation et des conditions de lumière. Une photographie prise au smartphone peut soutenir cette surveillance, à condition que le système d'intelligence artificielle ne réduise pas le problème à une simple étiquette de classe.

Les modèles de vision récents atteignent des niveaux élevés d'exactitude sur des corpus publics. Les architectures légères, notamment MobileNetV3, ont été conçues pour limiter les coûts de calcul sur les dispositifs mobiles (Howard et al., 2019), tandis qu'EfficientNet propose une mise à l'échelle conjointe de la profondeur, de la largeur et de la résolution (Tan & Le, 2019). Dans le domaine agricole, des adaptations de MobileNetV2 ont montré qu'un modèle compact peut conserver une performance élevée (Peng et al., 2024 ; Chen et al., 2024). Ces progrès ne suffisent toutefois pas à garantir une décision sûre dans les conditions de terrain.

Trois difficultés sont particulièrement importantes. Premièrement, les probabilités softmax sont fréquemment surconfiantes : un score de 90 % n'implique pas nécessairement une fréquence empirique de 90 % de bonnes réponses. Le temperature scaling constitue une correction post-entraînement simple et bien établie (Guo et al., 2017). Deuxièmement, un système qui répond à toutes les images, y compris aux images floues ou hors distribution, transforme une incertitude technique en risque décisionnel. Troisièmement, une prédiction exacte en moyenne peut masquer une faible sensibilité pour une classe agronomiquement critique.

L'explicabilité visuelle apporte une couche d'audit supplémentaire. Grad-CAM met en évidence les régions qui contribuent au score d'une classe (Selvaraju et al., 2017). Une carte d'activation ne constitue ni une preuve causale ni une annotation de détection, mais elle peut révéler que le modèle s'appuie sur une feuille, une lésion, un organisme, un filigrane ou un arrière-plan. Pour une application destinée à la surveillance agricole, cette distinction est essentielle.

La présente étude s'inscrit dans une logique de conception responsable en contexte contraint. Des travaux menés dans d'autres domaines ont montré l'importance d'adapter les modèles aux réalités locales et aux ressources disponibles : CoBERT contextualise la recommandation dans des environnements socio-économiques instables (Mpia et al., 2023), tandis qu'un système léger de reconnaissance des cris de nourrissons démontre la faisabilité d'une inférence et d'une notification en temps réel sur matériel embarqué à faible coût (Mpia et al., 2025). Ces articles ne portent pas sur les ravageurs ; ils éclairent néanmoins le principe méthodologique selon lequel la valeur d'un modèle dépend de son intégration au contexte, au matériel et au processus de décision.

L'objectif est donc de concevoir et d'auditer AgroPest-Trust Edge, un système de reconnaissance multi-classe qui combine frugalité, calibration, abstention sélective, explication et restitution prudente. Les contributions sont : (i) un audit taxonomique et documentaire du corpus ; (ii) une sélection stratifiée et reproductible de 1 500 images ; (iii) une architecture MobileNetV3-Small orientée edge ; (iv) un mécanisme de contrôle qualité et de rejet ; (v) une analyse critique d'un benchmark ViT public ; et (vi) une interface qui distingue la prédiction, la confiance et la recommandation. L'article signale explicitement l'absence de métriques propres au MobileNetV3 tant que l'expérience complète n'a pas été exécutée et archivée.

2. Cadre conceptuel et travaux connexes

2.1. Classification agricole et architectures légères

La tâche étudiée est une classification d'image entière. Elle diffère de la détection d'objets, qui exige des boîtes englobantes, et de la segmentation, qui exige des annotations par pixel. Le corpus Dangerous Farm Insects est organisé en dossiers de classes et ne fournit pas d'annotations spatiales. Revendiquer une détection à partir de ces seules données serait donc méthodologiquement incorrect. La carte Grad-CAM est utilisée comme localisation faible des zones influentes, sans être assimilée à une vérité terrain.

Peng et al. proposent PestNet, une amélioration légère de MobileNetV2 qui mobilise notamment AdamW et mixup (Peng et al., 2024). Akhtar et al. (2025) comparent MobileViT et EfficientNetV2B2, puis étudient leur quantification et leur déploiement mobile sur des corpus d'insectes agricoles. Ragu et Teo (2025) relient l'apprentissage à faible nombre d'exemples, le transfert et Grad-CAM dans la classification des ravageurs. Ces travaux montrent la progression vers des modèles plus compacts et plus interprétables, mais la calibration, le rejet sélectif et la traduction du résultat en conseil restent rarement réunis dans un même flux utilisateur.

L'apprentissage par transfert est adapté au nombre limité d'images disponibles. Le modèle est initialisé avec des représentations ImageNet, puis sa couche de sortie est remplacée par quinze neurones. Mixup régularise les frontières de décision en interpolant les images et les étiquettes (Zhang et al., 2018), tandis qu'AdamW sépare la décroissance des poids de l'étape d'optimisation adaptative (Loshchilov et Hutter, 2019). Ces choix ne garantissent pas à eux seuls la généralisation ; ils constituent des hypothèses de conception à tester sur un protocole strictement séparé.

2.2. Calibration, abstention et explicabilité

Un classifieur calibré associe ses niveaux de confiance à des fréquences observées de correction. Lorsque les prédictions annoncées à 70 % sont correctes environ sept fois sur dix, la confiance devient interprétable. L'Expected Calibration Error (ECE) résume l'écart entre confiance moyenne et exactitude dans des intervalles de probabilité. Le temperature scaling apprend un unique paramètre T sur l'ensemble de validation ; il modifie l'amplitude des logits sans changer l'ordre des classes (Guo et al., 2017).

La classification sélective introduit une décision supplémentaire : accepter ou refuser. La couverture mesure la proportion des images acceptées et le risque sélectif l'erreur parmi ces images. Une hausse du seuil de confiance réduit généralement la couverture mais peut améliorer la fiabilité des réponses délivrées. Dans AgroPest-Trust Edge, la décision dépend à la fois de la confiance calibrée et de la qualité de l'image. L'abstention ne constitue pas un échec du système ; elle est une réponse appropriée lorsque l'information visuelle est insuffisante.

Grad-CAM complète cette logique en permettant un audit de la zone influente (Selvaraju et al., 2017). Une activation centrée sur le ravageur ou le symptôme est plus plausible qu'une activation concentrée sur un bord, un texte ou un fond répétitif. L'explication demeure qualitative en l'absence d'annotations de régions. Elle doit donc être présentée comme un indice d'audit, non comme une justification biologique définitive.

2.3. Positionnement de l'étude

Tableau 1 : Positionnement critique des travaux connexes

Travail	Approche	Force principale	Limite observée
Peng et al. (2024)	MobileNetV2 amélioré, AdamW, mixup	Modèle léger	Calibration et abstention non centrales
Akhtar et al. (2025)	MobileViT, EfficientNetV2B2, quantification	Taille-latence-exactitude	Explicabilité et risque sélectif non intégrés
Ragu & Teo (2025)	Few-shot, transfert, Grad-CAM	Lien entre faibles données et XAI	Déploiement edge peu développé
Dima806 (2023)	ViT de classification	Rapport détaillé par classe	Modèle lourd ; rappel faible pour Fall Armyworms
Présente étude	MobileNetV3, calibration, rejet, Grad-CAM, UI	Fiabilité opérationnelle et transparence	Validation locale et métriques propres à compléter

3. Matériels et méthodes

3.1. Démarche de science de la conception

La recherche suit une démarche de science de la conception, adaptée à la production et à l'évaluation d'un artefact informatique (Peffers et al., 2007). L'artefact comprend le pipeline de données, le modèle, les mécanismes de confiance, l'interface et les fichiers de traçabilité. Le cycle comporte la définition du problème, l'audit des données, la conception, l'implémentation, l'évaluation et la discussion des limites.

Une distinction est maintenue entre trois catégories de preuves : les données natives du corpus, les variables dérivées par le pipeline et les performances provenant d'un modèle public indépendant. Cette séparation prévient l'attribution erronée d'un résultat externe au modèle proposé. Les fichiers attendus pour l'expérience propre portent le préfixe `experiment_` afin de ne pas écraser les sorties de référence.

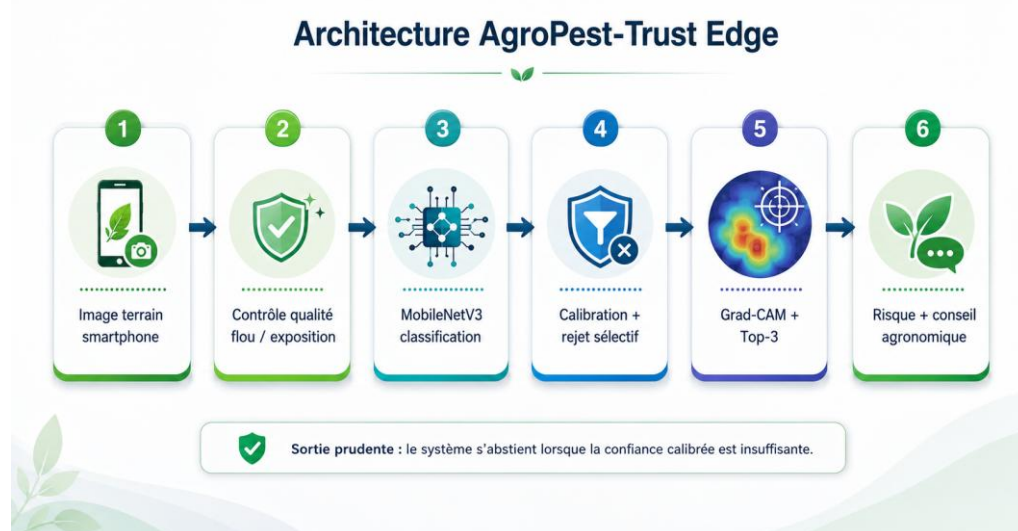


Figure 1 : Architecture fonctionnelle d'AgroPest-Trust Edge.

3.2. Corpus, provenance et audit taxonomique

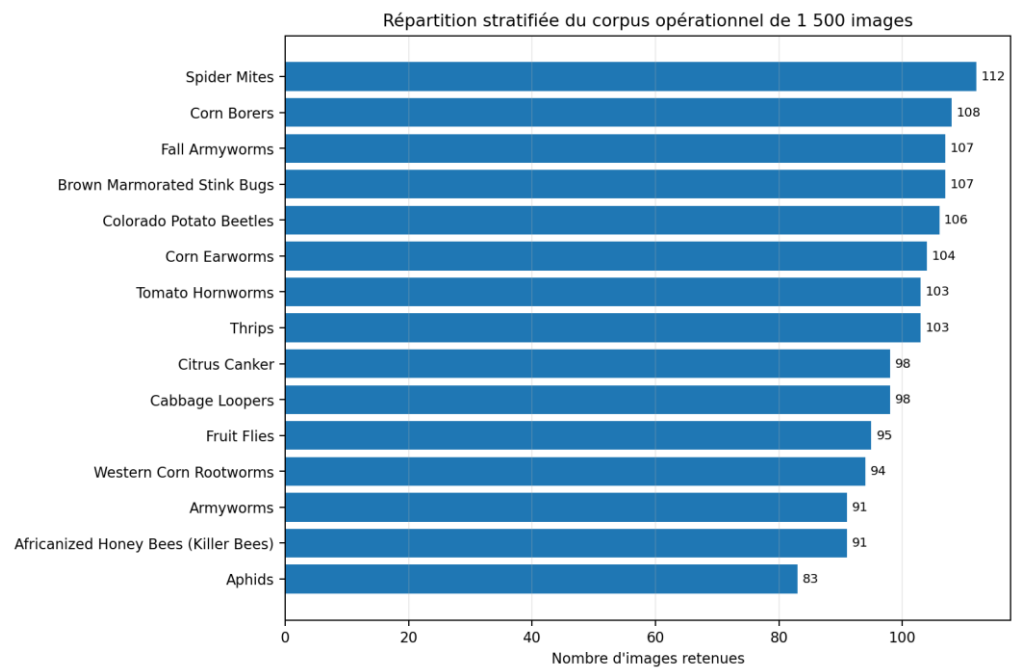
Le corpus primaire est le Dangerous Farm Insects Dataset, accessible sur Kaggle et attribué à Tarun R. Jain avec un collaborateur (Jain et collaborateur, 2023). La page du corpus décrit quinze catégories. Le notebook de Saxena a servi de point d'entrée documentaire pour examiner l'organisation des données (Saxena, s. d.), sans que ses sorties soient reprises comme résultats propres.

Le répertoire source contient 1 591 images. Une sélection stratifiée en retient exactement 1 500, afin de respecter la contrainte du projet tout en conservant la distribution relative des classes. L'unité d'analyse est une image unique. La variable cible `class_name` correspond au nom du dossier de classe. Les empreintes SHA-256 servent à repérer les doublons exacts ; les dimensions, formats et scores de qualité sont des métadonnées de gouvernance et ne sont pas utilisés comme prédicteurs par le CNN.

L'audit taxonomique révèle une incohérence utile à documenter : Citrus Canker est une maladie végétale et Spider Mites correspond à des acariens. Le système est donc défini comme un classifieur de ravageurs et de stress biotiques, non comme un classifieur exclusif d'insectes. Cette reformulation évite d'étendre artificiellement la portée biologique des données.

Tableau 2 : Répartition stratifiée du corpus opérationnel

Classe	Total	Train	Val.	Test	Statut
Africanized Honey Bees	91	63	13	15	Insecte
Aphids	83	58	12	13	Insecte
Armyworms	91	63	13	15	Insecte
Brown Marmorated Stink Bugs	107	74	16	17	Insecte
Cabbage Loopers	98	68	14	16	Insecte
Citrus Canker	98	68	14	16	Maladie
Colorado Potato Beetles	106	74	15	17	Insecte
Corn Borers	108	75	16	17	Insecte
Corn Earworms	104	72	15	17	Insecte
Fall Armyworms	107	74	16	17	Insecte
Fruit Flies	95	66	14	15	Insecte
Spider Mites	112	78	16	18	Acarien
Thrips	103	72	15	16	Insecte
Tomato Hornworms	103	72	15	16	Insecte
Western Corn Rootworms	94	65	14	15	Insecte

**Figure 2 : Effectifs retenus dans les quinze classes.**

3.3. Séparation des données et prétraitement

La sélection est réalisée avec une graine fixe de 42. Les 1 500 images sont réparties en 1 042 images d'entraînement, 218 de validation et 240 de test. L'ensemble de test est isolé avant toute optimisation. Les statistiques de normalisation proviennent d'ImageNet. Les images sont redimensionnées à 224×224 pixels ; l'entraînement utilise un

recadrage aléatoire modéré, un retournement horizontal, une rotation limitée et une variation contrôlée de luminosité et de couleur.

Le contrôle qualité d'inférence combine la variance du Laplacien pour la netteté et des indicateurs de luminance pour l'exposition. Ces scores n'établissent pas une vérité absolue sur la qualité ; ils servent de filtre conservateur. Une image manifestement floue ou surexposée doit déclencher une demande de nouvelle capture plutôt qu'une prédiction forcée.

3.4. Architecture et protocole d'entraînement

Tableau 3 : Hyperparamètres du modèle principal

Paramètre	Valeur
Entrée	224 × 224 pixels
Architecture principale	MobileNetV3-Small
Comparatif prévu	EfficientNet-B0
Initialisation	Poids ImageNet
Optimiseur	AdamW
Taux initial	3×10^{-4}
Weight decay	1×10^{-4}
Batch	32
Époques maximales	30
Patience	7
Mixup α	0,2
Label smoothing	0,05
Graine	42

MobileNetV3-Small est choisi comme architecture principale parce qu'il a été recherché en tenant compte du matériel mobile (Howard et al., 2019). La couche de classification est remplacée par une couche linéaire de quinze sorties. La perte d'entropie croisée est pondérée selon la fréquence des classes et reçoit un label smoothing de 0,05. L'arrêt anticipé utilise la perte de validation avec une patience de sept époques.

Le choix du modèle ne présuppose pas qu'il surpassera le ViT en exactitude. L'hypothèse testée est différente : un modèle léger peut offrir un compromis plus favorable entre exactitude, taille, latence, calibration et autonomie. Le résultat ne pourra être établi qu'après exécution du notebook et archivage des métriques propres.

3.5. Calibration, rejet sélectif et risque

Après l'entraînement, la température T est estimée sur l'ensemble de validation en minimisant la log-vraisemblance négative. Les probabilités calibrées sont obtenues par $\text{softmax}(z/T)$. L'ECE est calculée avant et après calibration. La prédiction est acceptée lorsque la confiance calibrée atteint 0,62 et que le score de qualité dépasse 0,32. Ces seuils sont des paramètres initiaux configurables, non des constantes universelles.

Le niveau de risque affiché par l'interface combine la confiance, la qualité et la classe. La recommandation privilégie l'inspection, l'hygiène de la parcelle, le retrait ciblé, la consultation d'un spécialiste et la lutte intégrée. Aucun pesticide, dosage ou calendrier d'application n'est prescrit automatiquement.

3.6. Métriques et protocole d'audit

Tableau 4 : Critères d'évaluation retenus

Métrique	Interprétation
Exactitude	Proportion globale de prédictions correctes
Précision	Fiabilité des alertes d'une classe
Rappel	Capacité à retrouver les cas réels
F1 macro	Poids égal accordé à chaque classe
ECE	Écart moyen confiance-exactitude
Couverture	Part des images acceptées
Risque sélectif	Erreur parmi les décisions acceptées
Latence	Temps moyen par image
Taille	Poids du fichier du modèle

La mAP n'est pas utilisée comme métrique principale, car aucune boîte englobante n'est prédite. La matrice de confusion, les F1 par classe, l'ECE, la couverture et le risque sélectif répondent directement à la tâche. Pour Grad-CAM, l'audit est qualitatif : la carte est inspectée avec l'image, la classe et le Top-3. Une validation quantitative de localisation exigerait des annotations spatiales absentes du corpus.

4. Résultats

4.1. Audit du corpus

Le manifeste final contient exactement 1 500 lignes. Spider Mites est la classe la plus représentée avec 112 images ; Aphids est la moins représentée avec 83 images. Le rapport maximal/minimal est d'environ 1,35, ce qui correspond à un déséquilibre modéré. La séparation comprend 69,5 % d'images d'entraînement, 14,5 % de validation et 16,0 % de test.

Treize classes correspondent à des insectes, une aux acariens et une à une maladie végétale. Cette information est conservée comme variable `taxonomic_status`. Aucune variable locale sur la culture hôte, la date, le site, la sévérité ou le traitement n'est disponible dans la source. Le corpus permet donc un prototypage de classification, mais pas une analyse causale, spatio-temporelle ou agronomique propre à Butembo.

4.2. Benchmark ViT public

Le modèle public `dima806/farm_insects_image_detection` est une adaptation de ViT-base-patch16-224-in21k et publie un rapport de classification sur 714 images (Dima806, 2023). Il est utilisé comme référence externe, non comme résultat du MobileNetV3 proposé. L'exactitude est de 90,90 %, la précision macro de 91,47 %, le rappel macro de 90,85 % et le F1 macro de 90,22 %.

Ces agrégats suggèrent une bonne performance moyenne, mais ils ne décrivent pas la variabilité entre classes. L'analyse par classe montre une dissociation marquée entre Fall Armyworms et Armyworms, deux catégories visuellement proches et potentiellement confondues.

Tableau 5 : Performances par classe du benchmark ViT public

Classe	Précision %	Rappel %	F1 %	n	FN
Fall Armyworms	78,95	31,91	45,45	47	≈32
Western Corn Rootworms	97,87	97,87	97,87	47	≈1
Colorado Potato Beetles	100,00	97,92	98,95	48	≈1
Thrips	97,62	87,23	92,13	47	≈6
Corn Earworms	90,70	81,25	85,71	48	≈9
Cabbage Loopers	93,88	95,83	94,85	48	≈2
Armyworms	61,43	91,49	73,50	47	≈4
Brown Marmorated Stink Bugs	100,00	100,00	100,00	48	0
Tomato Hornworms	97,92	100,00	98,95	47	0
Citrus Canker	90,38	100,00	94,95	47	0
Aphids	90,20	95,83	92,93	48	≈2
Corn Borers	81,48	91,67	86,27	48	≈4
Fruit Flies	100,00	100,00	100,00	48	0
Africanized Honey Bees	100,00	100,00	100,00	48	0
Spider Mites	91,67	91,67	91,67	48	≈4

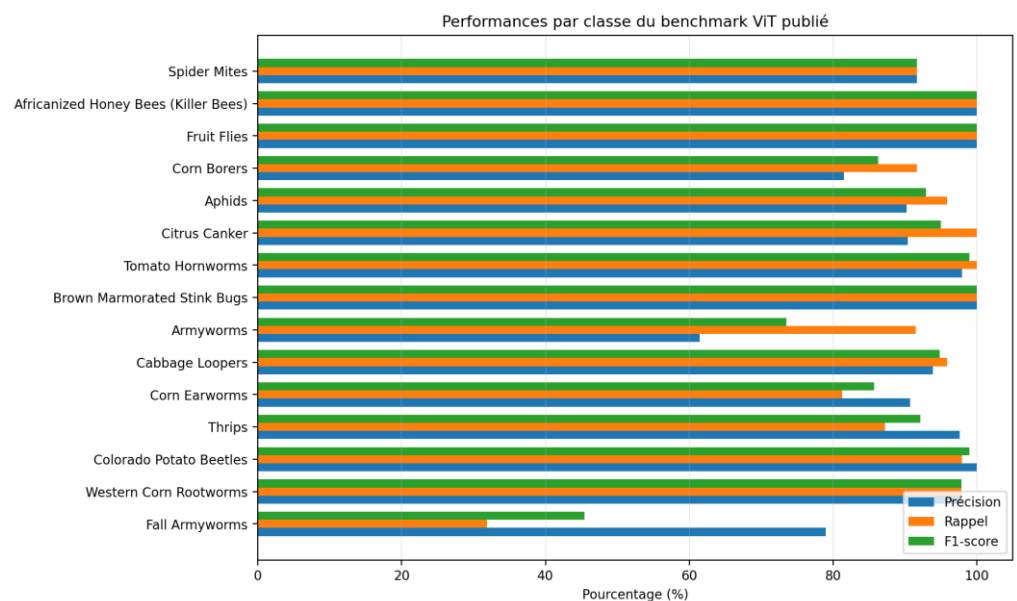


Figure 3 : Précision, rappel et F1-score par classe du benchmark ViT.

4.3. Analyse des erreurs

Fall Armyworms présente un rappel de 31,91 %. En appliquant ce rappel au support de 47 images, environ 32 cas ne sont pas retrouvés. Cette estimation est fondée sur des valeurs arrondies et doit être interprétée comme un ordre de grandeur. Le F1 de 45,45 % confirme que la classe ne peut pas être considérée comme résolue.

Armyworms atteint au contraire 91,49 % de rappel, mais seulement 61,43 % de précision. Le modèle récupère donc la plupart des vrais Armyworms au prix de nombreux faux positifs. Cette asymétrie est compatible avec une confusion entre Armyworms et Fall Armyworms, mais seule la matrice de confusion complète permettrait de vérifier précisément les directions d'erreur.

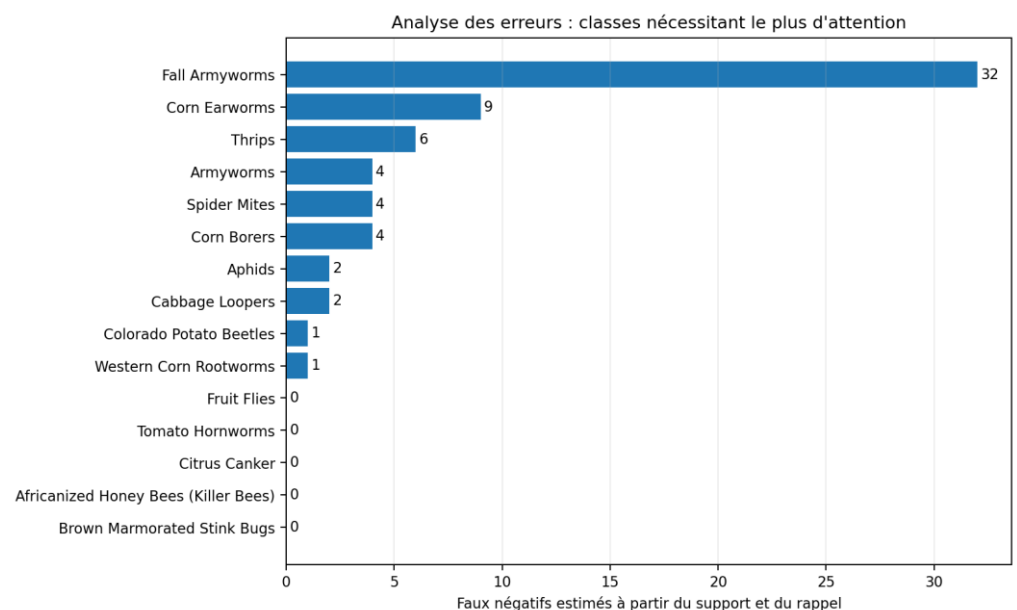


Figure 4 : Faux négatifs estimés à partir du support et du rappel publié.

4.4. Compromis de déploiement edge

Tableau 6 : Comparaison de résultats publiés sur le corpus

Modèle	Exactitude	FP32	Quantifié	Latence	Source
MobileViT	73,4 %	22,0 Mo	5,4 Mo	0,34 s	Akhtar et al.
EfficientNetV2B2	77,8 %	33,0 Mo	9,6 Mo	0,20 s	Akhtar et al.
ViT public	90,9 %	Non rapporté	Non rapporté	Non rapportée	Dima806

Dans l'étude d'Akhtar et al., EfficientNetV2B2 quantifié conserve 77,8 % d'exactitude avec une taille de 9,6 Mo et une latence rapportée de 0,20 seconde (Akhtar et al., 2025). MobileViT atteint 73,4 %, 5,4 Mo après quantification et 0,34 seconde. Le ViT public est plus exact, mais la référence consultée ne fournit pas dans le même tableau une comparaison homogène de latence et de taille sur le matériel cible.

La figure 5 ne constitue donc pas un classement universel. Les modèles, procédures et matériels ne sont pas strictement identiques. Elle illustre seulement le compromis : une exactitude supérieure peut être associée à un coût plus élevé, alors qu'un modèle quantifié peut être plus favorable au déploiement.

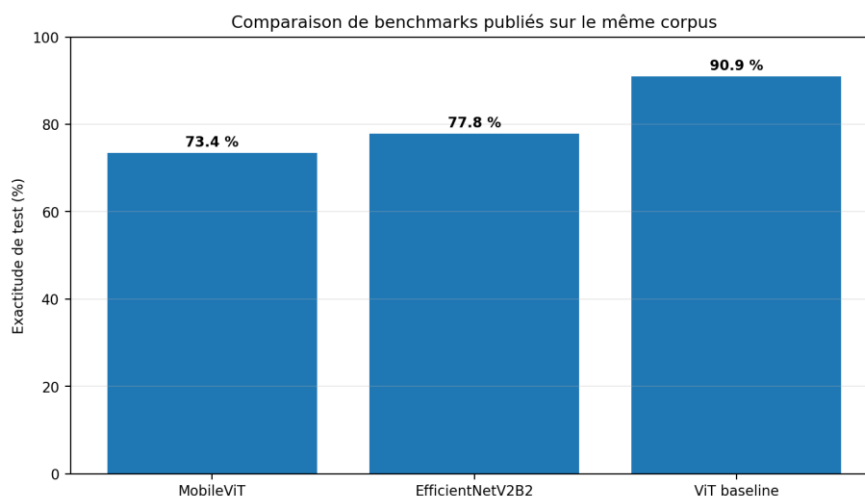


Figure 5 : Exactitudes de test publiées, à interpréter selon les protocoles respectifs.

4.5. Interface et restitution

L'interface du prototype présente l'image, la qualité, la classe Top-1, les trois hypothèses principales, la confiance calibrée, le niveau de risque et une recommandation. Le message d'accueil précise que le système fournit une aide à la surveillance et non un diagnostic agronomique certifié.

Lorsque le seuil de confiance ou de qualité n'est pas atteint, la sortie devient « analyse incertaine ». L'utilisateur est invité à rapprocher l'appareil, améliorer la lumière ou demander une confirmation. Cette règle rend visible l'incertitude et réduit la pression pour produire une réponse à tout prix.

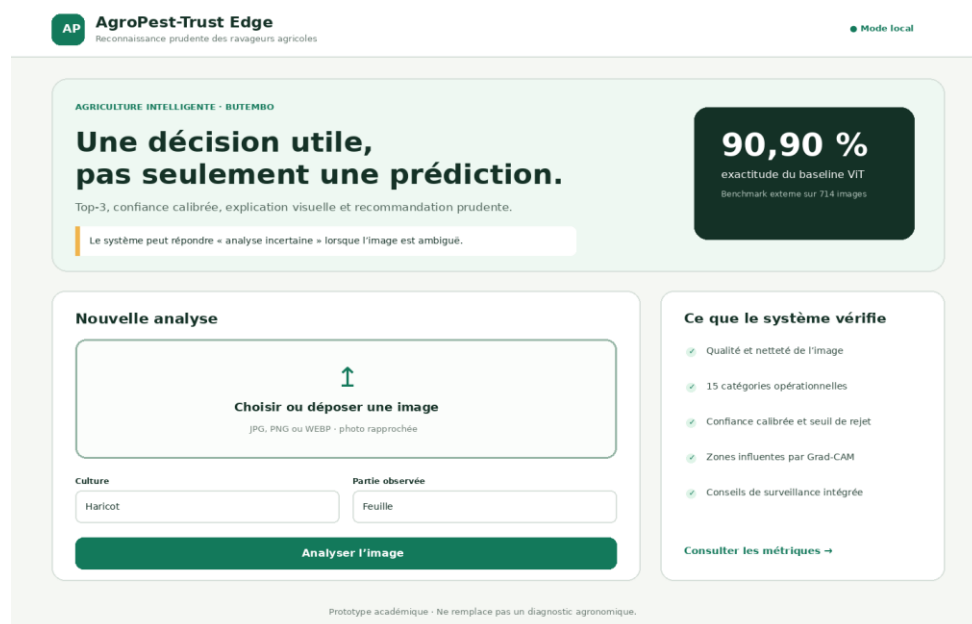


Figure 6 : Interface d'accueil du prototype AgroPest-Trust Edge.

5. Discussion

5.1. Pourquoi l'exactitude globale ne suffit pas

Le principal enseignement est méthodologique : une exactitude de 90,90 % peut coexister avec un rappel inférieur à 32 % pour une classe. Un système présenté uniquement à travers son score global donnerait une image incomplète de sa fiabilité. Le F1 macro, les résultats par classe et la matrice de confusion doivent accompagner l'exactitude, surtout lorsque certaines erreurs conduisent à une intervention tardive.

La calibration ajoute une dimension différente. Deux modèles de même exactitude peuvent produire des probabilités de qualité très différente. Une confiance mal calibrée rend les seuils arbitraires et peut encourager une interprétation excessive. L'ECE avant et après temperature scaling permettra de vérifier si les probabilités du MobileNetV3 deviennent plus cohérentes avec les fréquences de correction (Guo et al., 2017).

5.2. Apport de l'abstention et de l'explication

L'abstention introduit un compromis explicite entre couverture et risque. Dans un contexte où l'utilisateur peut reprendre une photographie, il est rationnel de refuser une image de faible qualité plutôt que de maximiser la couverture. Ce principe rapproche l'évaluation du coût réel des erreurs et transforme l'incertitude en action concrète : nouvelle capture, inspection manuelle ou consultation.

Grad-CAM permet d'identifier des raccourcis visuels, notamment les arrière-plans, textes ou bordures. Sa valeur dépend toutefois de la manière dont elle est utilisée. Une carte chaude ne prouve pas que le modèle a appris une caractéristique biologiquement correcte. Elle devient utile lorsqu'elle est confrontée à l'image, au Top-3, à la classe réelle et au contexte de capture.

5.3. Validité externe et contexte de Butembo

Le corpus public ne représente pas nécessairement les cultures, appareils, éclairages et ravageurs observés à Butembo. Le fait que l'interface soit destinée à ce contexte

ne démontre pas la généralisation locale. Une validation externe devra collecter des images datées, documenter la culture hôte, l'appareil, la résolution, le stade de développement, le site approximatif et l'identification experte.

Cette limite n'annule pas l'utilité du prototype. Elle délimite correctement sa contribution actuelle : un cadre de données, de confiance et d'interface prêt à recevoir des preuves locales. La démarche rejoint les travaux qui insistent sur l'adaptation des systèmes d'IA aux environnements contraints et aux usages réels (Mpia et al., 2023 ; Mpia et al., 2025).

5.4. Limites de l'étude

Tableau 7 : Menaces à la validité et stratégies de réduction

Type	Menace	Réduction
Validité interne	Images dupliquées, qualité variable ou erreurs d'étiquetage	SHA-256, manifeste, contrôle d'intégrité et inspection des erreurs
Validité de construit	Confusion entre classification et détection	Tâche définie comme classification ; aucune mAP sans boîtes
Validité statistique	Un seul split et métriques MobileNetV3 non encore archivées	Graine fixe ; futures répétitions multi-graines et k-fold
Validité externe	Corpus international, contexte local absent	Validation externe à Butembo et test hors distribution
Validité opérationnelle	Latence non mesurée sur plusieurs smartphones	Benchmark sur appareils cibles et export quantifié
Validité agronomique	Conseils non validés localement	Relecture par agronome et adaptation aux pratiques locales

La limite la plus importante est l'absence, dans le manuscrit source, d'un fichier de résultats propre à MobileNetV3. Les métriques du ViT public ne doivent pas être transférées au modèle proposé. Une version expérimentale ultérieure devra joindre au minimum les courbes d'apprentissage, le rapport de classification, la matrice de confusion, l'ECE avant et après calibration, les courbes risque-couverture, la taille des exports et la latence mesurée sur les appareils cibles.

6. Conclusion

AgroPest-Trust Edge répond à un problème souvent négligé dans la reconnaissance agricole : un modèle utile ne doit pas seulement prédire, mais également indiquer quand sa réponse est incertaine, montrer les régions influentes et restituer une recommandation prudente. Le travail a produit un corpus opérationnel stratifié de 1 500 images, un audit taxonomique, une architecture edge, un protocole de calibration et de rejet, ainsi qu'une interface orientée vers la surveillance.

L'audit du benchmark ViT montre que 90,90 % d'exactitude globale masque un rappel de 31,91 % pour Fall Armyworms. Ce résultat justifie l'évaluation par classe et l'intégration de la calibration, de la couverture et du risque sélectif. Il ne constitue pas une performance propre au MobileNetV3. La prochaine étape scientifique consiste à exécuter et archiver l'expérience complète, puis à valider le système sur des images locales expertisées à Butembo.

Contributions : Conceptualisation, M.S.K.; méthodologie, M.S.K. & M.H.N.; validation, M.S.K. & M.H.N.; Investigation, M.S.K.; ressources, M.S.K. & M.H.N.; Traitement des données, M.S.K.; écrire le manuscrit, M.S.K.; visualisation, M.S.K.; supervision, M.H.N, correction du manuscrit, M.S.K.

Les auteurs ont lu et approuvé la version publiée de ce manuscrit.

Sponsor financier : Cette recherche n'a reçu aucun soutien financier.

Disponibilité des données : Les données ne sont pas disponibles.

Remerciement : Non applicable.

Conflits d'intérêt : Les auteurs déclarent aucun conflit d'intérêt.

Références

1. Akhtar, M. H., Eksheir, I., & Shanableh, T. (2025). Edge-Optimized Deep Learning Architectures for Classification of Agricultural Insects with Mobile Deployment. *Information*, 16(5), 348. <https://doi.org/10.3390/info16050348>
2. Chen, H., Zhou, G., He, W., Duan, X., & Jiang, H. (2024). Classification and identification of agricultural products based on improved MobileNetV2. *Scientific Reports*, 14, 3454. <https://doi.org/10.1038/s41598-024-53349-w>
3. Dima806. (2023). farm_insects_image_detection [Model card and classification report]. Hugging Face. https://hugging-face.co/dima806/farm_insects_image_detection
4. Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On Calibration of Modern Neural Networks. *Proceedings of the 34th International Conference on Machine Learning*, 70, 1321–1330. <https://proceedings.mlr.press/v70/guo17a.html>
5. Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1314–1324. <https://doi.org/10.1109/ICCV.2019.00140>
6. Jain, T. R., & collaborateur. (2023). Dangerous Farm Insects Dataset [Dataset]. Kaggle. <https://www.kaggle.com/datasets/tarundalal/dangerous-insects-dataset>
7. Loshchilov, I., & Hutter, F. (2019). Decoupled Weight Decay Regularization. *International Conference on Learning Representations*. <https://openreview.net/forum?id=Bkg6RiCqY7>
8. Mpia, H. N., Mburu, L. W., & Mwendia, S. N. (2023). CoBERT: A contextual BERT model for recommending employability profiles of information technology students in unstable developing countries. *Engineering Applications of Artificial Intelligence*, 125, 106728. <https://doi.org/10.1016/j.engappai.2023.106728>
9. Mpia, H. N., Kavalami, M. M., Lusenge, G. K., Ushindi, K. P., Kyalengekana, D.-D. K., & Ciswa, O. M. (2025). Lightweight deep learning system for infant-cry recognition with real-time notification in resource-constrained environments. *Machine Learning with Applications*, 22, 100791. <https://doi.org/10.1016/j.mlwa.2025.100791>
10. Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
11. Peng, H., Xu, H., Shen, G., Liu, H., Guan, X., & Li, M. (2024). A Lightweight Crop Pest Classification Method Based on Improved MobileNet-V2 Model. *Agronomy*, 14(6), 1334. <https://doi.org/10.3390/agronomy14061334>
12. Ragu, N., & Teo, J. (2025). Pest classification: Explainable few-shot learning vs. convolutional neural networks vs. transfer learning. *Scientific African*, 27, e02512. <https://doi.org/10.1016/j.sciaf.2024.e02512>
13. Saxena, U. (s. d.). Farm Insect Detection | Transfer Learning [Notebook]. Kaggle. Consulté le 24 juillet 2026. <https://www.kaggle.com/code/utkarshsaxenadn/farm-insect-detection-transfer-learning>
14. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
15. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning*, 97, 6105–6114. <https://proceedings.mlr.press/v97/tan19a.html>
16. Zhang, H., Cissé, M., Dauphin, Y. N., & Lopez-Paz, D. (2018). mixup: Beyond Empirical Risk Minimization. *International Conference on Learning Representations*. <https://openreview.net/forum?id=r1Ddp1-Rb>