

**HIKVISION®**

# Livre blanc sur la technologie AcuSeek



 Homme à lunettes, casquette, manteau long gris



# TABLE DES MATIÈRES

## 1. Introduction 3

## 2. Contexte 3

## 3. Comment fonctionne AcuSeek 3

3.1 L'architecture technique d'AcuSeek .....	4
3.2 Fondements techniques d'AcuSeek .....	6
3.3 Modes d'application d'AcuSeek .....	7
3.4 D'AcuSeek à AcuSearch .....	7

## 4. Scénarios d'application 8

4.1 Résidence connectée   Etablissement de santé connecté .....	8
4.2 Usine intelligente et chantier connecté .....	9
4.3 Parc intelligent .....	9
4.4 Commerce connecté .....	10

## 5. Intelligence artificielle responsable 10

# 1. Introduction

Ce livre blanc propose une analyse approfondie du système technologique AcuSeek de Hikvision, avec pour objectif d'établir un cadre de compréhension partagé entre tous les acteurs impliqués dans la chaîne produit. Il met en lumière l'application innovante de modèles d'intelligence artificielle allégés dans des dispositifs embarqués, et explique de manière structurée les limites techniques, les seuils de performance et la capacité d'adaptation de la solution AcuSeek à différents scénarios d'utilisation.

Le modèle de transmission technologie–scénario–valeur présenté dans ce document constitue un cadre de connaissance standardisé, destiné à accompagner l'exploration des besoins clients et à faciliter la démonstration de solutions personnalisées.

## 2. Contexte

Les méthodes traditionnelles de recherche vidéo reposent généralement sur des règles fixes, telles que la lecture par plage horaire, la lecture par événement, la lecture centrée sur les personnes ou les véhicules, ou encore la lecture par séquences. Parmi ces approches, la lecture par événement et celle par personne/véhicule s'appuient sur une analyse structurée, qui ne permet pas d'interpréter les informations non structurées et présente une capacité de généralisation limitée.

Avec l'émergence des grands modèles multimodaux, les NVR (enregistreurs vidéo en réseau) peuvent désormais retrouver rapidement des cibles à partir de descriptions textuelles, tout en prenant en charge des recherches approximatives. Cela permet non seulement de détecter des objets en mouvement et les éléments présents dans leur environnement, mais aussi d'analyser des comportements simples et les relations entre cibles adjacentes, comme les styles vestimentaires, les types de véhicules, les marques de voitures ou encore certaines actions élémentaires. Cette approche se distingue par une grande capacité de généralisation.

AcuSeek peut être déployé dans de nombreux secteurs clés. Dans les espaces publics tels que les gares, les écoles, les quartiers résidentiels, les chantiers, les centres commerciaux ou les sites touristiques, il aide les utilisateurs à localiser des valises, des enfants disparus, des trottinettes, des véhicules, des sacs à dos, des outils et d'autres objets.

## **3. Comment fonctionne AcuSeek**

### **Modèles unimodaux vs. grands modèles multimodaux en intelligence artificielle**

Dans le domaine de l'intelligence artificielle, les systèmes d'analyse intelligente sont classés en deux catégories selon leurs capacités de traitement des données : les modèles unimodaux et les modèles multimodaux.

Les modèles unimodaux sont spécialisés dans le traitement d'un seul type de données. Par exemple : Les modèles de reconnaissance d'image analysent des données visuelles en extrayant des caractéristiques à partir des pixels afin d'identifier des catégories d'objets. Les modèles de classification audio traitent des données sonores en s'appuyant sur l'analyse du vocabulaire et de la structure syntaxique pour déterminer des catégories textuelles.

Ce type d'approche, limité à une seule modalité, entraîne une acquisition fragmentée de l'information, ce qui réduit leur capacité à comprendre pleinement des situations complexes du monde réel.

### **Modèles multimodaux de grande envergure**

À l'inverse, les grands modèles multimodaux représentent une avancée majeure : ils intègrent et traitent plusieurs types de données hétérogènes simultanément. En dépassant les limites des systèmes unimodaux traditionnels, ces modèles s'appuient sur des architectures et des algorithmes avancés pour permettre un apprentissage conjoint et une compréhension croisée de données issues de multiples sources. Cette approche permet de simuler de manière plus complète et plus précise les processus cognitifs humains.

### **La technologie AcuSeek**

La technologie AcuSeek, présentée dans ce livre blanc, fusionne de manière innovante les caractéristiques visuelles et textuelles afin de permettre une recherche croisée entre plusieurs modalités à partir de descriptions en langage naturel. AcuSeek révolutionne l'interaction avec les systèmes de recherche vidéo en permettant aux utilisateurs d'effectuer des recherches instantanées à partir d'une simple requête textuelle.

### **3.1 L'architecture technique d'AcuSeek**

Les capacités d'AcuSeek reposent sur le modèle d'IA à grande échelle Guanlan développé par Hikvision – un cadre de modélisation multimodale conçu en interne. Guanlan signifie littéralement « observer les vagues de l'eau ». Pour comprendre la nature et le mouvement de l'eau, il faut en observer les ondulations. Ce nom reflète une volonté de s'engager pleinement, de remonter à l'origine des phénomènes et d'en saisir l'essence même.

Le modèle Guanlan de Hikvision se distingue par une perception visuelle avancée et une compréhension fine image-texte. AcuSeek exploite notamment les fonctionnalités suivantes du modèle Guanlan :

- A. Compréhension cross-modale (fusion texte-image)
- B. Modélisation spatio-temporelle pour l'analyse vidéo
- C. Intégration de connaissances spécifiques au domaine de la sécurité
- D. Capacités de préentraînement continu

Grâce aux techniques de distillation de modèle, ces fonctions avancées ont été optimisées pour être déployées aussi bien sur des dispositifs embarqués que sur des serveurs, tout en restant compatibles avec des matériels aux capacités de calcul variées.

#### **Avantages techniques d'AcuSeek**

Porté par le modèle d'IA Guanlan, reconnu pour ses performances en perception visuelle et compréhension multimodale, AcuSeek propose une recherche vidéo efficace à partir de requêtes textuelles. Sa puissance repose sur trois piliers fondamentaux :

#### **Données d'entraînement de haute qualité**

En filtrant de vastes ensembles de données issus d'Internet, nous sélectionnons des échantillons d'apprentissage de grande qualité. Cela permet de réduire la consommation de ressources informatiques, tout en améliorant les performances du modèle et en lui conférant une solide capacité de compréhension générale.

#### **Architecture de modèle centrée sur la sécurité**

Grâce à une expertise approfondie dans les scénarios de sécurité, nous concevons des cadres d'entraînement spécialisés, entièrement développés sur mesure. Ces

modèles dédiés répondent avec précision à la diversité des besoins utilisateurs dans les applications de sécurité.

### Infrastructure d'entraînement optimisée

Notre architecture de cluster d'entraînement distribué est spécifiquement conçue pour traiter des données liées à la sécurité (images et textes), avec les caractéristiques suivantes :

- Répartition intelligente des données entre les nœuds
- Communication inter-clusters économe en ressources
- Cycles d'entraînement accélérés pour un déploiement rapide dans divers scénarios

### 3.2 Fondements techniques d'AcuSeek

L'algorithme d'AcuSeek s'appuie sur un très grand volume de paires image-texte annotées lors de son entraînement. À partir de ces données, des caractéristiques visuelles (par exemple : chien, neige, jaune, fourrure) et des éléments textuels (sous forme de tokens tels que neige/champ/chien) sont extraits grâce à des réseaux neuronaux cross-modaux. Ces éléments constituent la base du modèle de recherche texte-vers-image.

Lors du déploiement, le système analyse les flux vidéo pour détecter des cibles, puis utilise des modèles embarqués pour générer et stocker des vecteurs de caractéristiques encodés en binaire, permettant une mise en correspondance rapide. Lorsqu'un utilisateur saisit une requête textuelle, celle-ci est convertie en vecteurs sémantiques via l'algorithme d'AcuSeek. Une comparaison en temps réel est alors effectuée entre ces vecteurs textuels et les caractéristiques visuelles indexées à l'aide du calcul de similarité cosinus (seuil > 0,85).

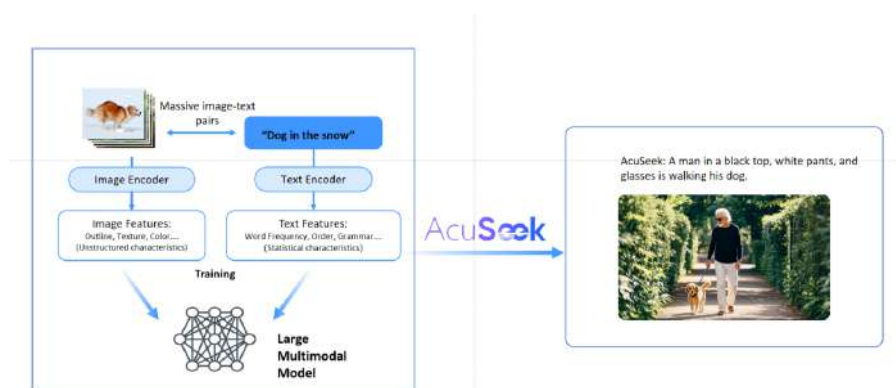


Figure 1. AcuSeek Technology

Ce mécanisme de recherche, optimisé pour les dispositifs en périphérie, permet de contourner la latence liée au cloud, tout en garantissant un haut niveau de précision grâce à un équilibrage adaptatif des seuils et à une architecture de stockage compacte des caractéristiques.

### 3.3 Modes d'application d'AcuSeek

Afin de garantir une utilisation simple et intuitive, l'application AcuSeek intègre une interface interactive bien pensée, mettant en avant les recherches liées aux personnes, véhicules et vélos, complétées par des suggestions du type « Devinez ce que vous cherchez ».

Ces recommandations sont issues d'études de marché sur les termes les plus fréquemment recherchés, avec des formulations validées et précises proposées directement à l'utilisateur.

Fonctionnalités clés :

- Collections personnalisées : les utilisateurs peuvent organiser le contenu de leurs dossiers en fonction de leurs propres scénarios d'utilisation.
- Cohérence multiplateforme : interface de recherche unifiée sur l'interface locale, le web et les différentes plateformes.
- Interaction minimale : une expérience utilisateur fluide, optimisée pour fonctionner efficacement dans les limites des capacités des appareils.



Figure 2. AcuSeek Web page

### 3.4 D'AcuSeek à AcuSearch

Le flux de travail d'AcuSeek repose sur une architecture de recherche en deux étapes. La première phase consiste en une mise en correspondance rapide entre les requêtes

saisies par l'utilisateur (texte) et les cibles indexées en temps réel sur les appareils en périphérie, grâce à un calcul léger de similarité cosinus (seuil > 0,85).

La seconde phase, assurée par AcuSearch, affine les résultats en exploitant une base de données temporelle contenant des vecteurs de caractéristiques historiques associés à des métadonnées spatio-temporelles.

Ce traitement en deux temps permet de :

1. Filtrer rapidement les données à l'aide de caractéristiques binaires compactes, avec une réponse en moins d'une seconde,
2. Améliorer la précision par l'analyse de modèles temporels et la vérification croisée entre caméras dans les enregistrements historiques.

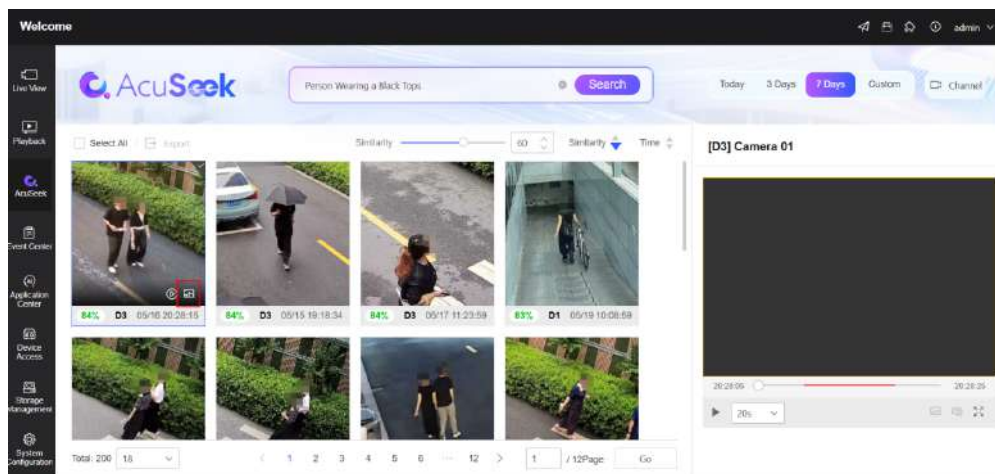


Figure 3. Résultats AcuSeek

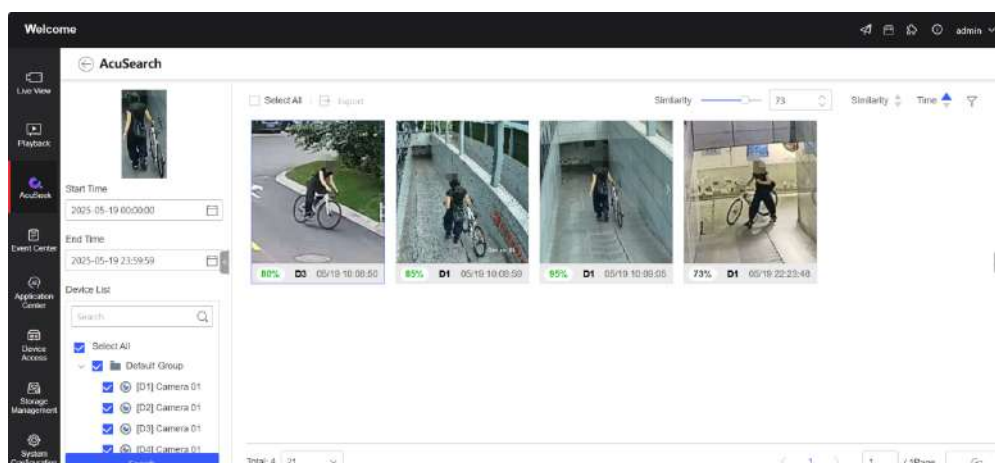


Figure 4. Résultats AcuSeek

## 4. Scénarios d'application

#### **4.1 Résidence connectée | Etablissement de santé connecté**

Dans les résidences et quartiers résidentiels, AcuSeek permet de résoudre efficacement des situations courantes telles que la disparition de trottinettes d'enfants, la présence de chiens non tenus en laisse ou la perte d'objets, grâce à des requêtes en langage naturel comme « garçon sur une trottinette » ou « personne promenant un chien sans laisse ».

Dans les environnements hospitaliers, le système facilite la localisation rapide de situations telles que la disparition de fauteuils roulants ou l'errance de patients, à l'aide de requêtes comme « personne poussant un fauteuil roulant » ou « personne âgée en tenue d'hôpital ».

#### **4.2 Usine intelligente et chantier connecté**

Dans les chantiers de construction, la gestion des pertes d'outils, le contrôle du port du casque de sécurité ou la supervision des opérations de grue nécessitent traditionnellement une mobilisation importante de personnel, toute négligence pouvant mettre des vies en danger. De même, dans les usines, la perte fréquente d'outils de grande valeur exige souvent une récupération rapide afin de limiter les pertes économiques.

Grâce à la solution AcuSeek, les opérateurs peuvent désormais saisir directement des requêtes telles que « personne avec chariot plat », localiser instantanément un outil égaré ou identifier le dernier utilisateur, sans avoir à effectuer de recherches manuelles à grande échelle sur le site.

#### **4.3 Parc intelligent**

Dans les parcs de bureaux, des incidents comme la perte de sacs, d'ordinateurs portables ou de téléphones sont fréquents. Bien que l'interdiction de fumer soit une règle couramment appliquée sur les campus, des infractions surviennent encore régulièrement. Dans ce type de scénario, les utilisateurs peuvent simplement saisir des requêtes en langage naturel comme « femme portant un sac » ou « personne tenant un ordinateur portable » pour localiser instantanément les individus concernés, facilitant ainsi la récupération rapide des objets perdus.

Lorsqu'une requête du type « personne en train de fumer » est saisie, les contrevenants sont immédiatement identifiables, ce qui permet d'appliquer des

sanctions et de réduire l'exposition au tabagisme passif pour les non-fumeurs.

La recherche de véhicules sur un campus constitue également un défi dans les systèmes de vidéosurveillance traditionnels. Grâce à l'analyse vidéo intégrée homme-véhicule, les utilisateurs peuvent considérablement réduire le temps de recherche. Par exemple, pour localiser une voiture dans l'enceinte du parc, il suffit de saisir « Mercedes-Benz rouge » en précisant un passage ou une zone, afin d'identifier le véhicule et retracer avec précision sa trajectoire.

#### **4.4 Commerce connecté**

Dans les centres commerciaux et supermarchés à forte affluence, des situations telles que la perte de valises, de sacs ou d'effets personnels, la séparation d'enfants ou la disparition d'objets appartenant au magasin surviennent fréquemment. Grâce à AcuSeek, les utilisateurs peuvent rapidement effectuer un premier tri dans une quantité massive de données afin de retrouver le dernier moment où l'objet a été vu. Cela permet d'aider efficacement aussi bien le personnel que les clients à récupérer rapidement les biens perdus.

## **5. Intelligence artificielle responsable**

Lors de la transition entre une recherche vidéo unimodale classique et l'utilisation de grands modèles multimodaux comme AcuSeek, il est essentiel de veiller à ce que le système d'IA respecte des principes éthiques et soit conforme aux exigences légales et réglementaires, tout en permettant aux utilisateurs de gérer efficacement les risques liés à son usage.

La responsabilité constitue un principe fondamental qui guide le développement de nos technologies d'IA. AcuSeek intègre un mécanisme de filtrage des requêtes de recherche destiné à empêcher les usages non éthiques. Ce mécanisme fera l'objet de mises à jour continues à travers les versions logicielles et les améliorations du système de filtrage, afin d'assurer une conformité constante.

Les utilisateurs sont invités à surveiller les mises à jour logicielles et à effectuer les mises à niveau dès leur disponibilité pour optimiser les performances et la sécurité d'AcuSeek.

Par ailleurs, chaque requête effectuée dans AcuSeek est enregistrée localement sous forme de journal d'activité hébergé sur l'appareil, ce qui facilite les audits de conformité

pour les utilisateurs.

Principales fonctionnalités de conformité :

- A. Gouvernance éthique de l'IA
- B. Conformité avec les normes RGPD, CCPA et ISO 27001 relatives à la protection des données.
- C. Filtrage dynamique
- D. Blocage en temps réel des requêtes interdites (ex. : usage abusif de données biométriques, termes discriminatoires).
- E. Traçabilité des requêtes
- F. Journal d'audit infalsifiable, avec horodatage chiffré (SHA-256) garantissant l'intégrité des données.
- G. Contrôle d'accès basé sur les rôles (RBAC)

Gestion des droits différenciés par niveaux d'utilisateur (administrateur / standard) pour limiter l'accès aux requêtes sensibles.

Restez informés des nouvelles versions et des certifications de conformité via notre portail dédié.