

# Il potere dei metadati della scena

Per un quadro della situazione, un'efficienza e  
approfondimenti migliori

Gennaio 2024

## Sommario

Nel contesto della videosorveglianza, i metadati descrivono in modo testuale i contenuti a video. Ad esempio, possono indicare quali oggetti di interesse sono visibili o descrivere la scena in modo molto approfondito. Inoltre, possono includere attributi associati agli oggetti o alla scena, come i colori dei veicoli e degli abiti, le posizioni esatte o la direzione di marcia. I metadati vengono creati in tempo reale, direttamente sulla telecamera o da un altro componente del sistema in grado di eseguire analitiche video.

Aggiungono anche un contesto agli eventi e consentono di ordinare ed esaminare rapidamente grandi quantità di filmati. A grandi linee, le funzioni che è possibile utilizzare grazie ai metadati possono essere classificate in tre aree:

- **Ricerche forensi dopo un evento**, per cercare oggetti o eventi di interesse in base a vari parametri che restringono la ricerca a un numero limitato di candidati. I dati di classificazione degli oggetti consentono di effettuare ricerche che prevedono un'ampia gamma di dettagli.
- **Uso in tempo reale**, per aiutare gli operatori a reagire rapidamente a una variazione della situazione. Oppure, possono dare un input per supportare il processo decisionale o consentire azioni automatiche.
- **Identificazione di tendenze, modelli e approfondimenti**. Per compilare report statistici, le piattaforme IoT e quelle dedicate all'efficienza operativa possono acquisire metadati tramite il conteggio dei visitatori, la misurazione della velocità, i dati sui flussi di traffico e altri tipi di raccolta automatizzata.

Alcune telecamere possono decodificare l'audio per recuperare metadati. I modelli acustici specifici possono essere rilevati ed etichettati in modo simile a quanto avviene per le classi di oggetti nel video. Ad esempio, un sistema di riconoscimento audio può identificare un'aggressione verbale o la rottura di un vetro.

Combinando i metadati provenienti da più input, ad esempio fonti visive, audio, relative alle attività e ai processi, si hanno molte più informazioni dettagliate di quante se ne otterrebbero da ogni singola fonte. I protocolli aperti e gli standard di settore sono essenziali per un'integrazione perfetta dei metadati.

# Sommario

1	Introduzione	4
2	Cosa sono i metadati?	4
3	Generazione di metadati in modalità edge	4
4	Applicazioni	5
	4.1 Uso in tempo reale per un intervento immediato	5
	4.2 Ricerca forense	5
	4.3 Identificazione di tendenze e modelli per approfondire	6
5	Dove vengono utilizzati i metadati?	6
6	Come vengono consegnati i metadati?	7
7	Metadati audio	9
8	Combinazione di metadati da più fonti	9

# 1 Introduzione

I metadati rappresentano il punto di partenza per raccogliere informazioni utili dai video. Assegnano un significato digitale ai contenuti descrivendo i dettagli principali della scena. Utilizzando i metadati, è possibile trovare e valutare contenuti importanti in grandi quantità di filmati, intervenendo come opportuno. Questo è il motivo per cui i metadati sono diventati un elemento sempre più importante per la sicurezza e l'efficienza delle operazioni aziendali.

Questo documento tecnico tratta dei metadati sia nel contesto della sorveglianza che in quello dell'efficienza operativa. Spiega in dettaglio quali sono i vantaggi dei metadati e come vengono utilizzati nei sistemi di gestione video e in altre applicazioni.

## 2 Cosa sono i metadati?

I metadati sono dati su altri dati. Nel contesto della videosorveglianza, indicano in modo testuale i contenuti del video, ad esempio quali oggetti di interesse sono visibili, o descrivono in modo molto approfondito la scena. La descrizione può includere attributi associati agli oggetti o alla scena, come i colori dei veicoli e degli abiti, le posizioni esatte o la direzione di spostamento. I metadati vengono creati in tempo reale, direttamente sulla telecamera o da un altro componente del sistema in grado di eseguire analitiche video.

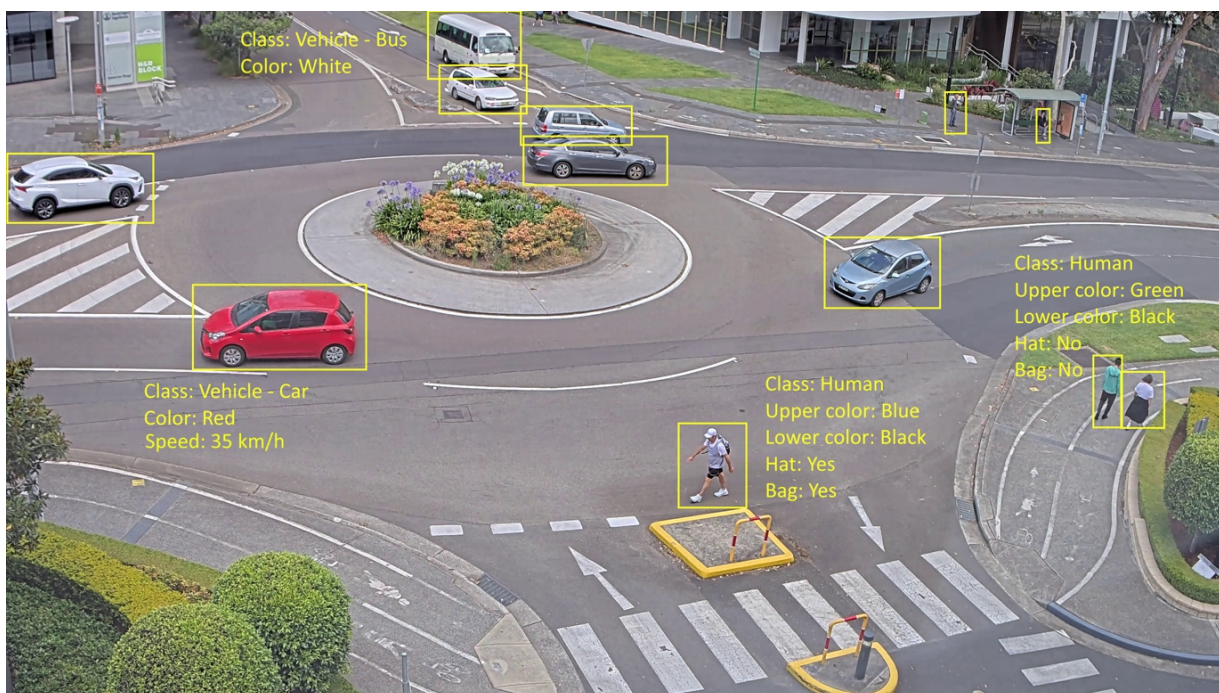


Figure 1. Esempio di fotogramma in cui gli oggetti vengono rilevati e analizzati per creare metadati.

## 3 Generazione di metadati in modalità edge

Le analitiche video ad alte prestazioni venivano eseguite su server perché in genere richiedevano una potenza di calcolo maggiore di quella di un dispositivo edge. Negli ultimi anni, però, lo sviluppo degli algoritmi e la maggiore potenza di calcolo dei dispositivi hanno permesso di eseguire analitiche avanzate in modalità edge. Questo significa che i metadati vengono generati sul dispositivo e possono essere utilizzati direttamente da altre analitiche. Il flusso video e il flusso di metadati possono anche essere trasmessi al VMS o a un'altra applicazione per un'ulteriore elaborazione.

Le analitiche edge hanno accesso al materiale video non compresso con una latenza molto bassa. In questo modo è possibile utilizzare applicazioni veloci in tempo reale, evitando costi aggiuntivi e complessità legate allo spostamento di tutti i video da elaborare in altre parti del sistema. Le analitiche edge-based hanno anche costi dell'hardware e di utilizzo più bassi perché riducono le risorse necessarie sul server del sistema.

Generare metadati in modalità edge significa estrarre dati dal video senza perdere informazioni durante la compressione o la trasmissione. Così facendo, si hanno metadati più precisi e un'analisi più precisa dei contenuti video. Migliore è la qualità dell'immagine, migliori sono i metadati.

## 4 Applicazioni

I metadati non forniscono solo dettagli sugli oggetti di una scena. Aggiungono anche un contesto agli eventi e consentono di ordinare ed esaminare rapidamente grandi quantità di filmati. In genere, le funzioni utilizzabili grazie ai metadati possono essere classificate in tre aree: ricerche forensi dopo un evento, uso in tempo reale e identificazione di tendenze, modelli e approfondimenti.

### 4.1 Uso in tempo reale per un intervento immediato

I metadati possono essere utilizzati in tempo reale per aiutare gli operatori a reagire velocemente quando una situazione cambia. Possono anche dare un input prezioso per supportare il processo decisionale o consentire azioni automatiche. Le analitiche edge in tempo reale che funzionano con metadati di alta qualità possono aiutare a proteggere persone, siti ed edifici e metterli al riparo da danni intenzionali o accidentali. È possibile rilevare, verificare e valutare rapidamente le minacce in modo da poterle gestire con efficienza.

### 4.2 Ricerca forense

I metadati consentono di cercare in modo rapido ed efficiente oggetti o eventi di interesse. Questo può far risparmiare molte ore agli investigatori, soprattutto quando devono esaminare grandi quantità di video provenienti da più fonti. È possibile cercare oggetti, come persone e veicoli, in base a vari parametri per restringere la ricerca a un numero limitato di candidati. Ad esempio, i parametri di ricerca possono includere movimenti, orari e caratteristiche degli oggetti.

Tabella 4.1 Le ricerche vengono agevolate utilizzando varie categorie di metadati.

Categoria di metadati	Cosa viene rilevato?	Esempi
Movimento	Come si muove un oggetto?	Direzione, velocità, altri comportamenti
Tempi	Quando compare un oggetto?	Giorno della settimana, ora del giorno, tempo di permanenza
Luogo	Dove si trova l'oggetto?	Luogo, campo visivo della telecamera
Classificazione degli oggetti	Che tipo di oggetto è?	Persone, veicoli (auto, autobus, camion, bicicletta/moto)
Attributi dell'oggetto	Quali caratteristiche ha?	Abbigliamento, accessori come cappelli o borse, caratteristiche fisiche come il colore degli abiti

Anche l'accesso a una sola categoria di metadati, ad esempio quella temporale, può rivelarsi decisivo per trovare i risultati desiderati.

I metadati di movimento consentono di effettuare ricerche in base alla velocità relativa dell'oggetto e alla sua direzione di marcia. I dati di classificazione degli oggetti consentono di effettuare ricerche che prevedono dettagli più ampi. In genere, le telecamere con processore per deep learning (DLPU) possono fornire metadati con una classificazione degli oggetti più precisa, consentendo ad esempio di cercare un camion verde o una persona con un cappotto blu.

### **4.3 Identificazione di tendenze e modelli per approfondire**

Per compilare report statistici, le piattaforme IoT e quelle dedicate all'efficienza operativa possono acquisire metadati tramite il conteggio dei visitatori, la misurazione della velocità, i dati sui flussi di traffico e altri tipi di raccolta automatizzata. I dati vengono analizzati per generare informazioni e approfondimenti utili.

## **5 Dove vengono utilizzati i metadati?**

Sono molti i vantaggi che si ottengono utilizzando i metadati per comprendere le caratteristiche e il contenuto di una scena. I principali consumatori di metadati possono essere classificati come segue.

**Applicazioni edge.** Le analitiche in esecuzione sulla telecamera possono applicare regole e filtri logici alle informazioni sugli oggetti della scena. In questo modo, possono attivare azioni basate su soglie definite o comportamenti specifici, controllando ad esempio una telecamera PTZ in base al rilevamento e al movimento di una persona nella scena.

**Software di gestione video (VMS).** Nel contesto della videosorveglianza, i metadati sono stati comunemente utilizzati all'interno di un VMS per visualizzare sovrimpressioni intorno ai potenziali oggetti di interesse nella scena. Con lo sviluppo di algoritmi di rilevamento e classificazione degli oggetti più avanzati, ora gli operatori possono anche localizzare gli oggetti in base a caratteristiche specifiche, come il colore degli indumenti. La possibilità di eseguire ricerche con questi dati riduce notevolmente la necessità di riguardare manualmente grandi quantità di filmati.

**Piattaforme IoT.** I metadati possono essere raccolti e presentati visivamente nelle piattaforme di business intelligence per generare informazioni utili analizzando tendenze in tempo reale e storiche. Le analisi statistiche basate sul flusso o sull'esperienza dei clienti consentono di prendere decisioni basandosi sui dati, in modo da migliorare le operazioni.

**Analitiche di secondo livello.** Alcune applicazioni richiedono un'elaborazione combinata (edge-based e server-based) per eseguire analisi più avanzate. La pre-elaborazione può essere eseguita sulla telecamera, con l'elaborazione successiva su server. Un sistema ibrido di questo tipo può agevolare la scalabilità

delle analitiche, che è economicamente vantaggiosa perché solo i video e i metadati rilevanti vengono trasmessi al server.

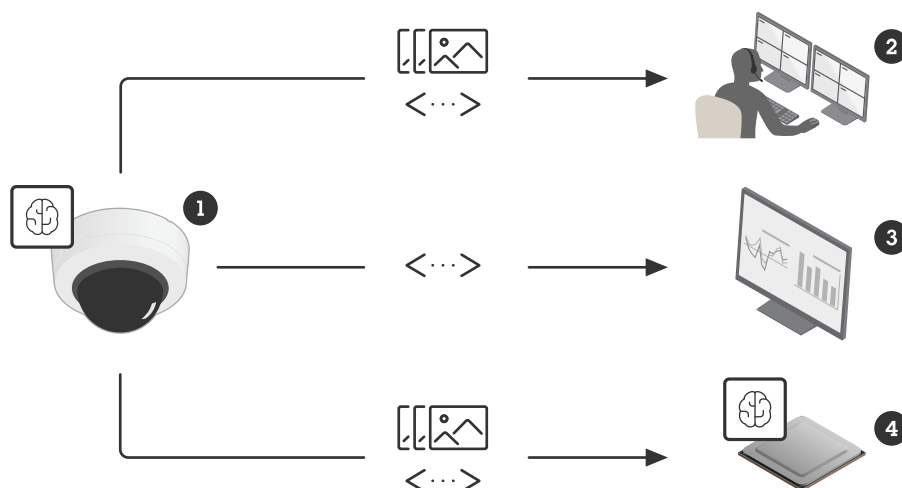


Figure 2. Consumatori di metadati

- 1 Applicazioni edge
- 2 VMS
- 3 Piattaforme IoT
- 4 Analitiche di secondo livello

## 6 Come vengono consegnati i metadati?

I metadati possono essere consegnati utilizzando approcci diversi in base agli usi previsti. Nelle applicazioni in tempo reale, i metadati devono essere costantemente trasmessi al consumatore su richiesta, perché ciò è fondamentale per garantire una risposta adeguata e un quadro corretto della situazione. In altre applicazioni meno critiche, che non richiedono un intervento in tempo reale, i metadati possono essere consolidati ulteriormente, ad esempio in base alla traccia di ciascun oggetto specifico nella scena, prima

di essere consegnati al consumatore. In questo modo si riduce la quantità totale di dati da archiviare ed elaborare.

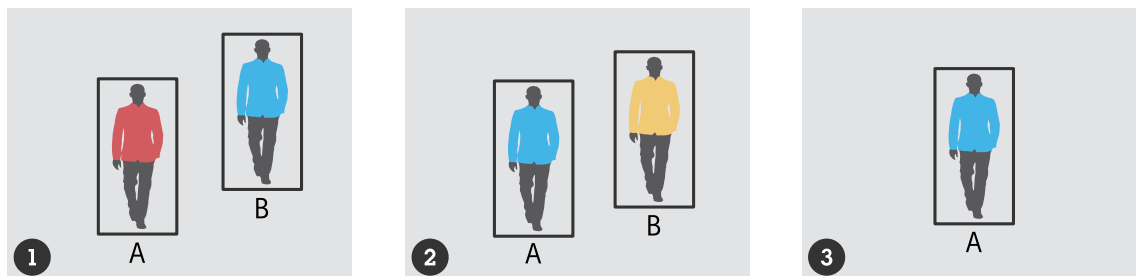


Figure 3. La figura illustra lo streaming dei metadati. I fotogrammi continui con metadati provenienti dalla telecamera forniscono informazioni sulla scena in tempo reale. Ogni fotogramma riprende la scena in un momento specifico, indipendentemente dagli eventi precedenti.

- 1 Il fotogramma 1 rileva gli oggetti A e B, classificando A come persona vestita di rosso e B come persona vestita di blu.
- 2 Nel fotogramma 2, la telecamera aggiorna la classificazione, determinando che l'oggetto A indossa effettivamente abiti blu e l'oggetto B indossa abiti gialli. Anche se gli oggetti rimangono identici al fotogramma 1, i loro attributi di colore cambiano e questo si riflette nei metadati.
- 3 Il fotogramma 3 mostra l'assenza dell'oggetto B. La telecamera segue solo l'oggetto A, ancora classificato come persona vestita di blu.

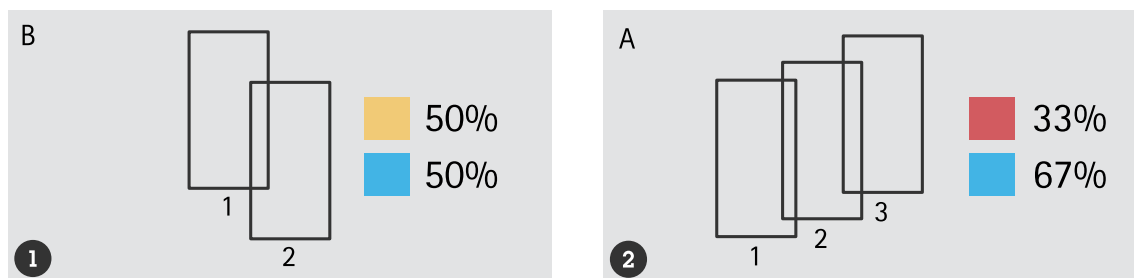


Figure 4. La figura illustra la distribuzione consolidata di metadati: la telecamera fornisce informazioni in formato unificato in base alla traccia degli oggetti nella scena. I fotogrammi di ciascun oggetto comprendono tutti i dettagli noti per tutta la durata della traccia.

- 1 Nel primo fotogramma vengono presentati i dettagli sull'oggetto B, incluso il primo e l'ultimo rilevamento, il riepilogo della traiettoria e gli attributi rilevati durante la traccia. L'oggetto B aveva una probabilità del 50% di indossare abiti gialli e una probabilità del 50% di indossare abiti blu.
- 2 Il secondo fotogramma rispecchia questo formato per l'oggetto A, rivelando una probabilità del 33% di abiti rossi e una probabilità del 67% di abiti blu.

Il vantaggio del metodo consolidato è che la telecamera riduce significativamente la quantità di dati inviati al consumatore, trasmettendoli solo quando sono presenti oggetti nella scena; in questo caso, i metadati



vengono riepilogati (consolidati) per agevolare l'interpretazione. Il metodo in streaming fornisce una descrizione completa della scena in ogni fotogramma, anche quando non sono presenti attività o oggetti e il consumatore deve interpretare i dati in base alle proprie esigenze specifiche. Come accennato, il metodo in streaming è vantaggioso per le applicazioni in tempo reale, mentre quello consolidato è ottimale per la post elaborazione quando il consumatore non deve eseguire azioni immediate.

Comprendere i punti di forza e i limiti di ciascun approccio è fondamentale per progettare l'architettura del sistema. Ad esempio, per una piattaforma IoT che genera approfondimenti basati sui metadati è vantaggioso ricevere un riepilogo post evento degli oggetti nella scena, perché in genere è soggetta a limiti di larghezza di banda e spazio di archiviazione.

Inoltre, i metadati possono essere forniti utilizzando protocolli di comunicazione e formati di file diversi in base alle esigenze e alle preferenze specifiche del consumatore.

## **7 Metadati audio**

Alcune telecamere possono decodificare l'audio per recuperare metadati. Le analitiche di riconoscimento audio sono in grado di rilevare modelli acustici ed evidenziare suoni di interesse nell'audio in diretta e registrato. In questo modo, i sistemi di riconoscimento audio abbinati ai dispositivi di videosorveglianza possono segnalare agli operatori potenziali eventi in corso, guidandoli verso le relative telecamere. Il sistema, ad esempio, può identificare un'aggressione verbale per evitare che la situazione degeneri, rilevare la rottura di vetri per prevenire effrazioni o trasmettere avvisi se un paziente è in difficoltà. Poiché gli operatori possono non solo vedere ma anche sentire ciò che accade in una scena, i sistemi di riconoscimento audio possono consentire un rilevamento precoce, un intervento rapido e, in molti casi, prevenire che una situazione peggiori. Il riconoscimento audio può anche essere utile come strumento di verifica secondario.

In genere, le analitiche addestrate con modelli acustici valutano una combinazione di caratteristiche come il livello di decibel e l'energia delle varie frequenze nel tempo. I modelli acustici specifici possono essere rilevati ed etichettati in modo simile a quanto avviene per le classi di oggetti nel video.

## **8 Combinazione di metadati da più fonti**

Le vere potenzialità dei metadati si esprimono quando vengono applicati a più input, ad esempio visivi, audio, relativi alle attività o ai processi. Fonti di dati come il tracciamento RFID, le coordinate GPS, gli avvisi di manomissione, i contatori (di temperatura, livelli chimici, ecc.), i rilevatori di rumore e i dati sulle transazioni di vendita sono molto utili nella gestione di qualunque sito. I dati provenienti da tutte le fonti possono essere allineati in base ai relativi timestamp.

Combinare metadati provenienti da fonti diverse significa avere molte più informazioni di quante se ne possano ottenere dalle singole fonti. I protocolli aperti e gli standard di settore sono essenziali per un'integrazione perfetta dei metadati.

# Informazioni su Axis Communications

Axis consente un mondo più intelligente e più sicuro creando soluzioni per migliorare la sicurezza e le prestazioni aziendali. Come società di tecnologie di rete e leader nel settore, Axis offre soluzioni nella videosorveglianza, controllo degli accessi, interfono e sistemi audio. Queste sono ottimizzate da applicazioni di analisi intelligente e supportate da formazione di alta qualità.

Axis ha circa 4.000 impiegati dedicati in più di 50 paesi e collabora con partner di tecnologia e integrazione di sistema in tutto il mondo per offrire soluzioni di clienti. Fondata nel 1984, Axis è con sede a Lund, in Svezia