



# Analisi video accelerate all'edge per le città IA

HPE Edgeline e NVIDIA IVA: acquisizione rapida di insight per le funzionalità di intelligenza artificiale all'edge

## Sommario

Perché scegliere l'analisi video all'edge.....	2
Panoramica della soluzione .....	4
Sistemi edge convergenti HPE: una nuova categoria di prodotti .....	4
Sistemi edge convergenti HPE Edgeline .....	4
GPU NVIDIA e analisi video intelligente.....	5
XJERA LABS .....	6
Diagramma dell'architettura .....	7
Obiettivi dell'analisi video e configurazione del sistema.....	8
Obiettivi del test .....	8
Configurazione hardware .....	8
Configurazione software .....	9
Video di input.....	9
Prestazioni dell'analisi video.....	10
Dimensioni del sistema.....	11
Conclusione.....	12
Notifiche.....	13

## Perché scegliere l'analisi video all'edge

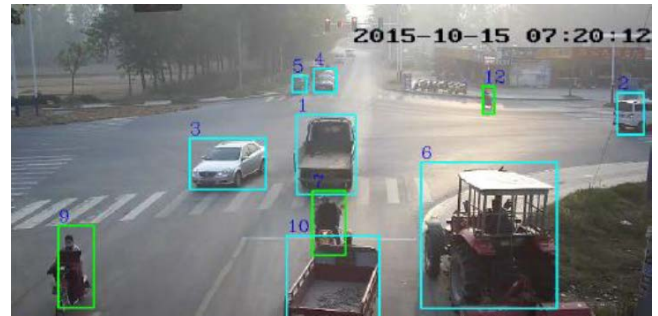
Trenta miliardi di immagini al secondo, pari a centomila miliardi di immagini all'ora!

Ecco l'incredibile quantità di contenuti che verranno acquisiti entro il 2020 dalle telecamere di sorveglianza in tutto il mondo. Questo miliardo di telecamere, il doppio rispetto ad oggi, si troveranno agli incroci stradali, nelle stazioni di transito e in altri luoghi pubblici, da dove contribuiranno a rendere le nostre città più sicure e intelligenti: come elementi fondamentale di qualsiasi iniziativa Smart City o di città IA<sup>1</sup>, saranno posizionate anche nei negozi, nei centri di assistenza, nei magazzini e in altri luoghi ancora per garantire protezione e sicurezza e raccogliere informazioni con cui incrementare le vendite, monitorare le scorte e migliorare il servizio.

Per dare un senso a questo numero sbalorditivo di pixel, i metodi tradizionali di elaborazione video che utilizzano il monitoraggio umano o gli algoritmi di visione artificiale "artigianali" non saranno in grado di soddisfare le esigenze di velocità e precisione. Invece, una branca dell'intelligenza artificiale (IA) chiamata deep learning offre un metodo potente e scalabile per estrarre informazioni dettagliate da questi sistemi di visione. Il deep learning richiede tuttavia ordini di grandezza con più elaborazione e memoria di quanto sia normalmente disponibile nei sistemi di server generici. Gli acceleratori NVIDIA® Tesla sono stati appositamente sviluppati per offrire le prestazioni di deep learning necessarie in un fattore di forma compatto a basso consumo energetico.



**Figura 1.** Informazioni estratte con la visione artificiale tradizionale<sup>2</sup>



**Figura 2.** Informazioni estratte con la visione artificiale del deep learning

Il deep learning è il settore dell'apprendimento automatico che cresce più rapidamente. Questa tecnologia utilizza numerose reti neurali a più strati (DNN) per apprendere livelli di rappresentazione e astrazione che danno un senso a dati come immagini, testo e suoni: le caratteristiche specifiche di un'immagine, come nell'esempio della **Figura 3**, vengono estratte e rese comprensibili a basso, medio e alto livello da una specifica rete neurale artificiale (ANN). Rispetto alla tradizionale visione artificiale artigianale, l'ANN è in grado di cogliere automaticamente caratteristiche sottili e importanti grazie all'apprendimento effettuato a partire da grandi volumi di set di dati etichettati, con una convalida sufficiente a fornire un livello di intelligenza più elevato per inferenza da immagini non elaborate. Un esempio dello scenario comparativo è riportato nella **Figura 1** e nella **Figura 2**, con livelli più elevati di precisione nel rilevamento sia per l'approccio tradizionale sia per quello di deep learning. Il risultato è un livello di precisione e prestazioni che supera di gran lunga le capacità umane. La soluzione NVIDIA Intelligent Video Analytics (IVA)<sup>3</sup>, completamente supportata sui sistemi edge convergenti HPE Edgeline con schede NVIDIA Tesla, consente a qualsiasi sviluppatore software di sfruttare facilmente la potenza del deep learning nella sua applicazione.

<sup>1</sup> NVIDIA AI City: [blogs.nvidia.com/blog/2016/09/20/ai-city](https://blogs.nvidia.com/blog/2016/09/20/ai-city)

<sup>2</sup> Presentazione Hikvision 2016 GTC: [on-demand.gputechconf.com/gtc/2016/presentation/s6840-shiliang-pu-intelligent-video.pdf](https://on-demand.gputechconf.com/gtc/2016/presentation/s6840-shiliang-pu-intelligent-video.pdf)

<sup>3</sup> Piattaforma NVIDIA Intelligent Video Analytics (IVA): [nvidia.com/object/intelligent-video-analytics-platform.html](https://nvidia.com/object/intelligent-video-analytics-platform.html)

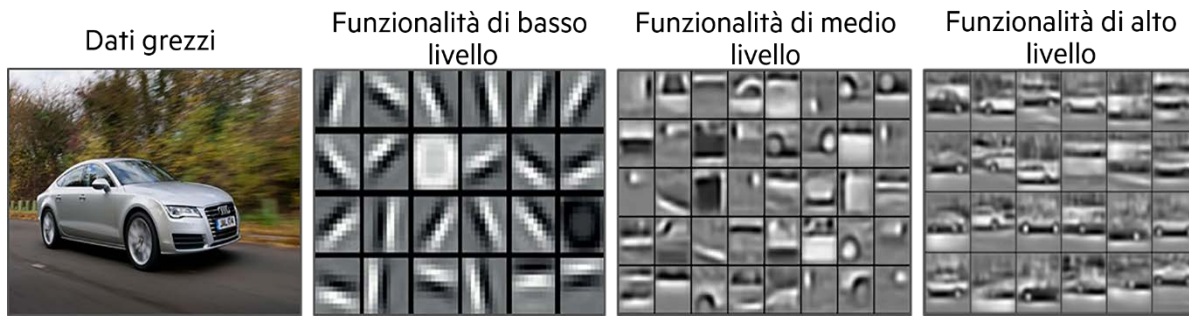


Figura 3. Come la rete neurale artificiale (ANN) vede un oggetto

La soluzione IVA è senza dubbio eccezionale nell'estrazione delle informazioni chiave, ma in numerosi scenari di sicurezza fisica e tecnica o di prevenzione dei guasti la velocità di reazione è altrettanto fondamentale per affrontare efficacemente la situazione. Ad esempio, la risposta a un grave incidente che ha coinvolto molti veicoli nell'ora di punta può richiedere l'invio rapido di polizia, ambulanze e mezzi antincendio, la modifica della segnaletica stradale per deviare il traffico e l'archiviazione automatica dei filmati preroll e postroll come prove per le indagini.

Il trasferimento in un data center tradizionale o nel cloud di ogni bit di dati video proveniente da migliaia di videocamere all'edge si rivela spesso:

1. troppo lento, causando una notevole latenza di reazione
2. troppo costoso, a causa della larghezza di banda elevata necessaria per tutti i feed video
3. ad alto rischio di danneggiamento o snooping se è coinvolta una rete pubblica.

Tuttavia, tutti questi problemi possono essere risolti facilmente tramite i solidi e compatti sistemi edge convergenti HPE Edgeline per l'elaborazione dei feed delle videocamere direttamente all'edge, in tempo reale, e con tutti i vantaggi delle reti neurali artificiali (ANN) accelerate da NVIDIA Tesla.

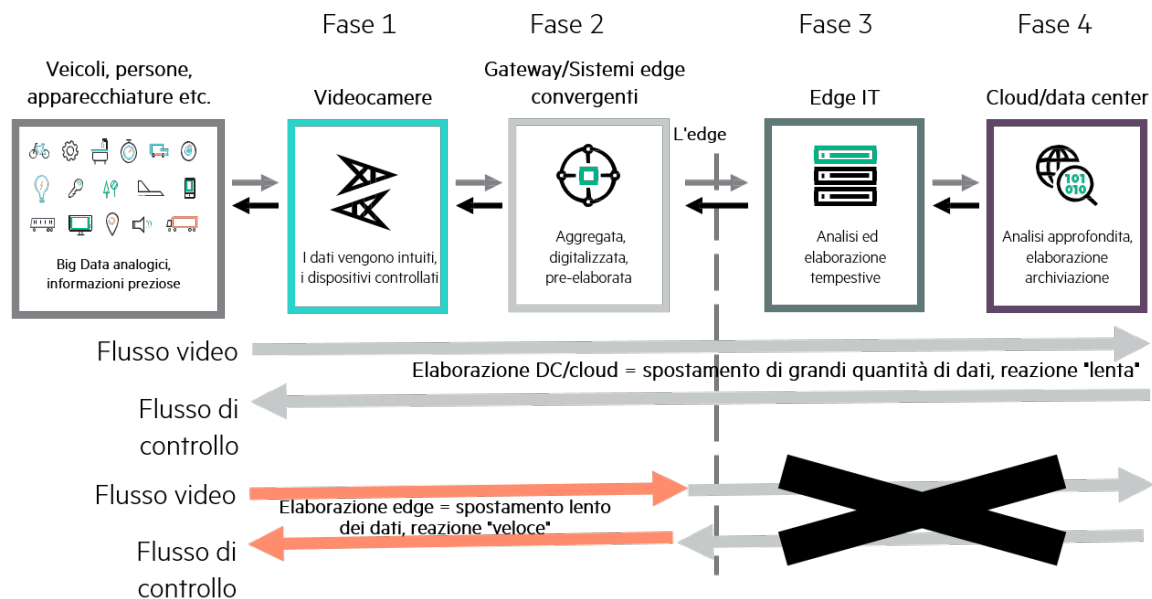


Figura 4. L'elaborazione edge determina una reazione più rapida agli eventi, un minore trasferimento di dati e rischi contenuti

## Panoramica della soluzione

L'analisi video è la dimostrazione perfetta del valore dell'edge computing, dal momento che l'elaborazione dei filmati vicino alla videocamera elimina la necessità di trasferire grandi volumi di dati video nel data center o nel cloud, contiene i costi della larghezza di banda, accelera notevolmente i tempi di reazione e riduce il rischio di danneggiamento o di spionaggio. I sistemi edge convergenti HPE Edgeline e NVIDIA Tesla P4 IVA consentono di acquisire gli insight più rapidamente per la videosorveglianza all'edge con funzionalità IA. Questi componenti hardware possono essere abbinati ad analisi video accuratamente progettate e basate sulle ANN, come quelle di XJERA LABS, per una piattaforma di analisi video edge estremamente intelligente, efficiente ed economicamente vantaggiosa. L'inclusione della connettività wireless HPE Aruba semplifica la scalabilità del sistema da decine a migliaia di videocamere, mentre i nostri software di sicurezza IoT leader di settore, come ClearPass e Niara, tengono a debita distanza le minacce e i potenziali dispositivi edge non autorizzati.

### Sistemi edge convergenti HPE: una nuova categoria di prodotti

I sistemi edge convergenti HPE, una nuova categoria di prodotti di settore in cui Hewlett Packard Enterprise è all'avanguardia, sono appositamente progettati per far convergere l'acquisizione dei dati in tempo reale, il controllo, l'analisi di classe enterprise all'edge e le funzionalità di gestione in remoto. HPE Edgeline è il primo prodotto di settore della categoria. Queste piattaforme robuste, compatte ed efficienti dal punto di vista energetico offrono un'ampia gamma di opzioni di connettività di rete e acquisizione o controllo dei dati per rispondere a quasi tutti i casi d'uso di elaborazione video all'edge. Il tipo di chassis, il numero di server, di GPU NVIDIA e di core CPU, la memoria e lo storage sono tutti adattabili al numero preciso e alla qualità dei feed video e all'intensità delle analisi necessarie.

### Sistemi edge convergenti HPE Edgeline

Il sistema edge convergente HPE Edgeline EL1000 supporta due schede GPU NVIDIA, una cartuccia server, come la potente HPE m510 che contiene fino a 16 core Intel® Xeon®, 128 GB di RAM, fino a 2 TB di unità SSD NVMe ultraveloci e una scheda NIC 10GbE a due porte. Può ospitare 2 schede PCIe full-height half-length (FHHL) come la scheda NVIDIA Tesla P4 utilizzata in questo test, 2 unità SATA SFF e anche un'ampia gamma di opzioni di connessione wireless (ad esempio Wi-Fi e LTE); è inoltre sufficientemente robusto da operare in ambienti ostili, con un intervallo di temperature da 0°C a 55°C.<sup>4</sup>



Figura 5. Sistema edge convergente HPE Edgeline EL1000

<sup>4</sup> L'estensione delle temperature operative è disponibile solo quando vengono utilizzate le opzioni supportate. Per maggiori dettagli, consultare la documentazione del prodotto: [hpe.com/info/edgeline](https://hpe.com/info/edgeline)

Il sistema edge convergente HPE Edgeline EL4000 supporta più schede GPU NVIDIA, fino a quattro cartucce server indipendenti, inclusi i server HPE m510, per un imponente Intel Xeon a 64 core, 512 GB di memoria, 8 TB di unità SSD e 8 porte 10 GbE in un fattore di forma 1U sottile. Può accogliere le schede NVIDIA Tesla P4<sup>5</sup>, ognuna connessa a una cartuccia server, offrendo un'incredibile potenza di elaborazione per eseguire velocemente gli algoritmi di analisi video più impegnativi all'edge. I doppi alimentatori ridondanti, l'irrobustimento (fino a MIL-STD attraverso partner HPE) e il supporto delle certificazioni del settore come NEBS Level 3 lo rendono un sistema estremamente affidabile, allo stesso livello delle apparecchiature server utilizzate da Telcos e altri operatori mission-critical. Sia HPE EL1000 che EL4000 supportano svariate opzioni di installazione adatte a qualsiasi ambiente: rack, a muro, scrivania o persino nelle apparecchiature del cliente.



Figura 6. Sistema edge convergente HPE Edgeline EL4000

### GPU NVIDIA e analisi video intelligente

Le GPU NVIDIA Tesla e la soluzione IVA (Intelligent Video Analytics) sfruttano la rivoluzionaria architettura NVIDIA Pascal™. La scheda NVIDIA Tesla P4<sup>6</sup> è appositamente realizzata per incrementare l'efficienza di elaborazione per i server scalabili in orizzontale che eseguono carichi di lavoro di deep learning,<sup>7</sup> rendendo possibili servizi basati su intelligenza artificiale avanzati e reattivi. Riduce la latenza per l'inferenza di 15 volte in qualsiasi infrastruttura hyperscale e offre un'incredibile efficienza energetica, 60 volte superiore alle CPU. (Cfr. la Figura 8). Ne deriva una nuova generazione di servizi di intelligenza artificiale, prima impossibili a causa delle limitazioni di latenza.

L'apprendimento automatico tradizionale ricorre all'estrazione artigianale delle caratteristiche e ad algoritmi specifici della modalità per etichettare le immagini o riconoscere le voci. Tuttavia, questo metodo presenta diversi svantaggi in termini di tempo e precisione. Oggi le reti neurali profonde avanzate utilizzano algoritmi, dati in streaming o video di grandi dimensioni, nonché la potenza di elaborazione della GPU, per modificare questa dinamica. La velocità, la precisione e la portata dell'apprendimento delle macchine hanno ormai raggiunto un livello tale da rendere possibile la vera intelligenza artificiale<sup>8</sup> e l'elaborazione IA.<sup>9</sup>

<sup>5</sup> Il test è stato effettuato su una sola scheda NVIDIA Tesla P4. È previsto un aggiornamento del sistema HPE Edgeline EL4000 che in futuro ospiterà fino a quattro schede NVIDIA Tesla P4.

<sup>6</sup> [images.nvidia.com/content/pdf/tesla/184457-Tesla-P4-Datasheet-NV-Final-Letter-Web.pdf](https://images.nvidia.com/content/pdf/tesla/184457-Tesla-P4-Datasheet-NV-Final-Letter-Web.pdf)

<sup>7</sup> [developer.nvidia.com/deep-learning](https://developer.nvidia.com/deep-learning)

<sup>8</sup> [blogs.nvidia.com/blog/2016/01/12/accelerating-ai-artificial-intelligence-gpus](https://blogs.nvidia.com/blog/2016/01/12/accelerating-ai-artificial-intelligence-gpus)

<sup>9</sup> [nvidia.com/object/ai-computing.html](https://nvidia.com/object/ai-computing.html)

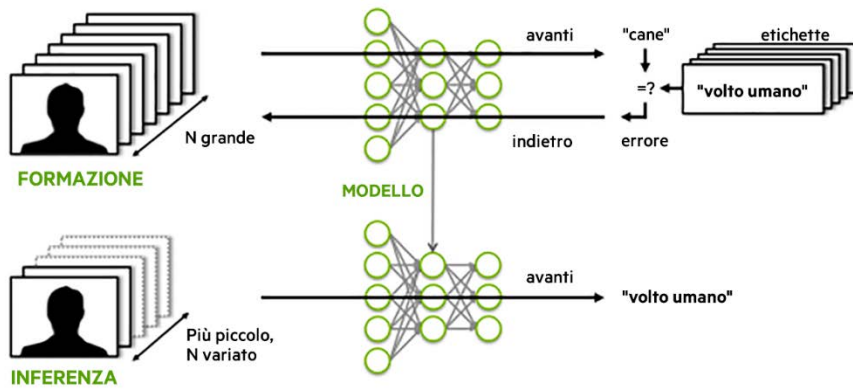
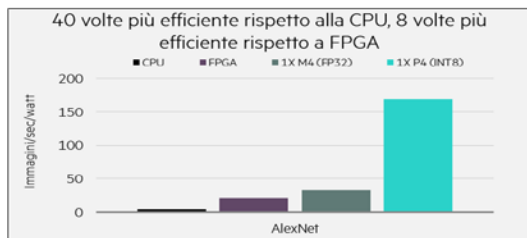


Figura 7. Il deep learning impara utilizzando grandi quantitativi di dati per offrire precisione straordinaria e inferenze rapide

I sistemi HPE Edgeline basati su ANN possono guardare, ascoltare e comprendere cosa accade su più videocamere, a velocità straordinarie; saranno in grado di riunire enormi set di dati provenienti da punti diversi della città e di trasformare rapidamente i semplici pixel in insight preziosi. La città IA che ne deriverà avrà la capacità di riconoscere e rispondere automaticamente in caso di incidente, smarrimento di bambini o animali o situazioni problematiche. La città IA diventa un robot pensante, con molti occhi, programmato per garantire la sicurezza e la felicità delle persone, per far funzionare senza problemi le apparecchiature e per molto altro ancora.

L'acceleratore NVIDIA Tesla P4 che supporta tutto questo all'edge offre i vantaggi e le funzionalità che seguono.

1. Reattività con inferenza in tempo reale
2. Sblocco dei nuovi servizi video basati su IA con un motore di decodifica dedicato
3. Efficienza senza precedenti per server a basso consumo e scalabilità orizzontale
4. Distribuzione più veloce con TensorRT e DeepStream SDK
5. Migliori prestazioni con funzionalità INT8



P4	
Numero di core CUDA	2560
Precisione di picco singola	5,5 teraflop
INT8 di picco	22 al primo posto
Bassa precisione	Prodotto scalare vettoriale 8 bit 4X con accumulazione a 32 bit
Motori video	Motore di decodificazione 1X, motore di decodificazione 2X
Memoria GDDR5	8 GB a 192 GB/s
Alimentazione	50 W e 75 W

AlexNet, dimensioni dei batch = 128, CPU: Intel E5-2690v4 che utilizza Intel MKL 2017, FPGA è Arria10-115 1x M4/P4 in nodo, alimentazione scheda P4 a 56 W, alimentazione GPU P4 a 36 W, alimentazione scheda M4 a 57 W, alimentazione GPU M4 a 39 W, grafico Perf/W che utilizza l'alimentazione GPU

Figura 8. NVIDIA Tesla P4 - progettata per l'inferenza in tempo reale e un'esperienza d'uso intelligente con i sistemi edge convergenti HPE

## XJERA LABS

XJERA LABS è specializzata nelle soluzioni per l'intelligenza artificiale (IA) che supportano l'analisi rivoluzionaria di contenuti video e immagini per svariate applicazioni commerciali. L'azienda si è aggiudicata diversi premi e sovvenzioni in settori come il rilevamento di movimenti per le piattaforme mobili e la ricerca sul deep learning. XJERA sfrutta le soluzioni e le piattaforme NVIDIA per fornire soluzioni di analisi video nella sorveglianza con livelli di prestazioni e precisione non disponibili nei prodotti standardizzati tipici.

Una delle principali problematiche tecniche dell'analisi video è il conteggio delle persone, dato che l'algoritmo deve gestire scenari diversi:

- varie caratteristiche dei singoli e della massa, come posizioni della testa, ornamenti, occlusioni e acconciature
- condizioni ambientali, come la direzione di ombre e luci, variazioni delle dimensioni e punti di osservazione
- le caratteristiche delle occlusioni dovute alla posizione e alle diverse angolazioni di installazione della videocamera.

XJERA LABS ha sviluppato algoritmi basati sul deep learning che individuano e seguono rapidamente una persona, contano il numero di persone che oltrepassano una linea in un ambiente ad alta densità o effettuano persino il riconoscimento facciale da una videocamera installata in un angolo direzionale. Tutto questo viene eseguito con l'88% di precisione nelle scene e il 95% dopo la riconfigurazione dei parametri, con implementazioni reali come un parco a tema a Singapore. (Cfr. la Tabella 1).



**Figura 9.** Il conteggio delle aree e delle persone che oltrepassano una linea effettuato da XJERA VCA

XJERA LABS utilizza le funzionalità di accelerazione di NVIDIA Tesla P4 per eseguire azioni di grande efficacia come l'estrazione automatica di tipi diversi di attributi da un feed video dal vivo in massimo 70 millesimi di secondo, con una precisione straordinaria del 90% circa.

#### Attributi delle persone

- |                          |                            |                            |
|--------------------------|----------------------------|----------------------------|
| • Maschio o femmina      | • Con o senza maglietta    | • Pantaloni lunghi o corti |
| • Capelli lunghi o corti | • Maniche lunghe o corte   | • A motivi o non a motivi  |
| • Con o senza occhiali   | • Con o senza jeans        | • Scontro (azione)         |
| • Con o senza cappello   | • Con o senza pantaloncini |                            |

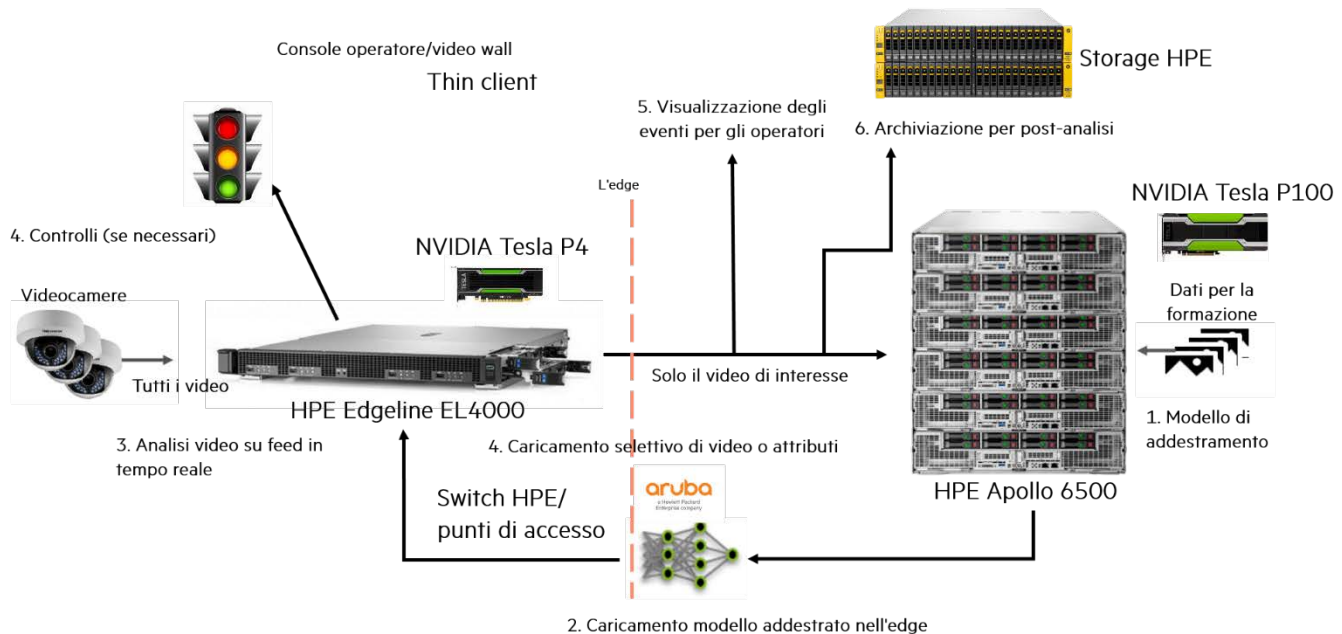
#### Attributi del veicolo

- Marca dell'auto
- Colore dell'auto
- Velocità dell'auto

### Diagramma dell'architettura

L'addestramento degli algoritmi di analisi video è un'attività ad estrema intensità di elaborazione che richiede l'accesso a grandi quantità di dati video storici di esempio per l'apprendimento. Per questa attività viene in genere utilizzato un sistema potente ad alta densità come HPE Apollo 6500, dotato di più acceleratori NVIDIA Tesla P100. Il modello addestrato viene quindi caricato nei sistemi di analisi edge come HPE Edgeline EL4000, che lo eseguono sui dati video in tempo reale con il supporto dell'acceleratore NVIDIA Tesla P4. Può quindi scegliere di caricare selettivamente solo i segmenti del video di interesse o solo gli attributi estratti (ad esempio volto, vestiario, tipo di autoveicolo, targa, incendio) nell'intero edge. L'uso dei risultati dell'analisi può essere inoltre consentito per intraprendere azioni di controllo immediate (e generalmente urgenti) come lo spegnimento delle apparecchiature, l'impostazione del semaforo sul rosso o l'attivazione di un allarme.

L'elaborazione nell'edge dei risultati video determina una latenza di reazione ottimale, mentre i video selettivi o l'uplink degli attributi riducono la larghezza di banda e i rischi per i dati. Il fatto che l'operatore non sia più costretto a guardare centinaia di feed video per individuare un potenziale problema migliora la capacità di riconoscimento e reazione in modo efficace. Inoltre, dato che la maggior parte delle informazioni di valore è già stata estratta dai feed delle videocamere, è necessario archiviare meno dati video nel back-end (in genere solo le sequenze pre e post evento), riducendo sensibilmente la necessità di grandi array di storage nel data center.



**Figura 10.** Architettura IVA (Intelligent Video Analytics) end-to-end di HPE e NVIDIA

Sebbene il grafico strutturale riporti un'architettura Intelligent Video Analytics end-to-end completa, questo white paper riguarda solo la valutazione delle prestazioni di HPE Edgeline EL4000 con schede NVIDIA Tesla P4. Il software di analisi XJERA LABS utilizzato a questo scopo è già stato addestrato e ottimizzato per l'esecuzione degli algoritmi di analisi video selezionati e analizzerà i flussi video preregistrati per un benchmarking stabile e ripetibile.

## Obiettivi dell'analisi video e configurazione del sistema

### Obiettivi del test

Il test delle prestazioni dell'analisi video è stato condotto con un sistema completamente indipendente; l'analisi è stata svolta su flussi video preregistrati per risultati coerenti e ripetibilità. L'obiettivo è misurare la precisione e la velocità del software di analisi video in due situazioni diverse (solo CPU e CPU con il supporto della GPU) per verificare la capacità di NVIDIA Tesla P4 di accelerare tali operazioni.

### Configurazione hardware

Il test del sistema prevede la seguente configurazione hardware.

1. Un (1) sistema edge convergente HPE Edgeline EL4000 con switched networking 10 GbE integrato
2. Due (2) alimentatori da 800 W
3. Una (1) scheda NVIDIA Tesla P4
4. Una (1) cartuccia server HPE m510
  - a. Intel Xeon D-1548/2,0 GHz (Turbo: 2,3–2,6 GHz)/8-core/12 MB/45 W
  - b. RAM DDR4 128 GB (4 x 32 GB)
  - c. SSD SATA M.2 2242 1 x 64 GB



## Configurazione software

La configurazione software è la seguente.

1. Ubuntu 14.04
2. XJERA XIntelligence 0.36
3. NVIDIA SDK (CUDA 8.0, cuDNN 5.1)
4. TensorRT 2.0

## Video di input

Il flusso video di input utilizzato per il test è il seguente.

1. **Descrizione:** scena affollata nel parco a tema
2. **Risoluzione:** 720p
3. **Frame al secondo (FPS):** 15
4. **Codec:** H.264 (AVC)

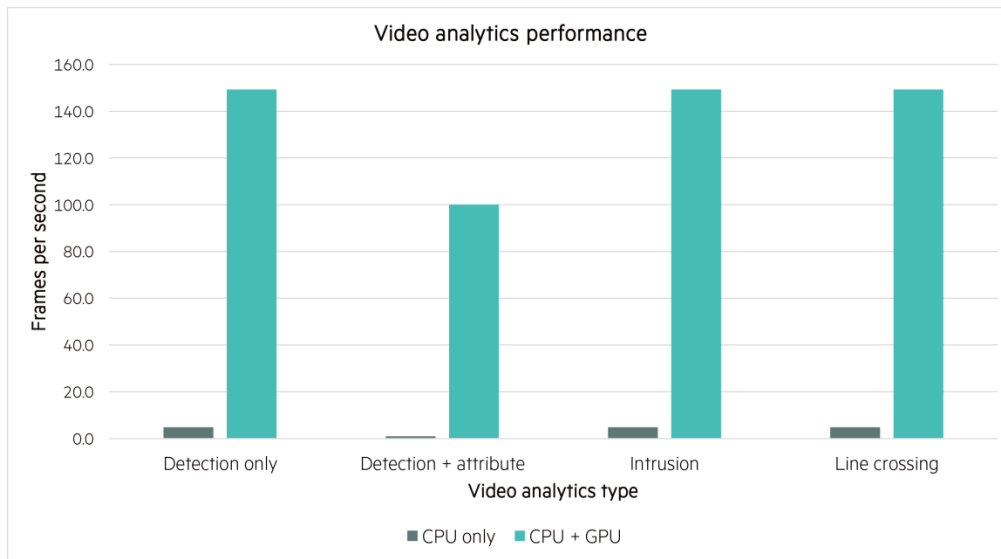


**Figura 11.** Frame di esempio dal video di input utilizzato per il test

## Prestazioni dell'analisi video

Gli algoritmi dell'analisi video di XJERA sono in grado di fornire un'eccellente precisione superiore al 95% per l'individuazione delle persone, il rilevamento delle intrusioni e il superamento delle linee, con un valore superiore all'85% per il rilevamento degli attributi. (Cfr. la Tabella 1).

Presumendo un target prefissato di precisione, l'analisi video di base (rilevamento, intrusione e superamento di linee) accelerata da GPU offre prestazioni 31 volte più elevate rispetto allo stesso algoritmo eseguito solo sulla CPU. Questo divario aumenta di addirittura 111 volte quando un'analisi video complessa come il rilevamento degli attributi viene eseguita sullo stesso flusso video. (Cfr. la Tabella 1).



**Figura 12.** Prestazioni di algoritmi di analisi video diversi eseguiti solo su CPU e su CPU + GPU

**Tabella 1.** Prestazioni di algoritmi di analisi video diversi eseguiti solo su CPU e su CPU + GPU

Tipi di analisi video	Attributo rilevato (se presente)	Precisione	Prestazioni (frame/sec) - solo CPU	Prestazioni (frame/sec) - CPU + GPU	CPU + GPU rispetto a solo CPU
<b>Solo rilevamento</b>		97%	4,8	149,3	31X frame in più al secondo
<b>Rilevamento + attributo</b>	Maschio	92%	0,9	100,0	111 di frame in più al secondo
	Capelli lunghi	93%			
	Occhiali	85%			
	Cappello	96%			
	Maglietta	84%			
	Maniche lunghe	86%			
	Pantaloncini	98%			
	Jeans	92%			
	Pantaloni lunghi	92%			
	Motivo	88%			
Scontro	94%				
<b>Intrusione</b>		97%	4,8	149,3	31X frame in più al secondo
<b>Superamento di una linea</b>		96%	4,8	149,3	31X frame in più al secondo

## Dimensioni del sistema

In base ai risultati del test, il numero stimato di feed video può essere analizzato da un sistema HPE Edgeline, come indicato di seguito.

Il dimensionamento definitivo per un caso d'uso reale è soggetto a variazioni, dal momento che dipende per molti aspetti dalla qualità del feed video (risoluzione, frame al secondo, tipo di codec), dalla complessità degli eventi nel video, dal tipo di analisi video che verrà eseguita etc.

**Tabella 2.** Numero stimato di feed video che può essere elaborato per sistema HPE Edgeline con analisi video CPU + GPU

Modello di sistema	Server per sistemi	GPU per sistema	Tipo di analisi video	N. stimato di feed video per sistema (720p @ 15 fps)
<b>HPE Edgeline EL1000</b>	Un (1) HPE m510	Una (1) NVIDIA Tesla P4	Rilevamento	<b>9</b>
			Rilevamento + attributo	<b>6</b>
			Intrusione	<b>9</b>
			Superamento di una linea	<b>9</b>
<b>HPE Edgeline EL4000<sup>10</sup></b>	Un (1) HPE m510	Una (1) NVIDIA Tesla P4	Rilevamento	<b>9</b>
			Rilevamento + attributo	<b>6</b>
			Intrusione	<b>9</b>
			Superamento di una linea	<b>9</b>
	Due (2) HPE m510	Due (2) NVIDIA Tesla P4	Rilevamento	18
			Rilevamento + attributo	12
			Intrusione	18
			Superamento di una linea	18
	Tre (3) HPE m510	Tre (3) NVIDIA Tesla P4	Rilevamento	27
			Rilevamento + attributo	18
			Intrusione	27
			Superamento di una linea	27
	Quattro (4) HPE m510	Quattro (4) NVIDIA Tesla P4	Rilevamento	36
			Rilevamento + attributo	24
			Intrusione	36
			Superamento di una linea	36

<sup>10</sup> Limitato a una sola scheda NVIDIA Tesla P4 per questo test. È previsto un aggiornamento del sistema HPE Edgeline EL4000 che in futuro ospiterà fino a quattro schede NVIDIA Tesla P4

## Conclusione

Non solo la tecnologia Intelligent Video Analytics all'edge consente latenza rapida di azione, ma la pre-elaborazione e l'invio selettivo determinano anche enormi risparmi sui costi di storage e networking. È importante scegliere la corretta combinazione di hardware e software, in modo che la piattaforma:

1. sia estremamente affidabile in termini di precisione dei risultati
2. sia in grado di elaborare i feed delle videocamere a costi contenuti
3. sia adatta all'implementazione in ambienti edge complessi.

I risultati di questo documento indicano chiaramente che, per un livello mirato di precisione, l'analisi video basata solo sulla CPU non offre lo stesso livello di prestazioni dell'analisi video basata su una solida CPU e accelerata dalla GPU. In effetti, un algoritmo di analisi video complesso, come il rilevamento degli attributi delle persone, può utilizzare un'intera CPU a 8 core solo per elaborare un solo frame video al secondo, il che è inaccettabile persino in presenza di un'unica videocamera.

Tuttavia, lo stesso server supportato da una GPU NVIDIA Tesla P4, pur utilizzando la stessa CPU a 8 core, può facilmente elaborare 100 frame al secondo (~6 videocamere a 15 fps), offrendo una soluzione molto più praticabile per l'analisi all'edge. Inoltre, l'offload libera la CPU per altre attività come la correlazione dei risultati dell'analisi video con i database, la raccolta di dati da altri sensori all'edge (ad esempio i lettori di badge) o il controllo delle apparecchiature (come segnali e allarmi).

I sistemi edge convergenti HPE Edgeline fanno convergere l'elaborazione di livello enterprise (CPU Intel Xeon e GPU Intel o NVIDIA), lo storage a elevate prestazioni (fino alle unità NVMe), il networking ad alta velocità (fino a 10 GbE), l'acquisizione dei dati, la tecnologia di controllo e la gestione leader di settore (HPE iLO) in un fattore di forma compatto e robusto adatto alla collocazione nei siti periferici. HPE Edgeline EL1000 è un singolo sistema server in grado di analizzare da sei a nove feed video simultanei; le dimensioni sufficientemente contenute ne consentono il posizionamento all'interno di un veicolo o una scatola di distribuzione. Le robuste funzionalità wireless (Wi-Fi e 4G LTE) lo rendono portatile e utilizzabile in siti privi di reti cablate. HPE Edgeline EL4000 è un sistema 1U sottile installabile su rack o a parete. La scalabilità da uno a quattro server nello chassis e la facilità di dimensionamento per l'analisi simultanea di un massimo di 36 feed video lo rende la scelta ideale per il monitoraggio di aree più vaste come vari incroci stradali, ingressi di edifici o posti di controllo di sicurezza.

In conclusione, si può notare che i sistemi edge convergenti HPE Edgeline, NVIDIA Tesla e l'ecosistema ISV basato sul deep learning come XJERA costituiscono una combinazione ideale. Tale piattaforma è perfetta per offrire soluzioni di analisi video all'avanguardia che consentono la sorveglianza intelligente nelle città IA del futuro.

## Notifiche

Software riservato. È necessario avere una licenza valida di Hewlett Packard Enterprise per il possesso, l'utilizzo o l'esecuzione di copie. In conformità con FAR 12.211 e 12.212, il software per uso commerciale, la documentazione del software e i dati tecnici relativi ai componenti commerciali sono forniti di licenza dal Governo degli Stati Uniti in base alla licenza commerciale standard del fornitore.

## Ulteriori informazioni alla pagina

[hpe.com/info/edgeline](http://hpe.com/info/edgeline)

[nvidia.com](http://nvidia.com)

[xjeralabs.com](http://xjeralabs.com)



---

**Registrati per ricevere gli aggiornamenti**

---

---

© Copyright 2017 Hewlett Packard Enterprise Development LP. Le informazioni contenute in questo documento sono soggette a modifiche senza preavviso. Le uniche garanzie valide sui servizi e i prodotti Hewlett Packard Enterprise sono espressamente indicate nelle dichiarazioni di garanzia che accompagnano tali prodotti e servizi. Nulla di quanto contenuto nel presente documento può essere ritenuto una garanzia aggiuntiva. Hewlett Packard Enterprise declina ogni responsabilità per eventuali omissioni ed errori tecnici o editoriali nel presente documento.



Intel e Intel Xeon sono marchi di Intel Corporation negli Stati Uniti e in altri Paesi. NVIDIA è un marchio e/o marchio registrato di NVIDIA Corporation negli Stati Uniti e in altri paesi. Tutti gli altri marchi di terzi sono di proprietà dei rispettivi titolari.