



 eestra

 eestracom

Il sistema di videosorveglianza 5G di Prato

15 Novembre 2019

IL SISTEMA DI VIDEOSORVEGLIANZA PRATO CITTA' SICURA



Realizzato da Estracom S.p.A., società del Gruppo Estra che si occupa di TLC

Il sistema di videosorveglianza «Prato città sicura» è composto da circa 150 telecamere da alta risoluzione, tutte collegate tramite un'infrastruttura dedicata in Fibra Ottica, con architettura FTTH-GPON.

Il sistema «Prato città sicura» è in uso a:

- ✓ **Comune di Prato.**
- ✓ **Questura di Prato.**
- ✓ **Comando Provinciale Carabinieri Prato.**
- ✓ **Polizia Municipale Prato.**
- ✓ **Protezione Civile.**
- ✓ **Comando Provinciale Vigili del Fuoco Prato.**

Tutte le centrali operative sono dotate di un centro di regia, tutte le sedi di regia sono collegate tramite un anello in Fibra Ottica a 10 Gbps.



UNO SGUARDO AL FUTURO



5G ↔ FIBRA OTTICA

INFRASTRUTTURE TLC + TECNOLOGIA + INNOVAZIONE = SMART CITY

REALIZZARE UNA CITTÀ “SMART” NON PUÒ PRESCINDERE DA UNA FORTE VALORIZZAZIONE DELLE RETI TLC, CHE DEVONO ESSERE SEMPRE PIÙ CAPILLARI E PERFORMANTI






ESTRA ED ESTRACOM NELLA SPERIMENTAZIONE 5G



CAPOFILA: **open fiber** | **WIND** | **3**

CITTA': PRATO E L'AQUILA



USE CASE LEADER DI “IoT E SENSORI” E “VIDEOSORVEGLIANZA ” A PRATO
in collaborazione con:



UNIVERSITÀ
DEGLI STUDI
FIRENZE



PIN

POLO
UNIVERSITARIO
CITTÀ DI PRATO

SERVIZI DIDATTICI
E SCIENTIFICI
PER L'UNIVERSITÀ
DI FIRENZE





UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

UC6: Videosorveglianza Cittadina



comune di
PRATO

L'obiettivo dello use case è quello di far evolvere e portare valore aggiunto al sistema di Videosorveglianza Prato Città Sicura già in esercizio.

Migliorare le funzionalità di analisi video esistenti ed aggiungerne di nuove, andando a soddisfare le esigenze delle Pubbliche Amministrazioni e dei cittadini.

Con la tecnologia 5G ogni punto in cui sia disponibile un'alimentazione elettrica diventa un possibile punto di installazione di una telecamera. Si possono così dislocare un maggior numero di telecamere nelle zone della città di maggiore interesse, semplicemente riposizionando una parte di quelle esistenti, prelevandole dalle zone in cui non sono più necessarie, senza dover sostenere il costo di acquistarne di nuove.

Lo Use Case Videosorveglianza è sviluppato in collaborazione con l'Università di Firenze-MICC (Media Integration and Communication Center), Prof. Marco Bertini

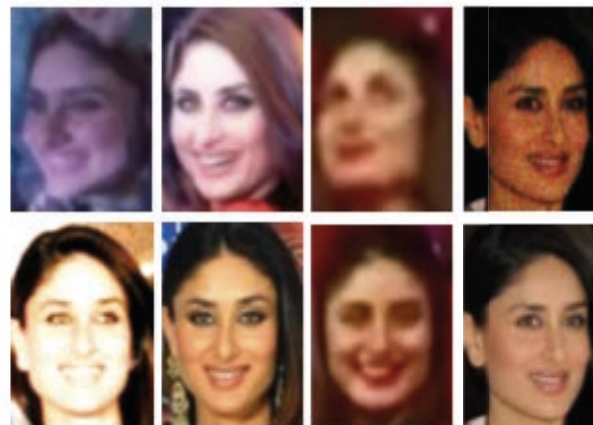


UC6: Videosorveglianza Cittadina

- Lo use case consiste nella realizzazione di un sistema di videosorveglianza cittadina con funzionalità avanzate di analisi automatica dei flussi video, sfruttando la possibilità data dal 5G e dalla fibra ottica di usare telecamere ad alta risoluzione.
- Il sistema di analisi riconoscerà:
 - Volti e persone (re-identificazione)
 - Targhe / veicoli (re-id)
 - Abbandono di rifiuti (IR)



Esempio di re-identificazione di persone



Riconoscimento di volti con diverse: pose, espressioni, illuminazioni, etc.



Esempio di analisi targhe e veicoli

UC6: Videosorveglianza Cittadina

Organizzazione in WP

- **WP1: Servizio di Monitoraggio e controllo del territorio**
- **WP2: Servizio di Controllo del traffico veicolare**
- **WP3: Servizio di Salvaguardia del territorio**



UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE



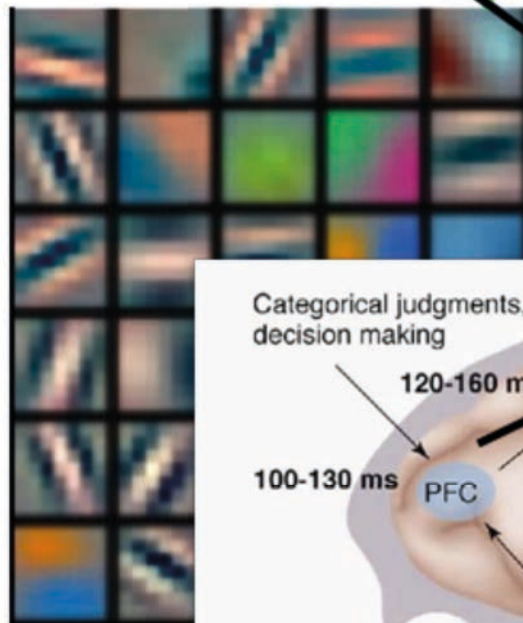
Il contributo del 5G

- Il valore aggiunto del 5G è banda disponibile per trasmettere i flussi video delle telecamere, es. da un network di telecamere HD per applicazioni di reidentificazione o valutazione degli spostamenti. eMBB (enhanced Mobile Ultra broadband)
- Facilità di installazione durante eventi temporanei o in zone sottoposte a vincoli per installazioni fisse per es. beni culturali.
- La banda disponibile nel 5G consente anche di usare in installazioni temporanee le più recenti telecamere termiche, che adesso hanno raggiunto risoluzioni HD.
- L'applicazione di tecniche di visione artificiale, come l'uso delle reti neurali convoluzionali, consente di automatizzare diversi compiti di video sorveglianza, rendendo effettivamente utilizzabili le installazioni di reti di telecamere.

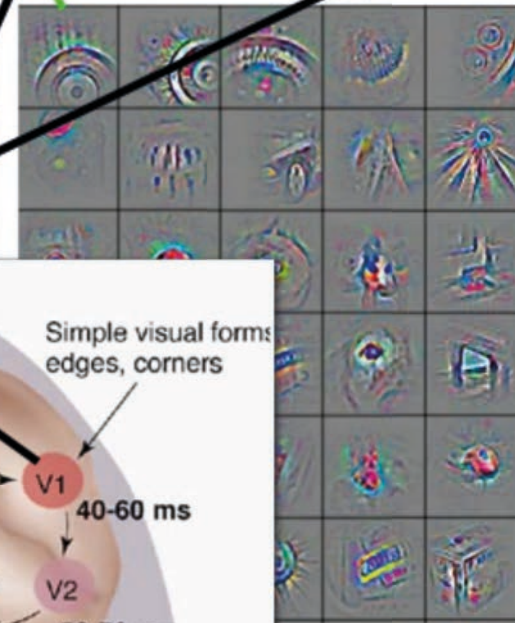


Le reti neurali profonde

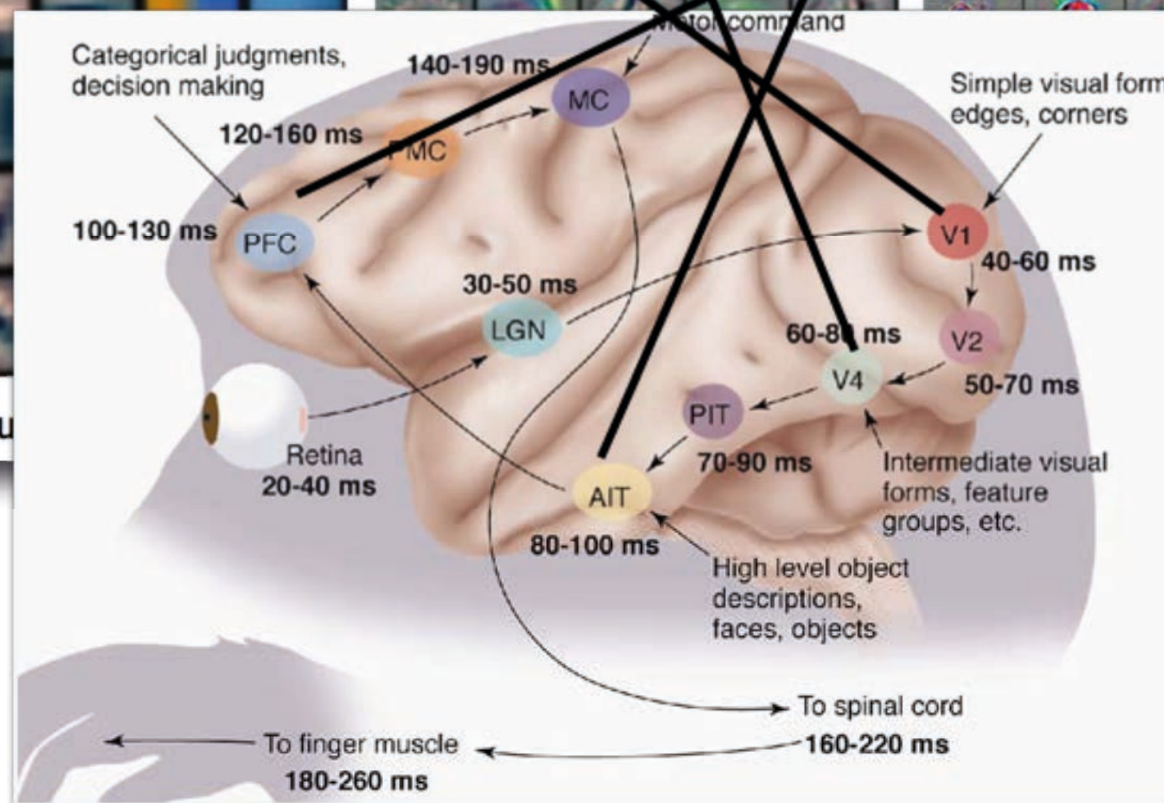
- Le reti neurali profonde, ovvero composte da un grande numero di strati di «neuroni», rappresentano lo stato dell'arte nella visione artificiale, per compiti di detection, recognition, segmentazione, ricerca e classificazione.
- La loro efficacia è basata sulla grande disponibilità di dati annotati, per imparare nei loro primi strati quali caratteristiche elementari dell'immagine sono più distintive, ed imparare negli strati successivi elementi sempre più di alto livello. L'ultimo strato della rete svolge il compito specifico della rete a partire dalle caratteristiche di alto livello.
- La disponibilità di strumenti di calcolo parallelo a basso costo come le GPU ha reso possibile la loro applicazioni anche in contesti di elaborazione real-time.



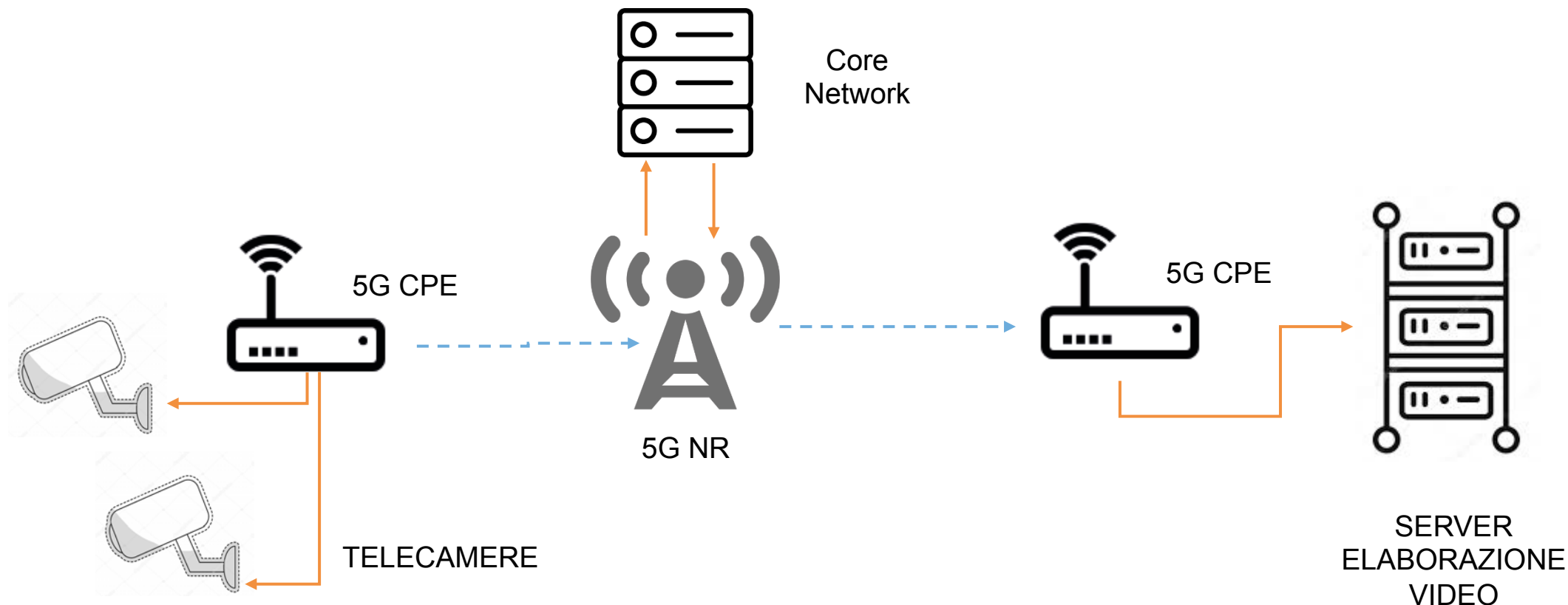
Feature visu



Fergus 2013]

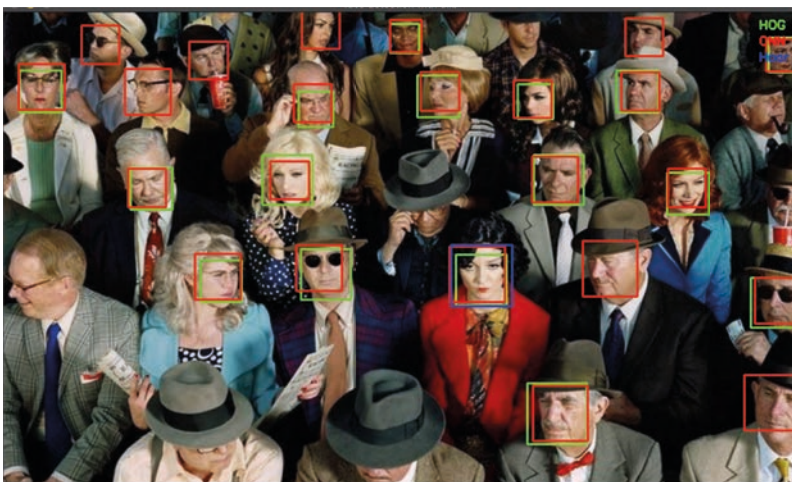


Use Case 6 «Videosorveglianza» - architettura



Face Recognition

- Acquisizione di video da telecamera
- Il video passa alla CORE net che lo inoltra alla 5G NR
- Il segnale della 5G NR viene recepito dal modem 5G collegato ad uno switch a cui è collegato un server con GPU
- Il server elabora il video per il riconoscimento della presenza di facce e loro eventuale identificazione in un database





Come funziona la face recognition

- Il sistema è composto da due componenti principali:
 - Un detector (localizzatore) di facce, implementato con due diversi algoritmi, che localizza dove sono nella scena.
 - Un estrattore di descrittori visuali di facce.
- Il sistema è basato sull'uso di reti neurali profonde (deep learning), la cui inferenza è eseguita su GPU.
- Per ogni volto identificato in un fotogramma viene estratto un descrittore visuale che viene confrontato con quelli presenti nel database, calcolandone la distanza.
- Nella demo si è usata una rete addestrata a riconoscere facce in un database di 3.3 milioni di facce.

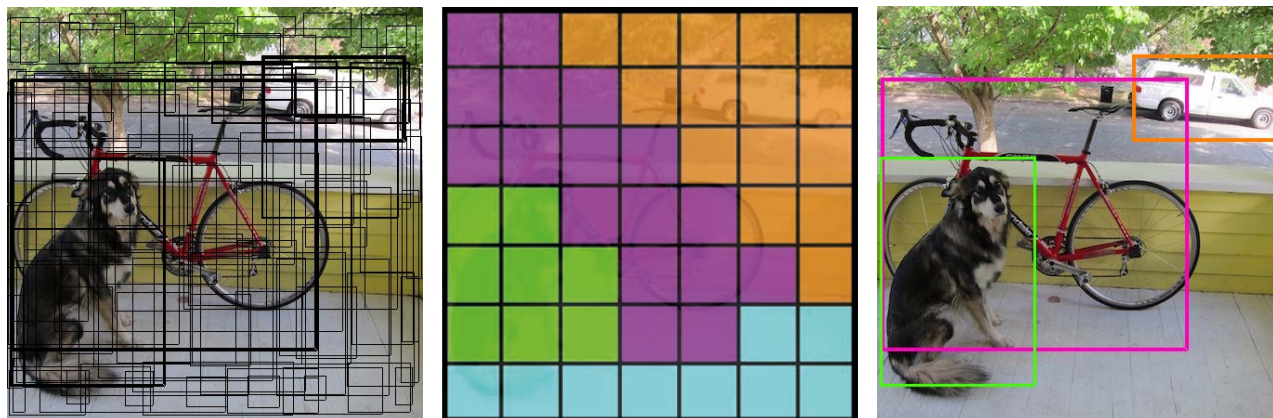
Riconoscimento tipo di veicoli

- **Acquisizione di video da telecamera che inquadra una strada di interesse.**
- **Il video passa alla CORE network che lo inoltra alla 5G NR.**
- **Il segnale della 5G NR viene recepito dal modem 5G collegato ad uno switch a cui è collegato un server con GPU.**
- **Il server elabora il video per il riconoscimento della presenza di veicoli, posizione sul piano immagine e loro classificazione (es.auto, moto, bus, camion,...).**



Come funziona il riconoscimento tipo di veicoli

- Il sistema usa una sola rete neurale profonda per fare la detection (localizzazione) di oggetti di interesse nel controllo del traffico e la classificazione degli stessi.
- La rete suddivide l'immagine in regioni da cui predice dei rettangoli (bounding boxes) che possono contenere degli oggetti, pesandoli per la probabilità della predizione. Ogni regione predice anche la probabilità di contenere uno specifico oggetto. Le due predizioni sono fuse insieme.



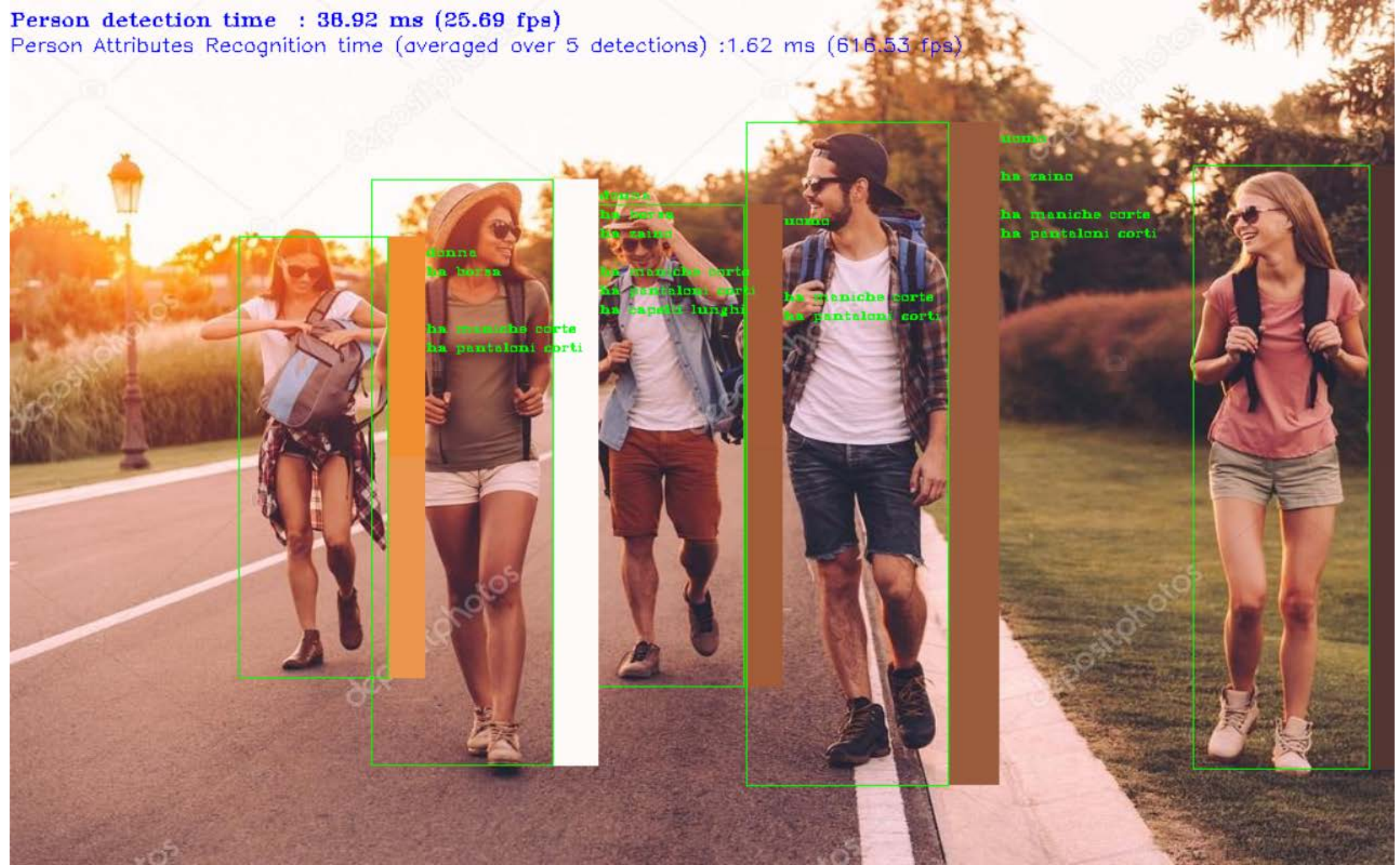
Demo Person Attribute Recognition

- **Acquisizione di video da telecamera.**
- **Il video passa alla CORE network che lo inoltra alla 5G NR.**
- **Il segnale della 5G NR viene recepito dal modem 5G cui è collegato un server.**
- **Il server elabora il video per il riconoscimento della presenza di persone e per ogni persona riconosce la presenza di attributi utili a descriverla.**



Person detection time : 38.92 ms (25.69 fps)

Person Attributes Recognition time (averaged over 5 detections) :1.62 ms (616.53 fps)



Descrizione Demo Person Attribute Recognition

- Il sistema è composto da due componenti principali:
 - Un detector (localizzatore) di persone, che localizza dove sono nella scena.
 - Un classificatore di attributi che riporta la loro presenza per ogni persona.
- Il sistema è basato sull'uso di reti neurali profonde (deep learning), la cui inferenza è eseguita in forma ottimizzata su CPU.
- Per ogni persona identificata in un fotogramma viene eseguita la rete neurale che valuta la presenza di attributi (sesso, zaino, borsa, tipo di vestiti, etc.).

Demo re-identificazione di persone

Demo re-identificazione di persone:

- **Acquisizione di video da due telecamere che inquadrano porzioni diverse di una strada di interesse.**
- **Il video passa alla CORE net che lo inoltra alla 5G NR.**
- **Il segnale della 5G NR viene recepito dal modem 5G cui è collegato un server.**
- **Il server elabora i video delle due telecamere per il riconoscimento della presenza di persone; per ogni persona calcola un descrittore visuale del corpo, le persone sono tracciate nei loro spostamenti all'interno di una singola ripresa e tra riprese diverse usando la similarità visuale calcolata sulla base del loro descrittore.**



Descrizione Demo re-identificazione di persone

- Il sistema è composto da due componenti principali:
 - Un detector (localizzatore) di persone, che localizza dove sono nella scena.
 - Un estrattore di caratteristiche visuali che modella l'apparenza visuale di una persona sulla base dei suoi vestiti.
- Il sistema è basato sull'uso di reti neurali profonde (deep learning), la cui inferenza è eseguita in forma ottimizzata su CPU.
- Per ogni persona identificata in un fotogramma viene eseguita la rete neurale che calcola il descrittore della loro apparenza visuale. Questo descrittore consente di riconoscere persone riprese da diversi punti di osservazione.
- A ogni persona individuata è assegnato un identificativo univoco, e il suo descrittore visuale è usato per ricercarla nei fotogrammi successivi, sia della stessa telecamera che delle altre telecamere.

Person detection in IR

- Il sistema è composto da una rete neurale inizialmente addestrata per lavorare su immagini RGB, quindi riaddestrata per lavorare su IR componenti principali:
 - Un detector (localizzatore) di persone, che localizza dove sono nella scena.
 - Un detector (localizzatore) di veicoli, che localizza dove sono nella scena.
- Il sistema è basato sull'uso di reti neurali profonde (deep learning), la cui inferenza è eseguita su server dotato di GPU.
- La videosorveglianza IR è privacy-friendly: si può applicare anche di giorno, garantendo alle persone osservate che la loro privacy non sia violata dato che non sono riconoscibili.

