

La visione artificiale (o computer vision o machine vision) è un campo dell'intelligenza artificiale (IA) che permette ai computer di ricavare informazioni significative da immagini digitali, video e altri input visivi e intraprendere azioni o formulare delle segnalazioni sulla base di tali informazioni. Se l'IA in generale permette ai computer di pensare, la computer vision permette loro di vedere, osservare e capire

La Visione Artificiale nel Machining

Massimiliano Annoni



La computer vision funziona più o meno come la vista umana, eccetto che gli umani hanno il vantaggio di disporre di anni e anni di esperienza in cui la vista si è allenata a distinguere gli oggetti, quanto sono lontani, se si stanno muovendo e se c'è qualcosa di sbagliato in un'immagine. La computer vision permette alle macchine di svolgere queste funzioni, ma deve farlo in molto meno tempo con telecamere, dati e al-

goritmi piuttosto che con retine, nervi ottici e una corteccia visiva. Dato che un sistema sviluppato per ispezionare i prodotti o osservare un asset di produzione può analizzare migliaia di prodotti o processi al minuto, notando difetti o problemi impercettibili, esso può superare rapidamente le capacità umane. La computer vision è usata in settori che vanno dall'energia e dai servizi pubblici all'industria manifatturiera e automobilistica e il

mercato continua a crescere. (1, 2) Gli usi più comuni della visione artificiale sono l'ispezione visiva e la rilevazione dei difetti, il posizionamento e la misurazione delle parti, e l'identificazione, la classificazione e il tracciamento dei prodotti. La visione artificiale è una delle tecnologie fondanti dell'automazione industriale. Ha contribuito a migliorare la qualità dei prodotti, accelerare la produzione e ottimizzare la ma-



nifattura e la logistica per decenni. Ora questa tecnologia consolidata si sta fondendo con l'intelligenza artificiale, guidando la transizione verso l'Industria 4.0.

Sistemi di Visione Artificiale classici

Le macchine potevano “vedere” prima dell'IA e del machine learning. All'inizio degli anni '70, i computer iniziarono a usare algoritmi specifici per elaborare immagini e riconoscere caratteristiche di base. Questa tecnologia classica di visione artificiale può rilevare i bordi degli oggetti per posizionare una parte, trovare differenze di colore che indicano un difetto e discernere blob di pixel connessi che indicano un foro.

La visione artificiale classica coinvolge operazioni relativamente semplici che non richiedono intelligenza artificiale. Il testo deve essere semplice e nitido, come un codice a barre. Le forme devono essere prevedibili e adattarsi a un modello esatto. Un sistema di visione artificiale classico non può leggere la scrittura a mano, decifrare un'etichetta spiegazzata o distinguere una mela da un'arancia. Tuttavia, la visione artificiale classica ha avuto un enorme impatto sulla produzione. Le macchine non si stancano, quindi possono individuare difetti più velo-

Applicazione di sistemi di visione artificiale presso il Laboratorio PoliMill del Politecnico di Milano

ciali specializzate possono utilizzare l'imaging termico per rilevare anomalie di calore e i raggi X per individuare difetti microscopici e fatica del metallo (3).

Le tecniche tradizionali di visione artificiale e di elaborazione delle immagini sono state utilizzate nel corso dei decenni in numerose applicazioni e ricerche. Tuttavia, l'avvento delle moderne tecniche di IA, utilizzando reti neurali artificiali che consentono una maggiore accuratezza delle prestazioni, e i progressi nel calcolo ad alte prestazioni grazie alle GPU (Graphics Processing Unit) che permettono un'accuratezza sovrumana, hanno portato a una diffusione capillare in settori come trasporti, retail, manifattura, sanità e servizi finanziari.

Che siano tradizionali o basati sull'IA, i sistemi di visione artificiale possono essere migliori degli esseri umani nel classificare immagini e video in categorie e classi ben definite, come accade per le piccole variazioni nel tempo

ce e in modo più affidabile rispetto agli occhi umani. Inoltre, le macchine non sono limitate come la visione umana. Le telecamere di visione arti-

nelle tomografie assiali computerizzate (TAC). In questo senso, la visione artificiale automatizza compiti che gli esseri umani potrebbero potenzialmente svolgere, ma con una precisione e una velocità molto maggiori.

Con la vasta gamma di applicazioni attuali e potenziali, non sorprende che le proiezioni di crescita per le tecnologie e le soluzioni di visione artificiale siano straordinarie.

L'ascesa dell'Intelligenza Artificiale: Inferenza di Deep Learning e Visione Artificiale Industriale

L'incremento della potenza dell'edge computing, dispositivi embedded e IoT, insieme a un universo crescente di modelli di deep learning per l'intelligenza artificiale, stanno espandendo radicalmente ciò che la visione artificiale può fare. Questa rapida crescita delle capacità sta guidando la trasformazione verso fabbriche intelligenti e Industria 4.0.

L'IA aumenta le prestazioni degli algoritmi classici di visione artificiale con modelli chiamati reti neurali. Quando un computer riceve un'immagine, o un flusso video di immagini, il software di visione artificiale confronta quei dati di immagine con un modello di rete neurale. Questo processo, chiamato inferenza di deep learning, consente ai computer di riconoscere differenze molto sottili come disallineamenti minimi nei pattern dei tessuti e difetti microscopici nelle schede elettroniche.

Per migliorare l'accuratezza e la velocità, i data scientist creano modelli di rete neurale specifici per applicazioni specifiche. Durante questo processo, chiamato addestramento supervisionato, un computer esamina decine di migliaia di campioni e identifica schemi significativi, inclusi schemi che un essere umano potrebbe non rilevare.

Esistono modelli per rilevare pixel non funzionanti e scoloriti nei display, vedere vuoti nelle saldature e individuare fili tirati nei tessuti. Naturalmente, vengono costantemente sviluppati e affinati nuovi modelli (3).

La storia della visione artificiale

Scienziati e ingegneri hanno cercato di sviluppare dei metodi che consentissero alle macchine di vedere e comprendere i dati visivi per circa

PROGRESSO CONTINUO NELLA TORNITURA



INDEX

Centri di tornitura mono e bimandrino, plurimandrino e a fantina mobile

CENTRI DI TORNITURA E FRESATURA INDEX SERIE G

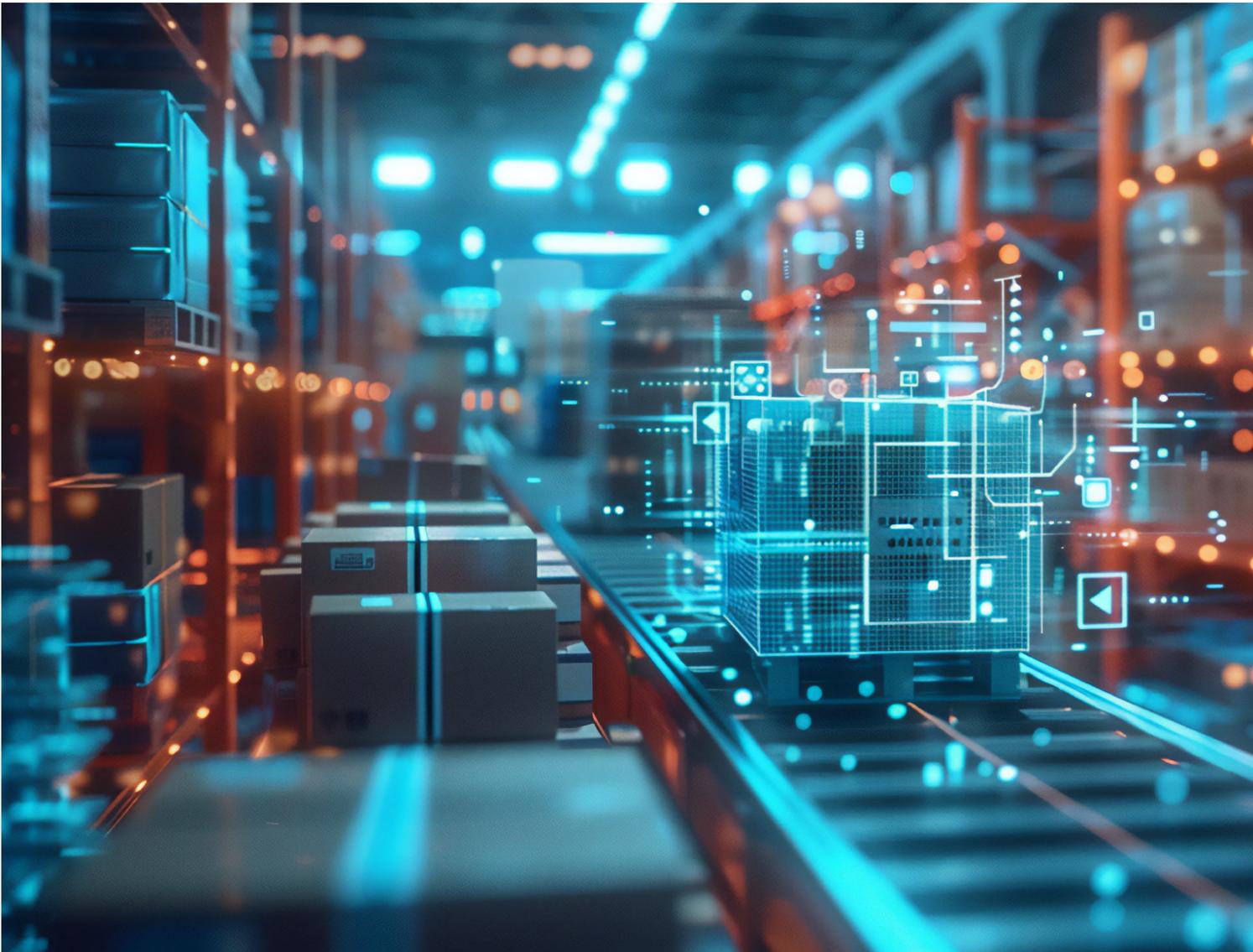
Questa nuova linea di macchine è la soluzione ideale per ogni tipo di esigenze per la produzione di pezzi di complessità da media a molto elevata. L'ampia gamma, che comprende i modelli G220/200 – G320/300 – G420/400 – G520/500, è in grado di soddisfare tutte le esigenze in vari settori: meccanica generale, aerospaziale, energetico, ma non solo. Versioni disponibili anche con isole di lavorazione robotizzate integrate.



Synergon

INSIEME PER IL VOSTRO SUCCESSO

info@synergon.it
www.synergon.it



60 anni. La sperimentazione è iniziata nel 1959 quando dei neurofisiologi mostrarono a un gatto una serie di immagini, cercando di correlarle con una reazione nel suo cervello. Essi scoprirono che reagiva prima ai bordi netti o alle linee e, scientificamente, questo implicava che l'elaborazione delle immagini iniziava con forme semplici come i bordi dritti (4).

Più o meno nello stesso periodo fu sviluppata la prima tecnologia di scansione delle immagini al computer, permettendo ai computer di digitalizzare e acquisire immagini. Un'altra pietra miliare fu raggiunta nel 1963 quando i computer furono in grado di trasformare immagini bi-

dimensionali in forme tridimensionali. Negli anni '60 l'IA divenne oggetto di ricerca accademica e vi furono i primi tentativi di utilizzarla per risolvere il problema della visione umana.

Il 1974 vide l'introduzione della tecnologia di riconoscimento ottico dei caratteri (OCR, Optical Character Recognition), che consentiva di riconoscere il testo stampato in qualsiasi font o carattere. Allo stesso modo, il riconoscimento intelligente dei caratteri (ICR, Intelligent Character Recognition) potrebbe decifrare il testo scritto a mano utilizzando delle reti neurali. Da allora, OCR e ICR hanno consentito di sviluppare applicazioni per l'elaborazione di documenti e fat-

ture, il riconoscimento delle targhe dei veicoli, i pagamenti mobili, la traduzione automatica e altre applicazioni comuni.

Nel 1982 il neuroscienziato David Marr dimostrò che la vista funziona in modo gerarchico e introdusse degli algoritmi per le macchine in grado di rilevare bordi, angoli, curve e forme elementari simili. Contemporaneamente l'informatico Kunihiko Fukushima sviluppava una rete di celle in grado di riconoscere i modelli. La rete, chiamata Neocognitron, integrava degli strati convoluzionali in una rete neurale.

Entro il 2000, l'attenzione degli studi si è concentrata sul riconoscimento degli oggetti e nel



2001 sono apparse le prime applicazioni di riconoscimento facciale in tempo reale. La standardizzazione di come i set di dati visivi sono etichettati, cioè classificati, e annotati è emersa nel corso degli anni 2000. Nel 2010 è diventato disponibile il dataset di dati ImageNet. Esso contiene milioni di immagini etichettate a mano attraverso un migliaio di classi di oggetti e fornisce una base per le CNN e i modelli di apprendimento profondo utilizzati oggi. Nel 2012 un team dell'Università di Toronto inseriva una CNN in un programma di riconoscimento delle immagini. Il modello, chiamato AlexNet, riduceva significativamente il tasso di errore nel rico-

Segmentazione	Classificazione	Rilevamento
Adatta a delineare oggetti	E' un gatto o un cane?	Dove è presente l'oggetto nello spazio?
Usata in veicoli a guida autonoma	Classifica con precisione	Riconosce gli oggetti per la sicurezza

noscimento delle immagini. Tale miglioramento fu ottenuto tramite una rete neurale decisamente più profonda che fu possibile grazie all'utilizzo di una GPU (Graphics Processing Unit) per il calcolo. Ciò avvenne già nel 2012, quando la potenza di calcolo era irrisoria rispetto ad oggi. Dopo questa svolta, i tassi di errore sono scesi a pochi punti percentuali (4).

Come Funziona la Visione Artificiale?

La computer vision ha bisogno di molti dati. Esegue analisi di dati più e più volte fino a quando non distingue e riconosce le immagini oppure, in linguaggio matematico, fino a quando una "loss function" non è minima. Per esempio, per addestrare un computer a riconoscere gli pneumatici di un'automobile, è necessario alimentarlo con una grande quantità di immagini di pneumatici e di elementi relativi agli pneumatici per imparare le differenze e riconoscere uno pneumatico, specialmente uno senza difetti.

Per realizzare ciò, vengono utilizzate essenzialmente due tecnologie: un tipo di machine learning chiamato deep learning e una rete neurale convoluzionale (CNN - convolutional neural network, un tipo di deep neural network).

Il machine learning utilizza algoritmi che permettono a un computer di imparare da sé il contesto dei dati visivi. Se vengono immessi abbastanza dati in questo algoritmo, il computer "guarderà" i dati e imparerà da solo a distinguere un'immagine da un'altra. Gli algoritmi permettono alla macchina di imparare da sola, senza che qualcuno la programmi per riconoscere un'immagine.

Un CNN aiuta un modello di machine learning o di deep learning a "guardare" scomponendo le immagini in pixel, a cui vengono dati dei tag o delle etichette. Usa le etichette per eseguire delle convoluzioni (un'operazione matematica su due funzioni per produrre una terza fun-

zione) e fa delle previsioni sulla base di ciò che "vede". La rete neurale esegue delle convoluzioni e controlla l'accuratezza delle sue previsioni in una serie di iterazioni fino a quando le previsioni iniziano ad avverarsi. È quindi in grado di riconoscere o vedere le immagini in modo simile a quello degli esseri umani.

Alla stregua di un essere umano che distingue un'immagine a distanza, una CNN distingue prima i contorni nitidi e le forme semplici, poi aggiunge delle informazioni mentre esegue le iterazioni delle sue previsioni. Una CNN è utilizzata per analizzare immagini singole. Una rete neurale ricorrente (RNN, Recurrent Neural Network) è usata in modo simile per le applicazioni video per aiutare i computer a capire come le immagini in una serie di fotogrammi sono collegate tra loro (1).

Schema di una Convolutional Neural Network (CNN)

Le reti neurali convoluzionali (CNN) (5) possono eseguire segmentazione, classificazione e rilevamento per una miriade di applicazioni:

- Segmentazione: La segmentazione delle immagini riguarda la classificazione dei pixel per appartenenza a una certa categoria, come un'auto, una strada o un pedone. È ampiamente utilizzata nelle applicazioni di veicoli a guida autonoma, incluso lo stack software NVIDIA DRIVE™, per mostrare strade, auto e persone.
- Classificazione: La classificazione delle immagini viene utilizzata per determinare cosa c'è in un'immagine. Le reti neurali possono essere addestrate per identificare cani o gatti, ad esempio, o molte altre cose con un alto grado di precisione. Più specificamente è in grado di stabilire con precisione che una data immagine appartiene a una certa classe. Questa capacità potrebbe essere utilizzata da una società di social media per identificare e separare automaticamente delle immagini di-

PROCESSI DI PRODUZIONE

scutibili caricate dagli utenti (1).

- Rilevamento: Il rilevamento delle immagini consente ai computer di localizzare in quale porzione dell'immagine sono presenti gli oggetti. In molte applicazioni, la CNN mette rettangoli di delimitazione intorno alla regione di interesse che contengono completamente l'oggetto. Un rilevatore può anche essere addestrato per vedere dove si trovano auto o persone all'interno di un'immagine.

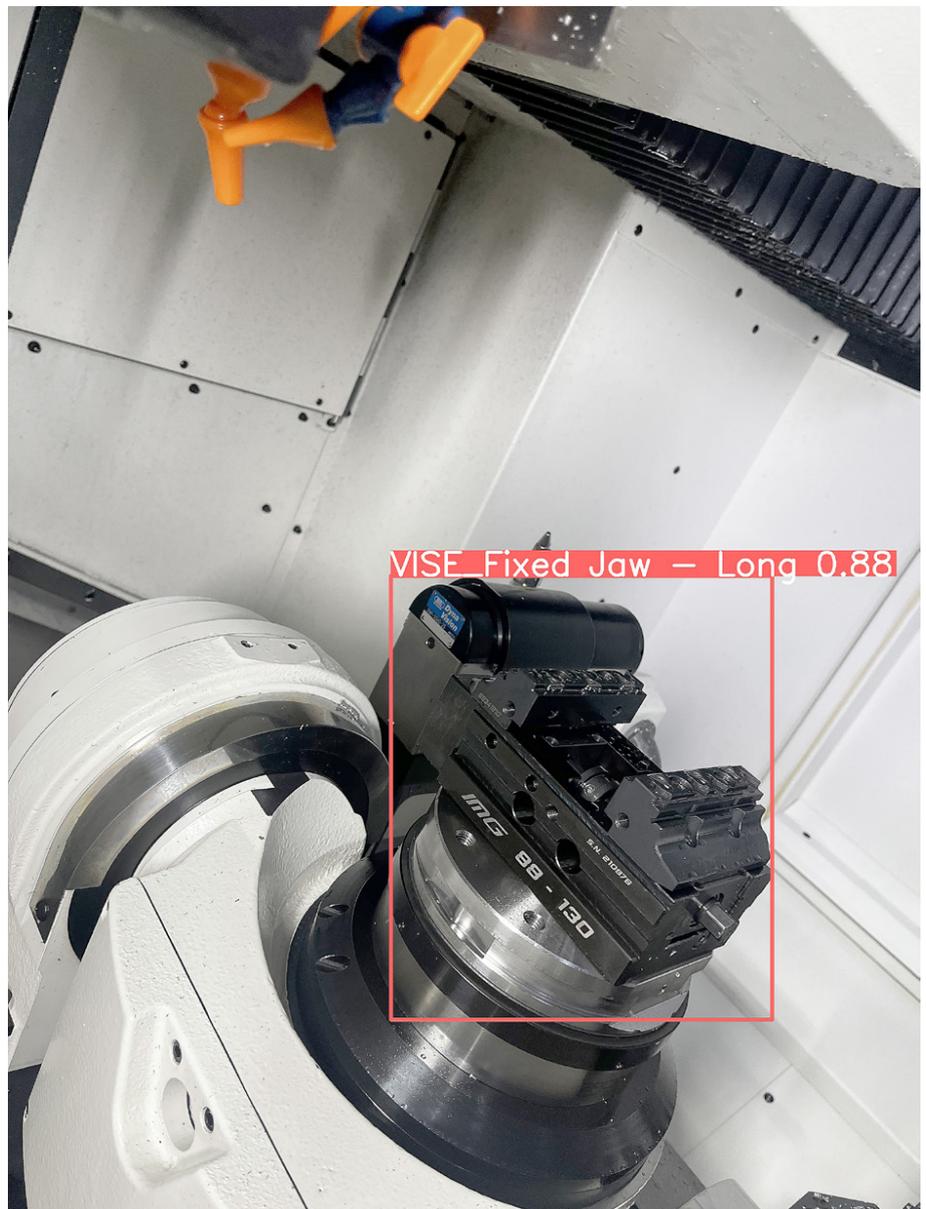
Altre funzioni utili sono:

- Object tracking: algoritmo che segue o traccia un oggetto una volta che è stato rilevato. Questa attività viene spesso eseguita con delle immagini catturate in sequenza o con dei feed video in tempo reale. I veicoli autonomi, ad esempio, hanno bisogno non solo di classificare e rilevare oggetti come pedoni, altre auto e infrastrutture stradali, ma anche di seguirne il percorso in movimento per evitare collisioni e rispettare il codice della strada (6).
- Il recupero delle immagini basato sul contenuto (content-based image retrieval) usa la computer vision per ricercare e recuperare immagini da grandi archivi di dati, basandosi sul contenuto delle immagini piuttosto che sui tag dei metadati ad esse associati. Questa attività può includere un'annotazione automatica dell'immagine che sostituisce quella manuale. Questi compiti possono essere utilizzati per i sistemi di gestione delle risorse digitali e possono aumentare la precisione della ricerca e del ritrovamento (1).

Visione artificiale intelligente e applicazioni strategiche

L'intelligenza artificiale sta espandendo la visione artificiale ben oltre l'ispezione visiva e il controllo qualità. Con la visione artificiale intelligente, i robot possono percepire in tre dimensioni, tenere oggetti l'uno per l'altro e controllare il lavoro reciproco. Possono persino interagire con i colleghi umani e assicurarsi che lavorino insieme in sicurezza.

Le macchine con visione intelligente possono usare l'elaborazione del linguaggio naturale per leggere etichette e interpretare segnali. I robot con visione intelligente possono comprendere forme, calcolare volumi e impacchettare perfettamente scatole, camion e persino container di spedizione con uno spreco minimo di spazio.



Applicazione di sistemi di visione artificiale presso il Laboratorio PoliMill del Politecnico di Milano

I lavoratori umani svolgono i compiti monotoni in modo che i lavoratori possano usare la loro competenza per concentrarsi sulle parti più importanti. L'applicazione della visione artificiale alla produzione può migliorare la qualità del

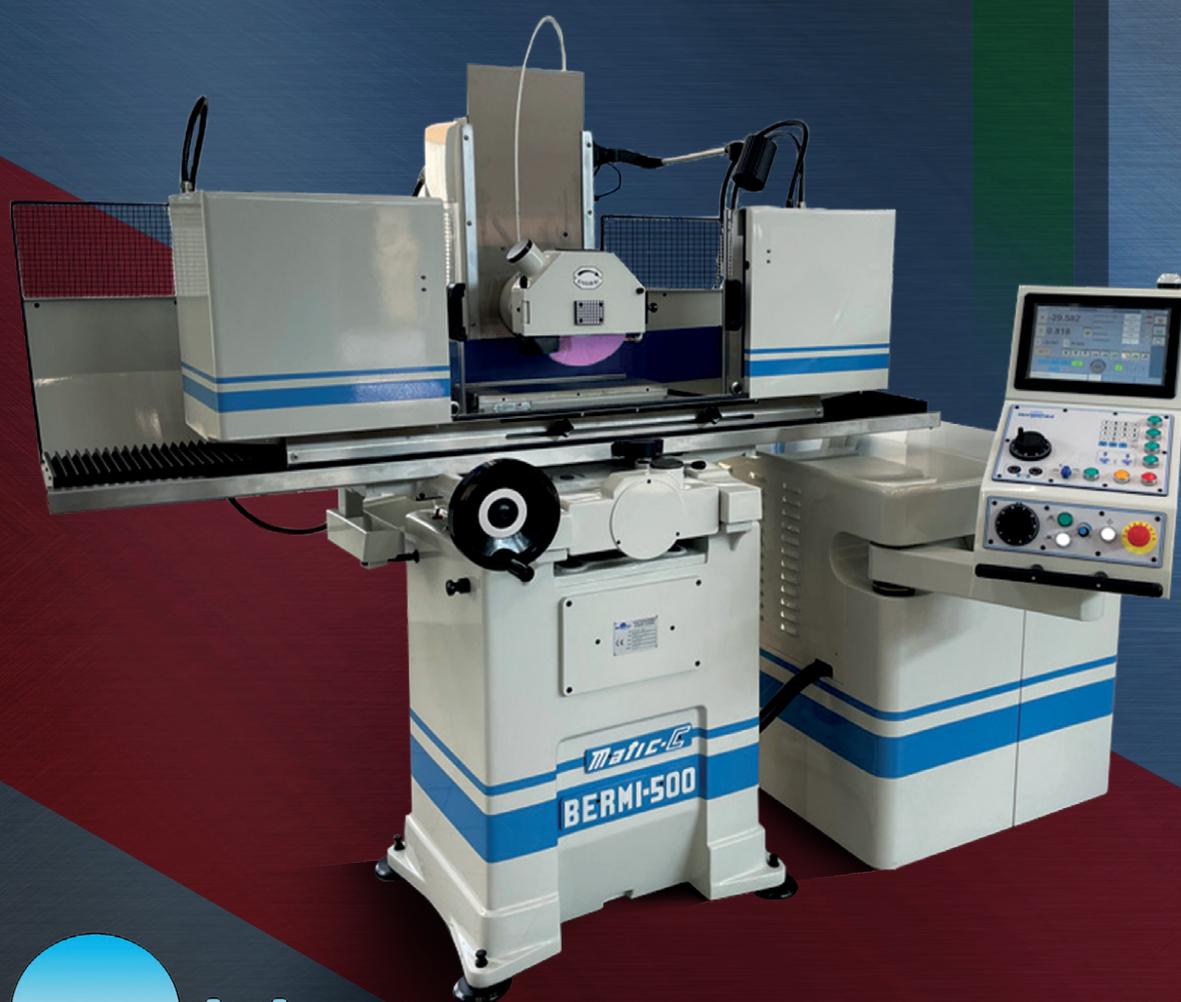
prodotto e l'efficienza complessiva del sistema, aumentando la produttività delle linee di produzione, riducendo i costi del lavoro e liberando il personale per concentrarsi su lavori di maggior valore. Gli usi attuali includono il controllo di qualità, in cui i sistemi di visione artificiale ispezionano le parti e i prodotti finiti alla ricerca di difetti. In agricoltura, i sistemi di visione artificiale utilizzano la selezione ottica per rimuovere materiali indesiderati dai prodotti alimentari. Per Audi, lavorare con Intel e Nebbiolo Technologies, integrare l'analisi predittiva e gli algoritmi

prodotto e l'efficienza complessiva del sistema, aumentando la produttività delle linee di produzione, riducendo i costi del lavoro e liberando il personale per concentrarsi su lavori di maggior valore. Gli usi attuali includono il controllo di qualità, in cui i sistemi di visione artificiale ispezionano le parti e i prodotti finiti alla ricerca di difetti. In agricoltura, i sistemi di visione artificiale utilizzano la selezione ottica per rimuovere materiali indesiderati dai prodotti alimentari. Per Audi, lavorare con Intel e Nebbiolo Technologies, integrare l'analisi predittiva e gli algoritmi

BERMI 500

PRO 2

Con tecnologia all'avanguardia e oltre 50 anni di esperienza nel settore, garantiamo lavorazioni di precisione estrema, zero vibrazioni e finiture superficiali di qualità superiore.



Bergamini
RETTIFICATRICI

Officine Meccaniche Bergamini srl
Viale Italia, 8 - 41037 Mirandola (MO) - Italy
Tel: 0535 20204 - Fax: 0535 23051
E-mail: info@bergamini.it
www.bergamini.it

Rettificatrici BERMI 500
Precisione nei dettagli,
eccellenza nel risultato

PROCESSI DI PRODUZIONE

mi di machine learning nell'ispezione delle saldature e nei processi critici di controllo qualità ha comportato un aumento del numero di saldature analizzate al giorno, una riduzione dei costi del lavoro nelle fabbriche permettendo ad Audi di passare a un monitoraggio più proattivo, evitando problemi piuttosto che semplicemente reagire ad essi.

“Alla fabbrica di Neckarsulm, stiamo già vedendo una riduzione dei costi del lavoro dal 30% al 50%”, dice Michael Häffner, Responsabile della Pianificazione della Produzione, Automazione e Digitalizzazione in Audi.

In settori fortemente regolamentati come quello farmaceutico, la visione artificiale fornisce controlli costanti sui contenuti del prodotto, sull'imballaggio e sull'etichettatura per la garanzia della qualità. Quando applicata alle catene di approvvigionamento, la visione artificiale può scansionare e tracciare automaticamente gli articoli in ogni punto del flusso di lavoro, fornendo un resoconto preciso e aggiornato dell'inventario (3). Nella medicina, l'elaborazione delle immagini mediche comporta l'estrazione rapida di dati vitali dalle immagini per aiutare a diagno-

sticare correttamente un paziente, inclusa la rilevazione rapida di tumori e l'irrigidimento delle arterie. Sebbene la visione artificiale non possa da sola essere considerata affidabile per fornire diagnosi, è una parte inestimabile delle moderne tecniche di diagnostica medica, supportando le valutazioni dei medici e, sempre più spesso, fornendo informazioni che i medici altrimenti non avrebbero visto.

C'è molta ricerca nel campo della visione artificiale, ma non è solo ricerca. Applicazioni concrete dimostrano quanto la computer vision sia importante nelle attività commerciali, nell'intrattenimento, nei trasporti, nella sanità e nella vita di tutti i giorni. Un elemento determinante per la crescita di queste applicazioni è rappresentato dalla enorme massa di informazioni visive che fluiscono da smartphone, sistemi di sicurezza, telecamere del traffico e altri dispositivi dotati di strumenti visivi. Questi dati potrebbero giocare un ruolo importante in tutti i settori, ma oggi rimangono inutilizzati. Le informazioni creano un banco di prova per addestrare le applicazioni di computer vision e un launchpad per farle diventare parte di una serie di attività umane:

- IBM ha utilizzato la computer vision per creare My Moments per il torneo di golf Masters 2018. IBM Watson ha esaminato centinaia di ore di filmati di Masters e sarebbe in grado di identificare le immagini (e i suoni) di colpi significativi. Ha elaborato questi momenti chiave e li ha offerti ai fan sotto forma di filmati personalizzati dei momenti salienti (1).
- Google Translate consente agli utenti di puntare la fotocamera di uno smartphone su una scritta in un'altra lingua e ottenere quasi immediatamente una traduzione della scritta nella loro lingua preferita (2).
- Lo sviluppo di veicoli a guida autonoma si basa sulla computer vision per interpretare gli input visivi delle telecamere e di altri sensori dell'auto. Questo è essenziale per identificare le altre auto, i segnali stradali, i segnali di corsia, i pedoni, le biciclette e tutte le altre informazioni visive che si incontrano sulla strada.
- IBM sta applicando la tecnologia di computer vision con partner come Verizon per portare l'IA all'edge e per aiutare i produttori automobilistici a identificare i difetti di qualità prima che un veicolo lasci la fabbrica (1).



Piattaforme di computer vision

Molte organizzazioni non hanno le risorse per finanziare dei laboratori di computer vision e creare modelli di deep learning e reti neurali. Possono anche non disporre della potenza di calcolo necessaria per elaborare enormi insiemi di dati visivi. Aziende come IBM, Microsoft, Amazon, ecc. stanno aiutando queste organizzazioni offrendo dei servizi di sviluppo di software di visione artificiale. Questi servizi forniscono dei modelli di apprendimento precostituiti disponibili dal cloud e diminuiscono la quantità di risorse di calcolo. Gli utenti si connettono ai servizi attraverso una API (Application Programming Interface) e li usano per sviluppare applicazioni di computer vision.

IBM ha anche introdotto una piattaforma di computer vision che risolve sia i problemi di sviluppo che di risorse di calcolo. IBM Maximo Visual Inspection include degli strumenti che permettono agli esperti in materia di etichettare, addestrare e distribuire modelli di visione di deep learning senza codifica o competenze di deep learning. I modelli di visione possono essere distribuiti nei data center locali, nel cloud e in dispositivi edge.

Anche Microsoft mette a disposizione degli utenti l'API di Visione Artificiale, che fornisce l'accesso ad algoritmi avanzati per l'elaborazione dei media e il ritorno delle informazioni. Caricando un asset multimediale o specificando l'URL di un asset multimediale, gli algoritmi di visione artificiale di Azure possono analizzare i contenuti visivi in diversi modi, basati sugli input e le scelte dell'utente, su misura per la propria attività (7).

L'API di Visione Artificiale di Azure permette varie funzioni:

- **Analisi delle immagini:** analisi del contenuto visivo per rilevare oggetti, volti, testo e molto altro.
- **Gestione dei video:** elaborazione dei video per estrarre informazioni chiave, rilevare attività e identificare oggetti.
- **Personalizzazione:** algoritmi su misura per soddisfare le esigenze specifiche della tua azienda.

Vision Studio di Microsoft svolge alcune funzioni utili di computer vision:

- **Riconoscimento Ottico dei Caratteri (OCR).** Il servizio di riconoscimento ottico dei caratteri



Riconoscimento di una morsa tramite machine vision presso il laboratorio PoliMill del Politecnico di Milano

consente di estrarre testo stampato o scritto a mano da immagini, come foto di segnaletiche stradali e prodotti, nonché da documenti quali fatture, bollette, report finanziari, articoli e altro ancora.

- **Analisi Spaziale.** Il servizio di analisi spaziale analizza la presenza e il movimento delle persone in un feed video e genera eventi a cui altri sistemi possono rispondere.
- **Face (Viso).** Il servizio Face fornisce algoritmi di intelligenza artificiale che rilevano, rico-

noscono e analizzano i volti umani nelle immagini. Il software di riconoscimento facciale è importante in molti scenari diversi, come la verifica dell'identità, il controllo dell'accesso senza contatto e la sfocatura dei volti per la privacy.

- **Analisi delle Immagini.** Il servizio di analisi delle immagini estrae molte caratteristiche visive dalle immagini, come oggetti, volti, contenuti per adulti e descrizioni di testo generato automaticamente per migliorare l'accessibilità.

Hardware e linguaggi di programmazione per il computer vision

Se non si vogliono utilizzare piattaforme online per motivi di privacy o di costi eccessivi quando si debbano analizzare grandi moli di dati, si può considerare l'acquisto di un computer adatto per l'elaborazione.

Dal punto di vista dell'architettura hardware, la CPU (Central Processing Unit) di un computer è composta da pochi core con molta memoria cache che possono gestire alcuni thread software alla volta. Al contrario, una GPU è composta da centinaia di core che possono gestire migliaia di thread simultaneamente. Poiché le reti neurali sono costituite da un gran numero di neuroni identici, sono altamente parallele per natura. Questo parallelismo si mappa naturalmente sulle GPU, che forniscono un'architettura aritmetica parallela sui dati e un significativo aumento della velocità di calcolo rispetto all'addestramento solo con CPU.

Questo tipo di architettura esegue un insieme simile di calcoli su una matrice di dati di immagine. La capacità single-instruction, multiple-data (SIMD) della GPU la rende adatta per eseguire compiti di visione artificiale, che spesso coinvolgono calcoli simili operanti su un'intera immagine.

In particolare, le GPU NVIDIA accelerano significativamente le operazioni di visione artificiale, liberando le CPU per altri compiti. Inoltre, più GPU possono essere utilizzate sulla stessa macchina, creando un'architettura capace di eseguire più algoritmi di visione artificiale in parallelo (5).

Python è il linguaggio di programmazione più popolare per il machine learning (ML), e la maggior parte degli data scientist conosce la sua facilità d'uso e il suo vasto store di librerie, molte delle quali gratuite e open-source.

I data scientist utilizzano Python nei sistemi di ML per l'estrazione e l'analisi dei dati, poiché Python supporta un'ampia gamma di modelli e algoritmi di ML essendo inoltre un linguaggio di alto livello facilmente leggibile.

Dato il rapporto tra ML e visione artificiale, i data scientist possono sfruttare l'universo in espansione delle applicazioni di visione artificiale per le aziende di ogni tipo per estrarre informazioni vitali da store di immagini e video e migliorare il processo decisionale basato sui dati.

Riferimenti bibliografici



1 - <https://www.ibm.com/it-it/topics/computer-vision>



2 - 7 Amazing Examples of Computer And Machine Vision In Practice, Bernard Marr, Forbes, 8 aprile, 2019, <https://www.forbes.com/sites/bernardmarr/2019/04/08/7-amazing-examples-of-computer-and-machine-vision-in-practice/#3dbb3f751018>



3 - <https://www.intel.com/content/www/us/en/manufacturing/what-is-machine-vision.html>



4 - <https://hackernoon.com/a-brief-history-of-computer-vision-and-convolutional-neural-networks-8fe8aacc79f3>



5 - <https://www.nvidia.com/en-us/glossary/computer-vision/>



6 - The 5 Computer Vision Techniques That Will Change How You See The World, James Le, Heartbeat, 12 aprile 2018, <https://jameskle.com/writes/computer-vision>



7 - <https://azure.microsoft.com/en-us/blog/introducing-vision-studio-a-uibased-demo-interface-for-computer-vision/>

Applicazione della visione artificiale presso PoliMill

Presso PoliMill crediamo che per ottimizzare si debba automatizzare. I sistemi di visione artificiale soffrono di due problematiche importanti che ne ritardano l'adozione:

- I dataset di immagini che servono per allenare i sistemi di visione artificiale sono più importanti della bontà dell'algoritmo stesso. Un buon data set deve essere variegato, ampio e rappresentativo dei casi uso del componente che si vuole riconoscere.
- Per questo, ottenere vaste quantità di immagini che rispettino le caratteristiche di cui sopra è complesso sia in termini di tempo che economici. Inoltre, le immagini devono essere adeguatamente etichettate per permettere agli algoritmi di deep learning di identificare correttamente i componenti di interesse. Questo comporta in molti casi un lungo e tedioso lavoro manuale in cui vengono manualmente riquadrati i componenti oggetto di

identificazione. Nel caso di decine di migliaia di immagini, questo processo può richiedere settimane se non mesi e costituisce lavoro profondamente alienante.

PoliMill sta ricercando metodologie innovative di automazione che limitino l'intervento umano in questa fase cruciale di pre-lavorazione dei dati per rendere meno oneroso e più automatico possibile il processo.

Lo scopo del lavoro in senso assoluto è creare un sistema di visione artificiale che possa non solo riconoscere se il setup macchina corrisponda a quello richiesto dal programma da eseguire, ma in futuro possa riconoscere se una lavorazione sia stata eseguita o meno, per esempio riuscendo a riconoscere se una punta a forare è integra o se il foro o un'altra feature siano presenti. Questo permette di non avere necessariamente operatori a bordo macchina presenti, evitare errori di analisi dovuti a distrazioni da parte degli operatori e diminuire potenzialmente i falsi positivi.