vinetur

Ricercatori sviluppano un modello di IA che mappa i tralci della vite nelle immagini dei vigneti con un’intersezione su unione dell’88,25%

Nei test invernali in vigneto il sistema ha superato gli strumenti consolidati di segmentazione delle immagini, un passo verso una potatura robotizzata più uniforme

sabato 14 novembre 2026



Un team di ingegneri agrari ha sviluppato un sistema di intelligenza artificiale in grado di distinguere i fusti della vite e i tralci da potare in immagini complesse di vigneti: un passo che potrebbe aiutare i robot del futuro a eseguire la potatura invernale con maggiore uniformità. Il lavoro, pubblicato sulla rivista Smart Agricultural Technology e presentato venerdì 13 novembre, si concentra sulla visione artificiale e non su una macchina per la potatura completamente autonoma.

La potatura è una delle operazioni ordinarie più importanti in vigneto. Aiuta a controllare la crescita della pianta, la produzione di frutti e il vigore della vite e, in ultima analisi, la qualità dell'uva. In molte aree viticole il lavoro dipende ancora in larga misura da squadre di operai specializzati. Di conseguenza è costoso, richiede tempo ed è difficile da standardizzare nei vigneti di grandi dimensioni, soprattutto quando la manodopera scarseggia. Poiché la potatura rappresenta una voce di costo importante per i produttori di uva, una migliore automazione potrebbe avere ricadute per il settore del vino, se la tecnologia si dimostrerà affidabile nei vigneti commerciali.

Il nuovo modello si chiama Dual-branch Transformer-Mamba Fusion Network, o DTMF. Il suo obiettivo è identificare le strutture legnose della vite in immagini riprese all'aperto durante il riposo vegetativo, quando le viti sono spoglie e pronte per la potatura invernale. Può sembrare un compito circoscritto, ma affronta un ostacolo fondamentale per la robotica in vigneto: una macchina non può effettuare un taglio se prima non è in grado di distinguere la vite, i tralci da potare e lo sfondo circostante.

Spesso il software fatica a interpretare lo sfondo. I tralci sottili possono confondersi con i fili della spalliera, il terreno, le erbe infestanti e altri elementi visivi di disturbo. Secondo i ricercatori, i sistemi standard di segmentazione delle immagini hanno spesso difficoltà a cogliere sia la forma complessiva della vite sia i dettagli più fini ai margini dei tralci sottili. Per affrontare il problema, hanno progettato un sistema a doppio ramo. Un ramo della rete analizza la scena nel suo insieme e le relazioni a lunga distanza nell'immagine, mentre l'altro si concentra sulle texture più fini, sui margini e sui pattern di frequenza, utili per seguire strutture legnose sottili.

Dal punto di vista tecnico, il modello combina un vision transformer, una trasformata wavelet discreta e moduli Mamba basati su modelli a spazio di stato. La componente transformer è progettata per cogliere il contesto globale dell'intera immagine. La componente wavelet e Mamba mira invece a preservare i dettagli strutturali più minuti e la continuità dei tralci sottili. Invece di fondere subito i due flussi, il sistema li mantiene separati per gran parte del processo e unisce le rispettive previsioni solo nella fase finale. I ricercatori hanno inoltre adottato un metodo di addestramento in due fasi, che passa dalla classificazione di base dei pixel nelle prime fasi a un affinamento più preciso dei contorni nelle successive, anche attraverso una funzione di perdita basata sui superpixel, concepita per allineare le previsioni ai margini naturali degli oggetti.

Il team ha testato il sistema su 2.150 immagini RGB ad alta risoluzione raccolte nello Xinjiang, in Cina, durante il riposo vegetativo invernale. Il set di dati comprendeva 1.250 immagini scattate nel novembre 2024 in un vigneto sperimentale della Shihezi University e 900 immagini scattate nel dicembre 2024 in un frutteto commerciale di Turpan. Tutte le immagini sono state etichettate manualmente a livello di pixel per individuare fusti e tralci. I ricercatori hanno suddiviso il set di dati in 1.505 immagini per

l’addestramento, 215 per la validazione e 430 per il test. Dopo l’aumento dei dati, il set di addestramento effettivo è salito a 4.515 immagini.

Secondo lo studio, sul set di test il modello DTMF completo ha raggiunto un’accuratezza media dei pixel del 94,24% e un’intersezione su unione media, una misura comune per la segmentazione, dell’88,25%. Nei confronti con diversi modelli consolidati, DTMF ha superato DeepLabV3, PSPNet, FCN, Segformer, U-Net e una U-Net basata su transformer. Il divario è stato particolarmente ampio nell’individuazione dei tralci, la parte più difficile del compito perché sono più sottili e più facili da confondere con lo sfondo. DTMF ha raggiunto un IoU dell’85,12% per i tralci, contro il 70,13% del modello di riferimento più performante in questa categoria, con un miglioramento di 14,99 punti. Anche la segmentazione dei fusti ha dato buoni risultati, con l’87,68%, sebbene su questa classe altri modelli moderni abbiano ottenuto valori più vicini.

I ricercatori hanno eseguito anche test di ablazione, nei quali si rimuovono alcune componenti del sistema per misurare il contributo di ciascuna. I risultati indicano che la struttura con doppio encoder e doppio decoder, il modulo di fusione wavelet-Mamba e la strategia di perdita in due fasi hanno tutti migliorato le prestazioni. Quando il modello utilizzava soltanto il ramo transformer, l’intersezione su unione media era del 71,98%. Con la configurazione completa, il valore è salito all’88,25%. In test separati sul metodo di addestramento, l’aggiunta della supervisione nelle fasi successive e dei vincoli basati sui superpixel ha portato l’IoU dei tralci dal 74,33% all’85,12%.

Gli esempi visivi presentati nell’articolo mostrano che il modello riesce a seguire i tralci anche su sfondi difficili, come fili e terreni irregolari, con meno interruzioni o segmenti mancanti rispetto ai sistemi di confronto. Gli autori hanno provato il modello anche su immagini esterne di altre piante legnose. Queste dimostrazioni suggeriscono che il metodo potrebbe avere un’utilità strutturale più ampia, ma l’articolo precisa che i risultati sono solo illustrativi e non costituiscono una verifica formale della capacità del sistema di generalizzare oltre i dati raccolti in vigneto.

Gli autori dello studio hanno sottolineato che il lavoro non rappresenta una soluzione completa per la potatura robotizzata. Una segmentazione accurata delle immagini è solo una parte del problema più ampio. Un robot per la potatura pronto per l’impiego in

campo dovrebbe ancora ricostruire in tre dimensioni la struttura della vite, individuare i punti di taglio, pianificare i percorsi di movimento e controllare un utensile da taglio nelle reali condizioni del vigneto. I ricercatori hanno inoltre rilevato che l'architettura parallela del modello aumenta i costi computazionali e potrebbe rendere più difficile l'impiego su hardware integrato montato su robot agricoli.

L'articolo segnala anche diversi altri limiti. Le immagini provengono solo da due aree dello Xinjiang e riguardano esclusivamente il riposo vegetativo invernale: il modello non è stato quindi ancora convalidato in altre regioni viticole, con sistemi di allevamento diversi o in condizioni estive caratterizzate da una folta vegetazione. La suddivisione dei dati è stata effettuata a livello di immagine, anziché per singola pianta, e il team non ha riportato prove di addestramento ripetute per misurare la variabilità tra diverse inizializzazioni casuali. Questi limiti indicano che gli ottimi risultati del benchmark vanno considerati una prova dei progressi nel riconoscimento visivo, non la dimostrazione che il sistema sia pronto per un ampio impiego commerciale.

Nonostante questi limiti, il lavoro indica una direzione concreta per l'automazione in vigneto. Se le versioni future saranno rese più rapide e convalidate in un maggior numero di contesti reali, l'approccio potrebbe contribuire a superare uno dei principali ostacoli tecnici alla potatura robotizzata nella produzione di uva. Per i vigneti che riforniscono il settore del vino e di altre bevande a base d'uva, questo potrebbe in futuro favorire una pianificazione più stabile della manodopera e una gestione più uniforme delle viti, anche se il risultato dipenderà ancora dall'integrazione sul campo e da test che vadano oltre la sola segmentazione delle immagini.