

Scaletta Vision Framework for iOS

- **Presentazione**

- Come argomento opzionale ho deciso di portare Vision Framework in quanto, oltre ad essere un amante di iOS, ritengo sia uno strumento molto potente e molto diffuso al giorno d'oggi.
- È un framework appunto per iOS abbastanza recente, presentato per la prima volta nella Apple conference del 2017 e aggiornato e ripresentato ogni anno una sua versione potenziata.

- **Che cos'è?**

- È un framework high-level nel campo della Computer Vision. Viene presentato da Apple come un'unica e consistente interfaccia, nonostante offra numerose modalità di tracking dell'immagine.
- Spesso viene integrato/utilizzato insieme ad un altro framework, ovvero Core ML (Machine Learning) per la creazione di modelli personalizzati. Verrà approfondito questo aspetto più avanti.
- Cosa si intende con "high-level"? Io, semplice programmatore, mi pongo la domanda "Voglio solo sapere... dove sono i volti?", Vision dice "ci penso io, non ti preoccupare del come lo faccio". Infatti, non è necessario essere degli esperti in Computer Vision, la complessità degli algoritmi di riconoscimento li gestisce lui.

- **Cosa posso fare con Vision?**

- Tra le principali funzionalità di Vision troviamo:
 - **Face Detection:** ovvero la capacità di riconoscere diversi volti in un'immagine, anche se sono di profilo o parzialmente coperti, ad esempio da occhiali o mascherina.
 - **Face Landmarks:** non solo è in grado di riconoscere i volti, ma anche di tracciarne i lineamenti, come il mento, gli occhi, il naso, la bocca e di tracciare così anche le espressioni facciali.
 - **Barcode Detection:** classico riconoscimento dei barcode, che conosciamo benissimo tutti.
 - **Text Detection:** in uno degli ultimi aggiornamenti del 2021 è stata aggiunta anche la funzionalità di riconoscere del testo nelle immagini. Con iOS 15.0 infatti è possibile prendere e copiare del testo presente in un'immagine direttamente dall'app Foto.
 - **Object Tracking:** è in grado di riconoscere degli oggetti animati, o comunque in movimento, e di tracciarli. Un esempio di applicazione potrebbe essere la messa a fuoco continua durante la cattura di un video di una persona che pratica dello sport.

- **Come funziona nella pratica?**

- Si prende un'immagine qualunque.
- Riconosce che c'è un rettangolo contenente del testo.
- Rielabora in background il rettangolo cercando di estrarre dalla prospettiva 3D e di trasportarlo su un piano 2D.
- Tramite algoritmi di intelligenza artificiale fornisce dei risultati riguardanti il contenuto del testo, dando anche un valore dell'accuratezza.

- **Lato codice**

- Se si guarda il codice salta subito all'occhio il discorso della semplicità, dell'high-level solution... in sole quattro righe si possono riconoscere dei volti in un'immagine.

- **Vision e Core ML**

- Riprendendo il discorso accennato prima, Vision viene spesso utilizzato insieme ad un altro framework che è Core ML, ovvero uno strumento che permette di applicare algoritmi di Machine Learning alle immagini nel nostro caso, ma ha anche altre applicazioni.
- È così possibile creare dei modelli per poter classificare le immagini. Ad esempio, con iOS 15.0, direttamente nell'app Foto è possibile ricercare le immagini presenti nella propria galleria in base al loro contenuto (es. fiumi, alberi, spiaggia, persone, documenti, ecc.).
- Questo procedimento di classificazione dell'immagine viene fatto automaticamente da Vision ogni volta che viene scattata una foto con la fotocamera o se ne seleziona una dalla libreria. In particolare, Vision ridimensiona e ritaglia la foto in modo da adattarla ad un input standard per Core ML. Quando Core ML genera una predizione sul contenuto dell'immagine, la fornisce a Vision che presenta i risultati all'utente.
- Ecco un esempio di riconoscimento degli oggetti all'interno delle immagini (1 – farfalla, 2 – broccoli, 3 – margherita gialla)

- **Conclusione**

- Questo è Vision, framework dell'Apple per iOS.
- Tutte le info/immagini sono state prese dalla pagina ufficiale Apple dedicata a Vision.