**1. Introduzione**

In un’epoca dove i grandi colossi dell’informatica hanno ormai monopolizzato i dati, raccogliendone enormi quantità dagli utenti delle loro piattaforme e accaparrandosi un sostanziale vantaggio su tutti i competitors, sviluppare un prodotto competitivo può risultare quasi impossibile per un’organizzazione con un bacino d’utenza non molto ampio e priva delle quantità di dati di cui dispone la concorrenza.

Proprio per risolvere il problema della carenza di dati, nascono i dataset sintetici: collezioni di dati generati da calcolatori in grado di simulare una raccolta dati in campo reale. Questi dataset rendono possibile per chiunque perfezionare i propri algoritmi in maniera efficace e veloce: non si devono, infatti, aspettare dati provenienti dal mondo reale per iniziare il training dei propri algoritmi.

Ovviamente, anche i dati sintetici hanno i loro difetti: il rischio, detto reality gap, di un algoritmo allenato con soli dati sintetici è quello di non sapersi adattare alle situazioni reali.   
Questo possibile problema è stato per anni un deterrente per l’uso dei synthetic data nelle applicazioni, tuttavia, con il miglioramento nel corso degli anni degli algoritmi di generazione di dati, il gioco ha cominciato a valere la candela grazie all’ormai ottima precisione che i dataset sintetici riescono a raggiungere a fronte di costi e tempi neanche minimamente comparabili a quelli di una raccolta dati sul campo.

Nell’ambito della computer vision, dove si colloca il lavoro di tesi che questa introduzione vuole presentare, un dataset sintetico può definirsi tale se composto da un’immagine, contenente gli oggetti la cui presenza è di interesse dell’algoritmo da allenare, e da un’etichetta che descriva quanto contenuto nell’immagine con l’obiettivo di “insegnarlo” all’algoritmo.

Questo lavoro di tesi si pone l’obiettivo di costruire, a tal proposito, un generatore di dati sintetici.   
Tale generatore verrà poi sfruttato per creare un dataset che andrà ad allenare un algoritmo di riconoscimento, con lo scopo di insegnargli ad individuare le mani che vede nelle immagini che andrà ad analizzare.

Per raggiungere tale scopo, sono stati utilizzati gli strumenti descritti nei prossimi paragrafi.

**1.1 Unity 3D**

Unity 3D è un motore grafico progettato principalmente per lo sviluppo in C# di videogiochi ambientati in uno spazio tridimensionale.  
Essendo lo spazio tridimensionale offerto dall’engine popolabile e modificabile a proprio piacimento, Unity risulta molto versatile e le sue applicazioni toccano svariati settori, tra i quali è presente anche il machine learning.

Unity, pur non essendo il miglior motore grafico in circolazione a livello di qualità dei render o di performance, risulta perfetto per la generazione di dataset sintetici per la computer vision: se configurato correttamente, infatti, è in grado di generare numerose immagini molto diverse tra loro e di labellarle in pochi e non molto complessi passaggi.

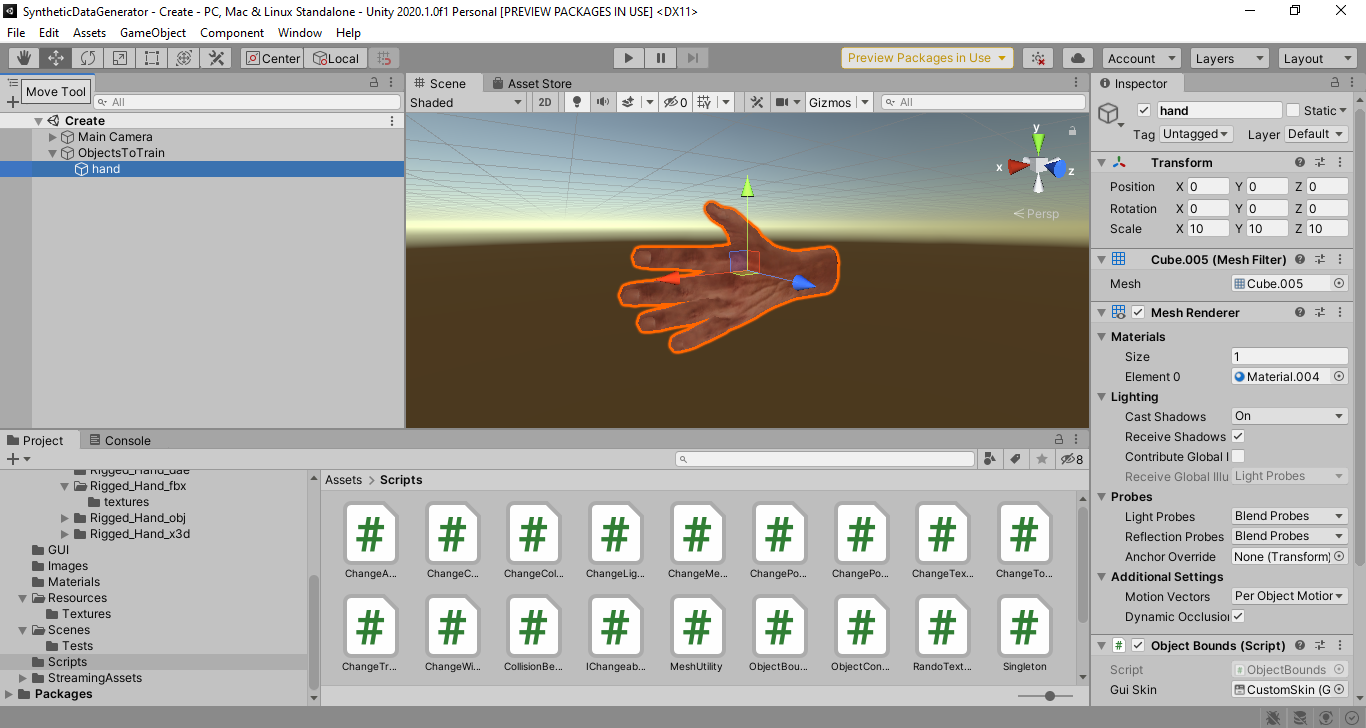


Figura 1.4 – schermata principale di Unity 3D

Fondamentali, per quanto riguarda questo programma, sono le *scene*, ossia la rappresentazione dell’ambiente (ad esempio un giardino con una mano), che permettono di visualizzare gli oggetti in 3D.

All’interno di essa ci si può muovere e cambiare il punto di vista, tenendo premuto il tasto destro del mouse e spostandosi nella direzione desiderata. In questo modo è come se si girasse la telecamera nelle 4 direzioni disponibili su di un piano immaginario. Inoltre è possibile effettuare lo zoom per avvicinare o allontanare la telecamere. Ovviamente sono consentite altre operazioni, ma ai fini del nostro progetto, queste sono le principali cui è stato fatto ricorso.

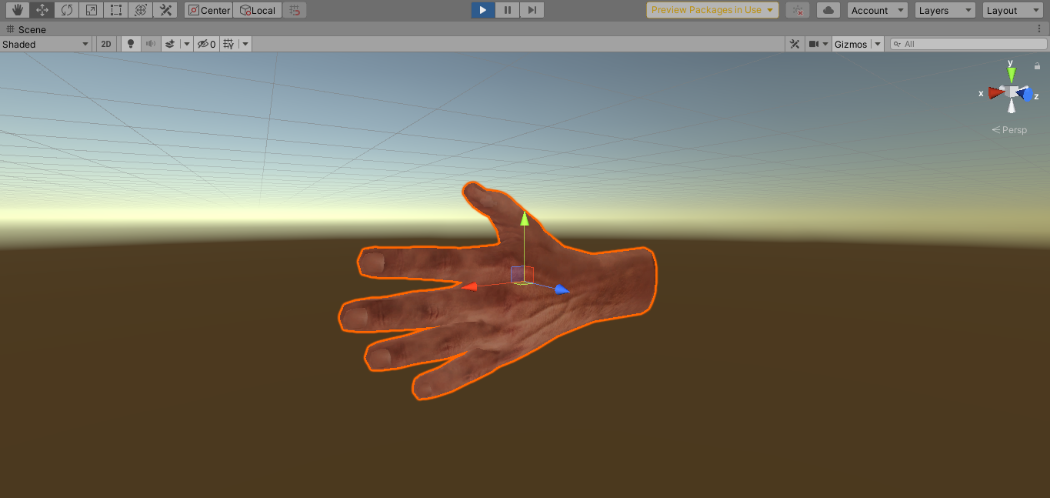


Figura 1.5 – esempio di scena

Aspetto importante sono i file .meta, creati da Unity nel momento in cui importiamo l’oggetto nel workspace. Dettagliano come avverrà l’importazione del file a cui fanno riferimento.

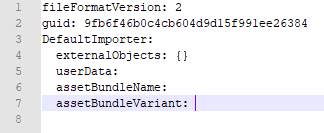


Figura 1.6 – esempio di file .meta

Simultaneamente a Unity è stato utilizzato Fatkun Batch DownloadImage, un’estensione di Google Chrome che consente di scaricare le numerose immagini sulle schede aperte durante il download.