

Detecció moviment d'una mà pel control d'un videojoc

Marina Bermúdez Granados, Adrian Vargas Orellana,
Laia Rubio Castro

Abstract — Per fer quan s'hagi completat el projecte

Keywords — Per fer quan s'hagi completat el projecte



1 INTRODUCCIÓ

En aquest document està detallat el desenvolupament del projecte final de l'assignatura Visió per Computador. El tema escollit tracta sobre el control dels moviments a un videojoc 2D amb diferents gestos d'una mà captats per una càmera. Per tant, el projecte es pot dividir en dues parts: La captació i classificació dels gestos de la mà, per després associar cada classe a una acció a fer al videojoc, com caminar o saltar uns obstacles.

La idea inicial va sortir de vídeos que el grup va trobar sobre altres persones intentant jugar a videojocs sense cap comandament, només amb una càmera i una mica de programació. A partir d'aquests projectes, es va decidir intentar controlar un videojoc amb imatges captades i després ens vam decantar per fer el control amb els gestos de les mans.

L'objectiu principal del projecte se centrarà a fer una classificació prou acurada i ràpida com a poder passar aquestes dades a un videojoc, similar a com es faria normalment amb un comandament.

2 ESTAT DE L'ART

Dins d'aquest projecte hi ha diverses tècniques que caldrà posar en pràctica. Primer cal estudiar com fer la detecció de la mà per poder captar i classificar les imatges corresponents. En acabar, s'haurà de passar aquesta classificació al videojoc per fer l'acció corresponent.

2.1 Detecció i Classificació d'una mà

Les tasques de detecció i classificació d'una mà segueixen normalment les mateixes tècniques. Tot i això, cal diferenciar entre classificar un gest o classificar els moviments necessaris per fer un gest. En aquest projecte es captarà frame a frame les imatges de la mà per detectar de manera binària si s'està fent un gest o no. Quan es detecti un gest per un mínim determinat de frames, llavors contarà com que si s'està fent.

Llavors, per detectar el gest d'una mà estàtica hi ha diverses implementacions que es poden fer servir. En primer lloc, es

capta el frame i es processa perquè sigui més fàcil de tractar. Les implementacions manuals investigades passen la imatge a escales de grisos, per aplicar un threshold i tractar d'obtenir el contorn de la mà. També es poden treure les característiques en compte dels contorns, depenent de com es vol fer la classificació amb la informació extreta. Les llibreries trobades que s'ocupen del procés enter aconsegueixen pintar les guies de les falanges amb els punts importants detectats, per donar el percentatge de classificació més alt segons com estan col·locats. Tot i que només es té previst fer servir una mà, les llibreries també són capaces de diferenciar entre les mans dretes i esquerres. Aquest procés s'executa mitjançant models d'aprenentatge computacional.

2.2 Tècniques de Segmentació

Per la detecció de les mans dins d'una imatge o frame, es poden utilitzar tècniques de segmentació per separar-les del fons, i després analitzar i detectar la forma de la mà per saber quina acció fer. Es poden usar tècniques de segmentació clàssica, però hi existeix una altra opció per poder centrar el projecte només a l'anàlisi de les mans.

L'opció més simple per la segmentació de les imatges és fer servir SAM, *Segment Anything Model*^[13], un model fet per Meta que com diu el nom, segmenta tot. Amb aquesta eina es podria aconseguir que aquest model segmenti les mans directament, per després poder fer una xarxa neuronal d'anàlisi la qual pot ser entrenada amb les nostres imatges generades i classificades. En cas que tingui un cost molt elevat computacionalment el fet de fer servir SAM, es tornaria a les tècniques clàssiques de segmentació.

2.3 Control del videojoc

Els vídeos mencionats anteriorment que han inspirat aquest projecte ja serveixen com a exemples sobre com controlar un videojoc sense comandament. Tot i això, els primers feien servir detecció de postures o el moviment d'un cap a l'esquerra i la dreta. En fer una mica més de recerca, s'han trobat un terme per descriure els videojocs on es fan servir les mans pel seu control: *Gesture-based Games*. La majoria dels projectes trobats s'han desenvolupat en *Python*, tot i que el llenguatge de programació depèn en el videojoc.

3 PROPOSTA

La nostra idea és utilitzar una base de dades creada per nosaltres, la qual constarà d'imatges de les nostres mans en diferents posicions així com, amb diferents fons. Això ens permetrà entrenar millor al nostre detector, ja que haurà de tenir en compte més posicions i possibles objectes o colors diferents de fons. Amb aquestes dades crearem un entorn de train i un de test per a detectar i captar les imatges per posteriorment fer la seva classificació.

Pel que fa al codi, farem ús de les llibreries trobades durant la recerca de referents (numerades al final del document) per a crear els nostres models.

[12] OpenMMLab. (n.d.). OpenMMLab. Retrieved April 28, 2023, from <https://openmmlab.com/>

[13] Meta AI. (n.d.). Segment Anything. Retrieved April 28, 2023, from <https://segment-anything.com/>

BIBLIOGRAFIA

[1] Tech, R. (2023). Uso Python e IA para jugar sin control y se rompe la silla [Video]. In *YouTube*.
<https://www.youtube.com/watch?v=vroIHqqz3v0>

[2] Tech, R. (2023). Intento hacer juegos con Python e IA y me explota la cabeza [Video]. In *YouTube*.
https://youtu.be/_BjL6W71mWY

[3] Aslananuragi. (2021, June 16). *Gesture controlled video game*. Analytics Vidhya.
<https://www.analyticsvidhya.com/blog/2021/06/gesture-controlled-video-game/>

[4] atul, kang &. (2019, March 4). *Snake Game using Tensorflow Object Detection API*. TheAILearner.
<https://theailearner.com/2019/03/04/snake-game-using-tensorflow-object-detection-api/>

[5] Wang, E., & Arif, M. (n.d.). *Gesture controlled gaming*. Devpost. Retrieved April 28, 2023, from
<https://devpost.com/software/gesture-controlled-gaming>

[6] *Free 3 cyberpunk characters pixel art*. (n.d.). CraftPix.Net. Retrieved April 28, 2023, from
<https://craftpix.net/freebies/free-3-cyberpunk-characters-pixel-art/?affiliate=163304>

[7] Haria, A., Subramanian, A., Asokkumar, N., Poddar, S., & Nayak, J. S. (2017). Hand gesture recognition for human computer interaction. *Procedia Computer Science*, 115(115), 367–374.
<https://doi.org/10.1016/j.procs.2017.09.092>

[8] Haroon, M., Altaf, S., Ahmad, S., Zaindin, M., Huda, S., & Iqbal, S. (2022). Hand gesture recognition with symmetric pattern under diverse illuminated conditions using artificial neural network. *Symmetry*, 14(10).
<https://doi.org/10.3390/sym14102045>

[9] Pol, S., Pagade, P., & Pati, D. (2017). Gesture Recognition Based Video Game Controller. *International Research Journal of Engineering and Technology (IRJET)*, 04(11).
<https://www.irjet.net/archives/V4/i11/IRJET-V4I11182.pdf>

[10] MediaPipe. (2023, January 19). *Gesture recognition task guide*. Google Developers.
https://developers.google.com/mediapipe/solutions/vision/gesture_recognizer

[11] hukenovs. (n.d.). *GitHub - Hukenovs/hagrid: HAnd Gesture Recognition Image Dataset*. GitHub. Retrieved April 28, 2023, from <https://github.com/hukenovs/hagrid>