

Rozpoznávání gest pomocí NVIDIA Jetson Nano

D. Pojhan¹, D. Majrich², and M. Sojka³

¹Gymnázium, Plzeň, Mikulášské nám. 23, danpojhan@gmail.com

²SPŠSE Dukelská 13, České Budějovice, denis.majrich@gmail.com

³SPŠSE Dukelská 13, České Budějovice, miroslav.sojka@spssecb.cz

Abstrakt

Rozpoznávání gest, které přímo souvisí s odhadem pózy a gest, je dnes hodně důležitý obor v hlubokém učení a umělé inteligenci. Cílem našeho projektu bylo seznámit se s vývojovým prostředím NVIDIA Jetson Nano a algoritmy odhadu pózy jako je ResNet18 a DenseNet121. Výstup naší práce není jen pochopení pomůcek pro vývoj rozpoznávání gest, ale vytvořili jsme i program, který s vámi dokáže hrát kámen, nůžky, papír. Výsledky naší práce se nachází na:

<https://github.com/HelloWorld7894/GestureDetection>

1 Úvod

Odhad pózy a gest, který vytváří skelet objektu, je stále jeden z nejtěžších oborů počítačového vidění. Teprve nedávno (NeurIPS 2022) byl předveden nový state-of-art model ViTPose, který dosahoval pouze 81 AP * na datasetu COCO. Při našem testování jsme používali algoritmy ResNet18 a DenseNet121.

* (Average Precision → metrika, používající se pro měření přesnosti modelů hlubokého učení)

2 NVIDIA Jetson Nano

Zařízení, na kterém jsme celý projekt testovali, byl malý jednodeskový počítač s GPU. Právě díky tomu je náš projekt přenosný a díky GPU s akcelerací výpočtů AI i dostatečně rychlý, NVIDIA zároveň nabízí SDK (software development kit) pro deep learning.

specifikace:

| | |
|------------|-----------------------------|
| GPU | 128-core Maxwell GPU |
| CPU | Quad-core ARM Cortex-A57 |
| paměť | 4 GB LPDDR4 1600MHz |
| CUDA | 10.2. |
| komunikace | GPIO, I2C, SPI, UART, USART |
| OS | Ubuntu 18.04 LTS |

GPU - Graphical Processing Unit

CPU - Central Processing Unit

CUDA - Compute Unified Device Architecture (nástroj pro akceleraci trénování a inference neuronové sítě)

GPIO - General Purpose Input/Output (elektrický kontakt přes který prochází komunikace v integrovaném obvodu či jednodeskovém počítači)

I2C - Inter-Integrated Circuit *

SPI - Serial Peripheral Interface *

UART - Universal Asynchronous Receiver/Transmitter *

USART - Universal Synchronous/Asynchronous Receiver/Transmitter *

* (Komunikační protokol mezi čipy a dalšími elektronickými součástkami na plošných spojích)

3 Neuronové sítě

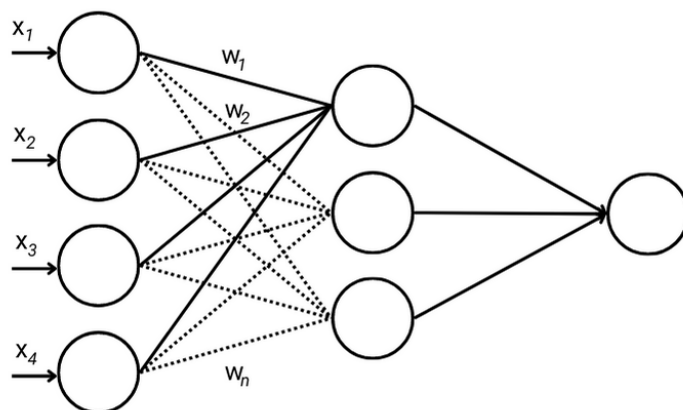
S rozšiřováním dostupných informací na internetu a komplexitou všech dat se kterými člověk musí pracovat, se také zvětšila časová náročnost různých úloh, které standardní algoritmy nezvládají. Díky těmhle problémům vznikly neuronové sítě.

Architektura těchto algoritmů je inspirována fungováním lidského mozku, a právě díky tomu se dokážou velmi dobře učit neurčitá data. Jejich nejběžnější aplikace jsou v klasifikaci a regresi, odkud se pak odvíjí více oborů jako např. sémantická segmentace, odhad pózy, klasifikace objektu, detekce objektu, atd.

Struktura

Stavební bloky neuronové sítě jsou tzv. neurony, které spolu vytváří vazby s různými váhami. Takže pokud bychom např. chtěli vypočítat hodnotu neuronu do kterého vedou váhy w_1, w_2, \dots, w_n , rovnice bude vypadat:

$$y_1 = f_a((w_1 * x_1 + b_1) + (w_2 * x_2 + b_2) + \dots + (w_n * x_n + b_n)) = \sum_{i=1}^n w_i * x_i + b_i \quad (1)$$



kde význam symbolů je následující:

$y_1 \leftarrow$ výstup neuronu

$f_a \leftarrow$ aktivační funkce neuronu, zavádí "nelinearity" do neuronové sítě

$b_i \leftarrow$ bias neuronu *

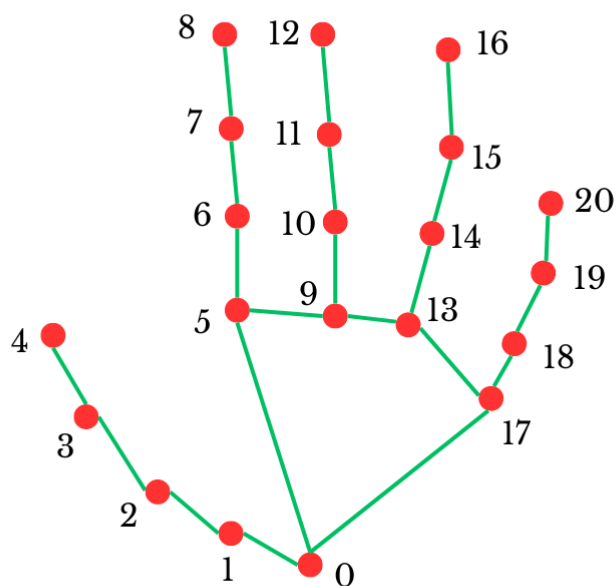
$w_i \leftarrow$ váha vazby *

$x_i \leftarrow$ vstup

* (nastavovatelné parametry neuronové sítě, které se trénováním mění)

Všechny parametry, které jsou značeny *, se pak následně mění pomocí algoritmu sestupu gradientu.

4 Odhad pózy ruky



Náš přístup pro odhad pózy ruky používá pouze předdefinované matematické rovnice z výstupu neuronové sítě u některých bodů ruky. Tohle řešení má své meze, protože používá pouze přesně dané hodnoty, avšak výstup se může podle prostředí měnit. Tenhle problém by šel eliminovat natrénováním další neuronové sítě, která by brala výstup z té dosavadní a fungovala by jako nadstavba. K tomuhle řešení jsme však nestihli dojít, protože jsme neměli dostatek času na natrénování dané sítě.

