

Прв парцијален испит по предметот Методологија на истражување во ИКТ

Андреј Миловановиќ 171033

Абстракт—Во овој документ се наоѓаат одговорите на првите две прашања од првиот колоквиум по предметот Методологија на истражување во ИКТ. Одговорот на третото прашање можете да го пронајдете во Jupyter Notebook документот прикачен на мојот GitHub проект.

I. Вовед

Трудот кој го одбрав за овој колоквиум е Fan et al. [1], каде се зборува за сличноста помеѓу репрезентациите на објекти при цртање и при визуелна обсервација на истите. Се користат слики од 105 различни објекти и рачно нацртани скици од истите објекти. Со овие податоци се креира модел на машинско учење кој користи длабока невронска мрежа со цел да ги препознае објектите од сликите и скиците, и им користи на научниците за добивање на интересни заклучоци за перцепцијата на луѓето, за начинот на кој визуелниот кортекс креира репрезентација за објектите при цртање и при набљудување на истите и корелацијата на таквата репрезентација во двата случаи. Повеќе детали за овој труд може да добиете во вебеспоменатиот Jupyter Notebook.

II. Методологија

Во оваа секција ќе ги дадам одговорите од второто прашање од колоквиумот, односно опис на методологијата на избраниот труд. Најпрво, истражувањето опишано во [1] е квантитативно, бидејќи користи статистички методи и тестови за добивање на заклучоци, врз податоци во форма на слики, кои фундаментално претставуваат матрици. Како што

претходно напоменах, сликите се делат на рачно нацртани скици, и фотографии. Овие податоци се собирани на два начини: 8400 скици од 105 вообичаени реални објекти се нацртани од страна на платени учесници во истражувањето, додека останатите 22,843 слики се фотографии преземени од Imagenet податочното множество [2].

Во овој труд се разгледуваат неколку хипотези, од кои главната (од која се подоцна изведени и останатите хипотези, со цел вршење на подетална анализа и добивање поточни заклучоци) е следната: Човековиот визуелен кортекс ги користи истите репрезентации на карактеристиките на објектите при визуелно препознавање (набљудување) и исцртување на истите. Ова се испитува користејќи длабока невронска мрежа, која ги споредува рачно нацртаните скици од објектите со фотографии од истите, разгледувајќи ја корелацијата помеѓу нивните одлики, претставени во форма на вектори на карактеристики. Како резултат на истражувањето, хипотезата се прифаќа, а статистичкиот критериум кој се користи за нејзиното прифаќање е Пирсоновиот тест за корелација помеѓу векторите на карактеристики на скиците и фотографиите. Големите коефициенти на корелација не носат до заклучокот дека истите карактеристики на високо ниво на објектите кои ги набљудуваме (визуелно препознаваме) се наоѓаат и во цртежите на истите објекти.

Трудот е исто така полн со визуелизации. Се визуелизира моделот на длабоко учење заедно со неговите слоеви. Се визуелизираат матриците

на Пирсоновата корелација помеѓу карактеристиките на објектите во различни слоеви на длабоката невронска мрежа. Се користи scatter plot за да се исцрта корелацијата помеѓу препознавањата на објектите од страна на човекот (x-оска) и длабоката невронска мрежа (y-оска). Се користи line plot со додадени error bars за да се визуелизира растот на корелацијата на карактеристиките на објектите со зголемување на бројот на слоеви во длабоката невронска мрежа.

Литература

- [1] J. E. Fan, D. L. Yamins, and N. B. Turk-Browne, "Common object representations for visual production and recognition," *Cognitive science*, vol. 42, no. 8, pp. 2670–2698, 2018.
- [2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.