

Употреба неуронске мреже у класификацији слика

Љубић Јанко SW-38/2016
Факултет Техничких Наука у Новом Саду

Увод

Улазни подаци састоје се од 320 различитих фотографија цвећа. Укупно се појављује 5 различитих типова цвета а сваки тип, заједно са називом слике на којој се налази описан је у улазном документу који садржи исправне информације о називу сваког цвета на свакој слици. Овај скуп података се користи за тренинг класификатора.

Тест скуп података састоји се од 80 различитих фотографија цветова из истих оних 5 различитих типова. Такође уз тест директоријум са сликама налази се и тест документ са информацијама о називу сваког цвета на свакој тест слици.

Тренинг скуп и тест скуп су међусобно дисјунктни.



Методологија рада

На почетку су подаци учитани на начин који је обезбедио правилан и паралелан поредак назива цветова са одговарајућим фотографијама из тренинг скупа. То је учињено тако што су потребни подаци прво читани из 'csv' датотеке и коришћени да се паралелно пронађе фотографија са датим називом и назив њеног цвета (односно лабела) и да се паралелно похране у одговарајућу колекцију података.

На идентичан начин на почетку програма учитани су и тест подаци. Како су прочитане фотографије? Свака од фотографија има своју независну резолуцију, било је од користи скалирати их на подједнаку резолуцију што је и учињено и након читања резолуција похрањених фотографија износи **100X100**. Када говоримо о карактеристикама и посебности сваког цвета морамо узети у обзир и његову боју, из тог разлога слике су учитане у боји, 'RGB'.

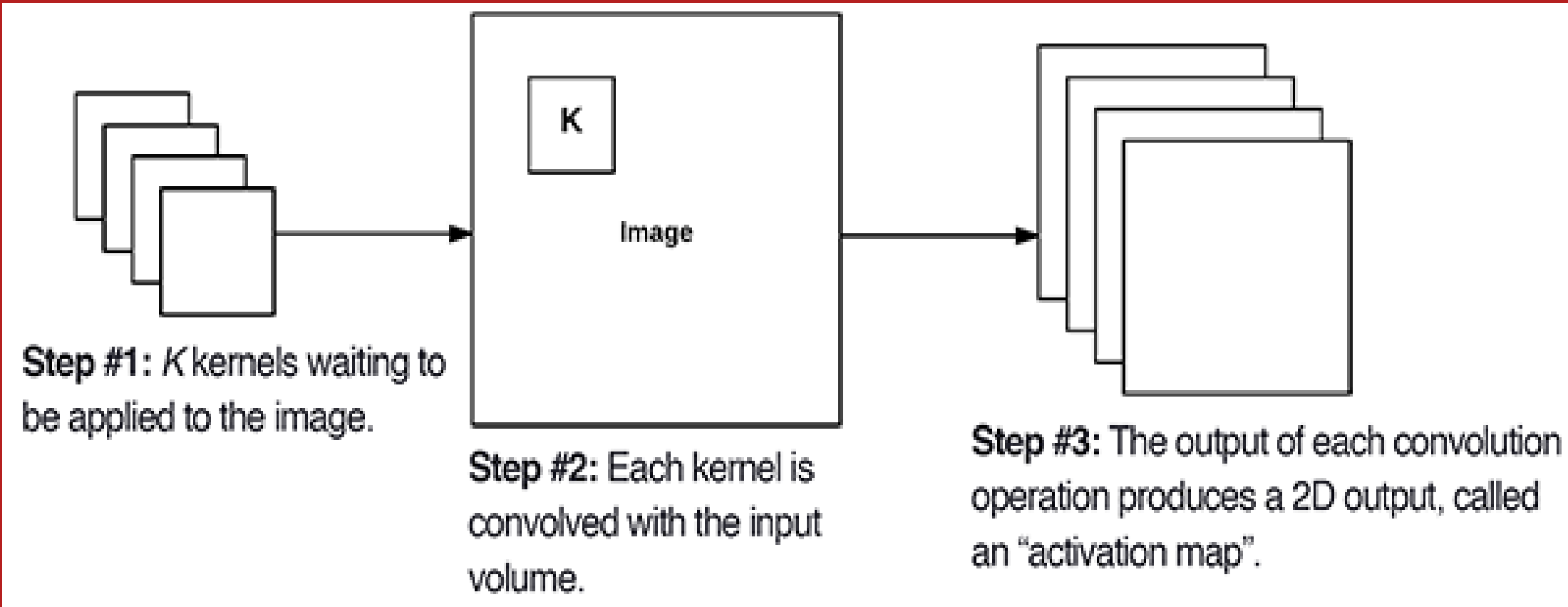
Било је потребно направити модел неуронске мреже и у њега убацити слојеве који ће прочитане податке обрађивати.

За издвајање карактеристика задужен је конволуциони слој **CONV2D** који применом одређеног броја (K) филтера матрице фотографија пресликава у активационе мапе које на себи имају израженије карактеристике (вертикална ивица, хоризонтална ивица, кругови, сенка...).

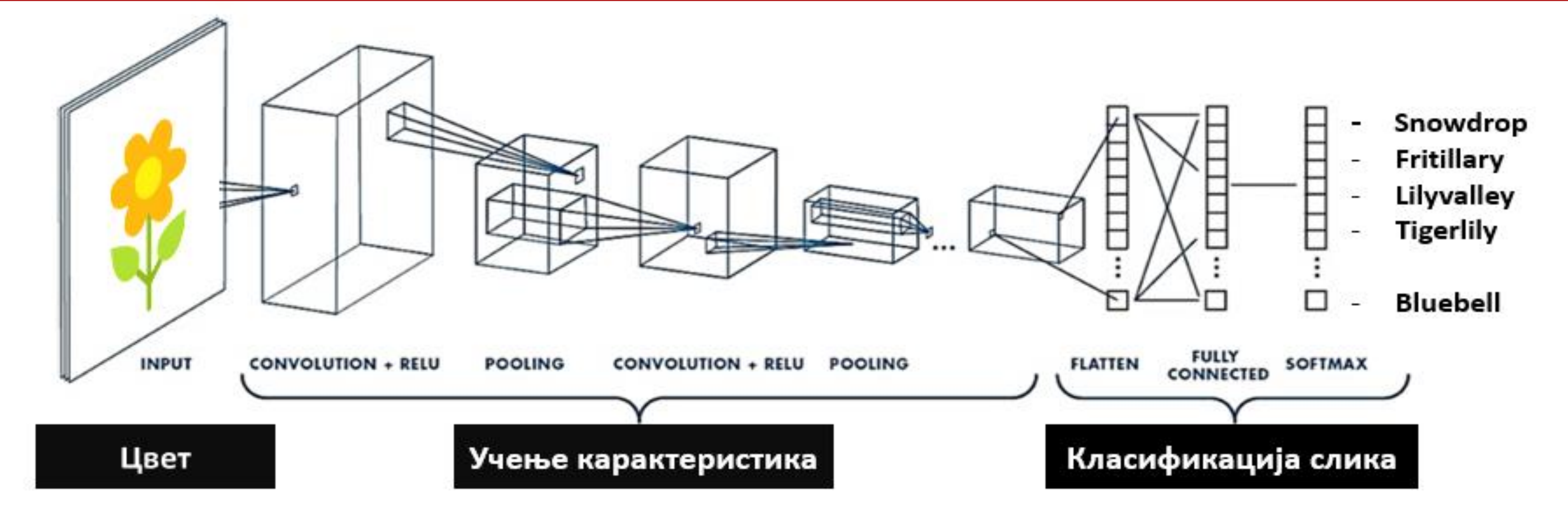
Методологија рада

Након овог слоја примењује се нелинеарна активациона функција 'ReLU'. Да бисмо од излаза могли извући неке закључке потребно је да их скалирамо и задржимо најважније карактеристике одбацујући оне небитне заостале од претходних слојева, за то се користи слој под именом 'Max Pooling'.

Након три итерације са комбиновањем горенаведених слојева улазимо у 'Fully connected' слој који „изравна“ излазе из претходних слојева и претвори их у јединствени вектор.



Шематски приказ протока фотографије кроз неуронску мрежу од улаза до излаза



Резултати

Након што се мрежа истренира, на њу се може применити тестирање уз помоћ 80 фотографија из тест скупа и упоредити излаз мреже са исправним именима сваког од 80 цветова.

Резултат тог поређења изражава се у проценту тачности. Неуронска мрежа коришћена у овом решењу није конзистентна у оцени решења. Најбољи резултат који постиже је 86,25%. То значи да од 80 фотографија она исправно препознаје 69, односно да греша у 15.75% случајева, код 11 цветова.

У поређењу са решењем направљеним користећи 'HOG DESCRIPTOR' неуронска мрежа је доста успешнија јер је у просеку 20% тачнија.

Закључак

Неуронска мрежа представља најбоље решење за класификацију слика када је у питању више различитих класа ентитета на фотографијама.

Конкретно решење које користи 'HOG DESCRIPTOR' може бити коришћено уколико радимо са на пример 2 ентитета.

Неуронска мрежа је повољна за мноштво комбинација јер се распоред слојева у њој може мењати и тиме у многоме варијирати понашање читавог скупа.

Тренинг скуп у овом случају није био довољно обиман јер скуп од 320 фотографија већ у 13. епохи доводи до прецизности тренинга од 100% и једино решење за то у овом тренутку јесте повећање тренинг скупа података.