

Klasifikacija bolesti biljaka na osnovu slike lista

Uroš Petrić, Ivan Luburić

Veljko Maksimović

UVOD

U ovom projektu bavili smo se temom detekcije bolesti biljaka na osnovu slike lista biljke. Za izradu projekta korišćen je Python a od tehnologija je korišćeno Duboko učenje (**Deep learning**) zajedno sa Konvolucionim neuronskim mrežama (**Convolution neural network – CNN**).

Duboko učenje smo koristili zbog rada sa velikom količinom podataka koji su predstavljeni putem slika a CNN smo koristili da bi lakše prepoznali karakteristike listova kao i bolesti sa slike.

DATASET

Prilikom izrade projekta je korišćen dataset koji se sastoji od **10000 slika** podeljenih u **39 vrsta** koje predstavljaju oboljenja na bilkama.

Dataset je na samom početku podeljen u 3 grupe, odnosno skupa:

- **Trening skup**
- **Validacioni skup**
- **Testni skup**

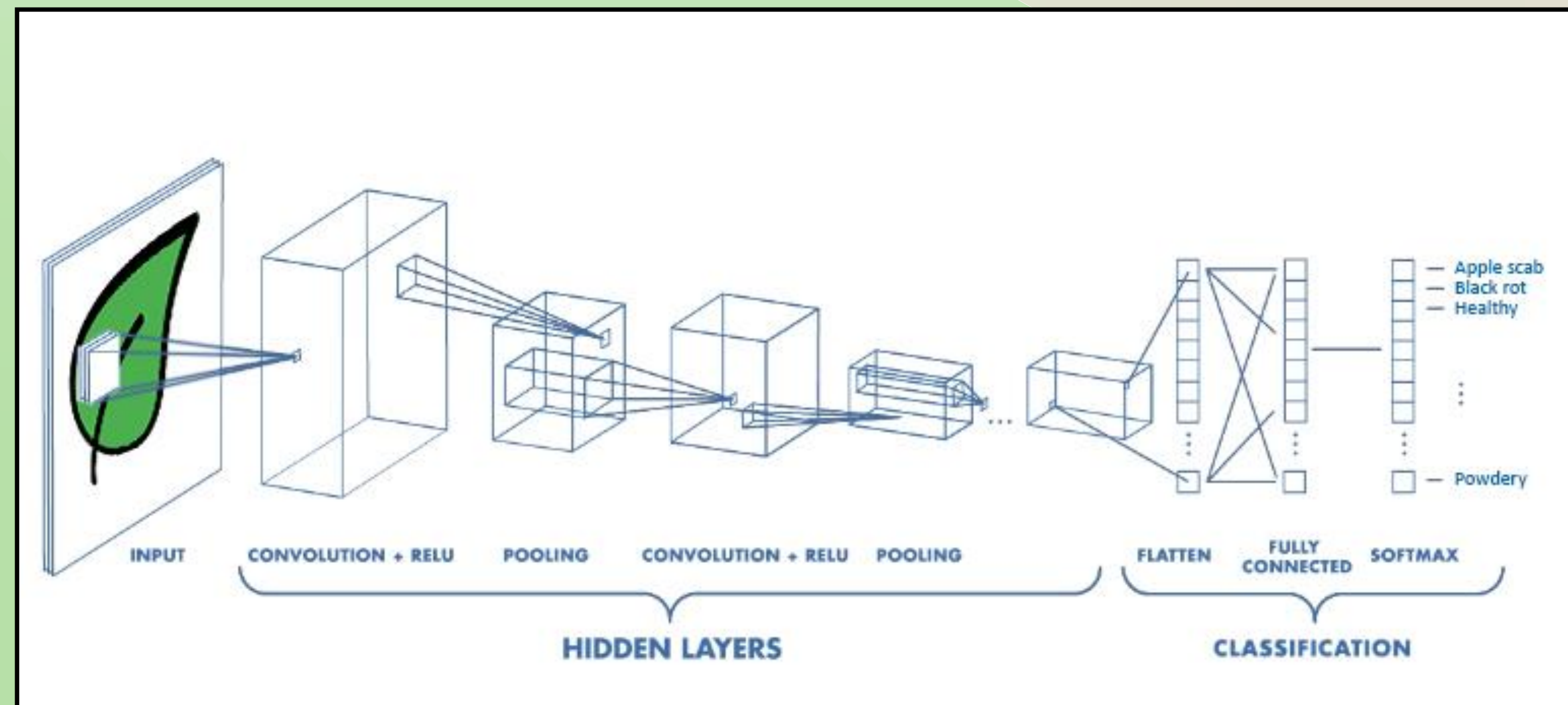


KREIRANJE MODELA CNN

Na samom početku je veličina svake slike promenjena na 224 x 224 nakon čega su ubačene u konvolucione neuronske mreže. Korišćene slike su u boji da bi mogli da koristimo 3 kanala RGB skale boja.

U prvom konvolucionom sloju uzimamo da veličina filtera, odnosno izlaznih kanala, bude 32. Ovim smo učinili da se 32 različita filtera primenjuju na slici pokušavajući da nađu i izdvoje karakteristike sa slike nakon čega kreiramo mapu karakteristika koja ima 32 kanala. Dakle od slike veličina 3 x 224 x 224 dobijamo sliku 32 x 222 x 222. Nakon toga se primenjuje **ReLU aktivaciona funkcija** koja će otkloniti ne-linearnost. Potom se koristi **Batch Normalization** koja normalizuje težinu neurona. Na kraju ovog sloja se na slici primenjuje **max pool sloj** koji će izdvojiti samo najbitnije karakteristike vezane za sliku da bi na kraju imali sliku veličina 32 x 112 x 112 koja se prosleđuje u sledeće slojeve konvolucionog algoritma koji rade po istom principu kao i prethodno opisani.

Na samom kraju će se izravnati izlazni podaci max pool sloja i proslediće se u sledeći linearni sloj koji se zove **fully connected sloj**, i konačno, kao finalni sloj, predviđamo **39 kategorija**. Što znači da kao izlaz modela imamo skalu 1x39. Kao izlaz iz modela dobijamo niz dužine 39, i iz tog niza uzimamo indeks elementa sa maksimalnom vrednošću. Taj indeks predstavlja našu predikciju.



REZULTATI TRENINGA

Kao rezultat treniranja dobijamo fajl pod nazivom **‘plant_disease_model_1_latest.pt’** u kome se nalaze podaci potrebni za predikciju bolesti lista. Tačnost predikcije se meri uz pomoć funkcije accuracy gde smo kao rezultat dobili:

- Train Accuracy: **87.5%**
- Test Accuracy: **83.8%**
- Validation Accuracy: **84.4%**

Tačnost sistema se može unaprediti povećanjem dataseta-a na veći broj podataka koji će biti korišćeni za treniranje mreže ali zbog hardverskih ograničenja nismo bili u mogućnosti za povećanje istog.

REFERENCE

- „Plant Disease Detection Using Convolutional Neural Networks with PyTorch“ by Manthan Bhikadiya
- <https://github.com/manthan89-py/Plant-Disease-Detection>