

Академија струковних
студија Шумадија
одсек у Крагујевцу

од 2019. године

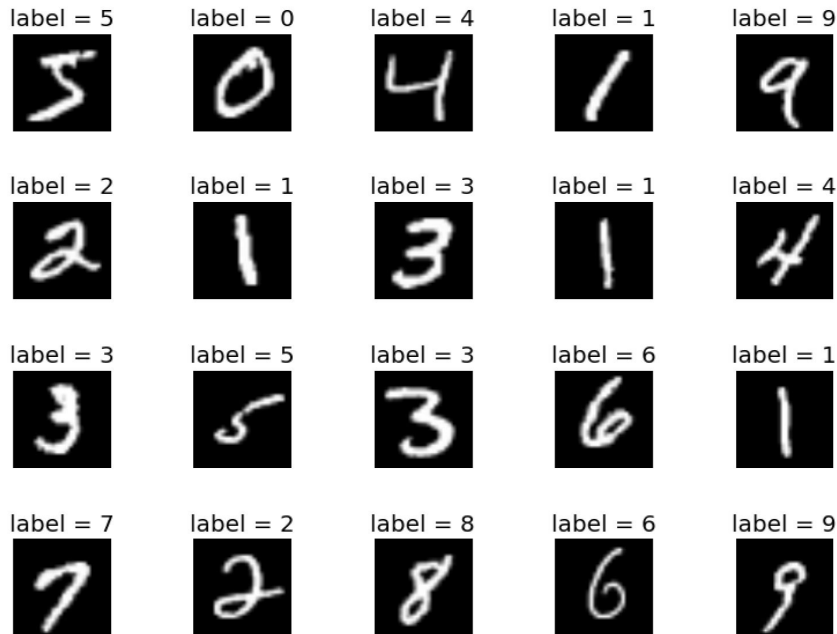
ПРЕПОЗНАВАЊЕ РУЧНО ПИСАНИХ ЦИФАРА ПОМОЋУ НЕУРОНСКИХ МРЕЖА

Ментор:
Др. Владимир Недић

Студент:
Филип Лазовић
094/2014

Проблем

- МНИСТ база података
- 65,000 означених примерака
цифара од 0 до 9
- Црно-беле слике резолуције 28x28



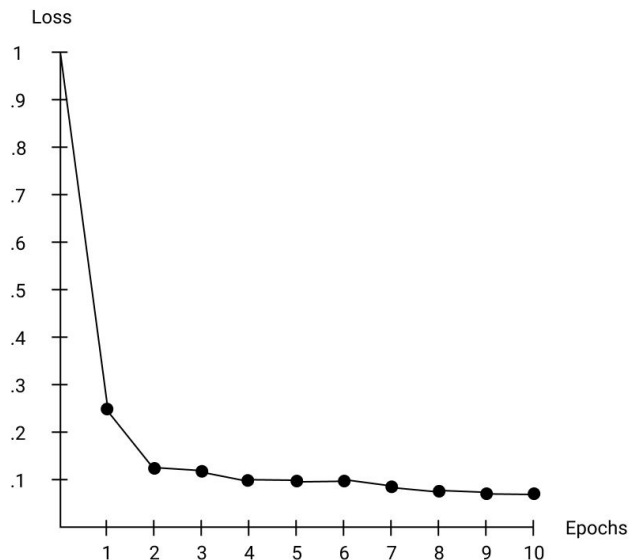
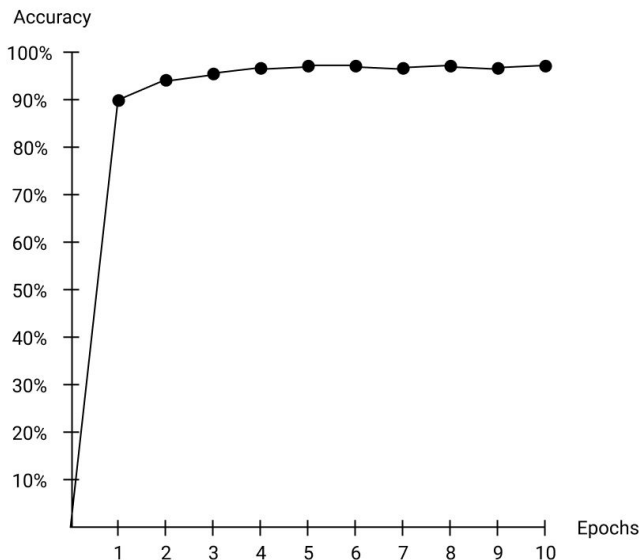
Архитектура модела

Модел има 3242 параметара и састоји се од:

- 1) 32 филтера величине 5x5. Активациона функција је РЕЛУ.
- 2) Слој за удруживање. Филтер величине 2x2 са померајем од 2 пиксела. Излаз је матрица димензија 14x14.
- 3) 32 филтера величине 5x5. Активациона функција је РЕЛУ.
- 4) Слој за удруживање са филтером величине 2x2 и померајем од 2 пиксела. Излаз је матрица величине 7x7.
- 5) Слој за поравнавање. Излаз је вектор са 1568 вредности.
- 6) Потпуно конектован слој са 10 неурона. Сваки неурон представља класу цифре.

Тренирање

Тренинг траје 10 епоха. За оптимизациону функцију користимо Спуштање Градијентом са стопом учења од 0.15. Прецизност модела је 97.6% над подацима за валидацију. Грешка је смањена на 0.08.



Примери грешака

Велики проценат цифара који су погрешно класификовани имају карактеристике више цифара, због чега модел није у стању да са сигурношћу одреди класу цифре.



Predicted: 3
True: 5



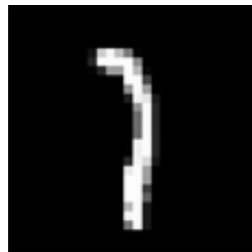
Predicted: 0
True: 6



Predicted: 9
True: 4



Predicted: 4
True: 9



Predicted: 1
True: 7

Шумови

Оптичка илузија за неуронске мреже. То се дешава јер модел не разуме апстрактне појмове као што су круг, линија, крива итд., већ долази до резултата преко низа математичких операција.



Class: 1



Class: 2



Class: 8

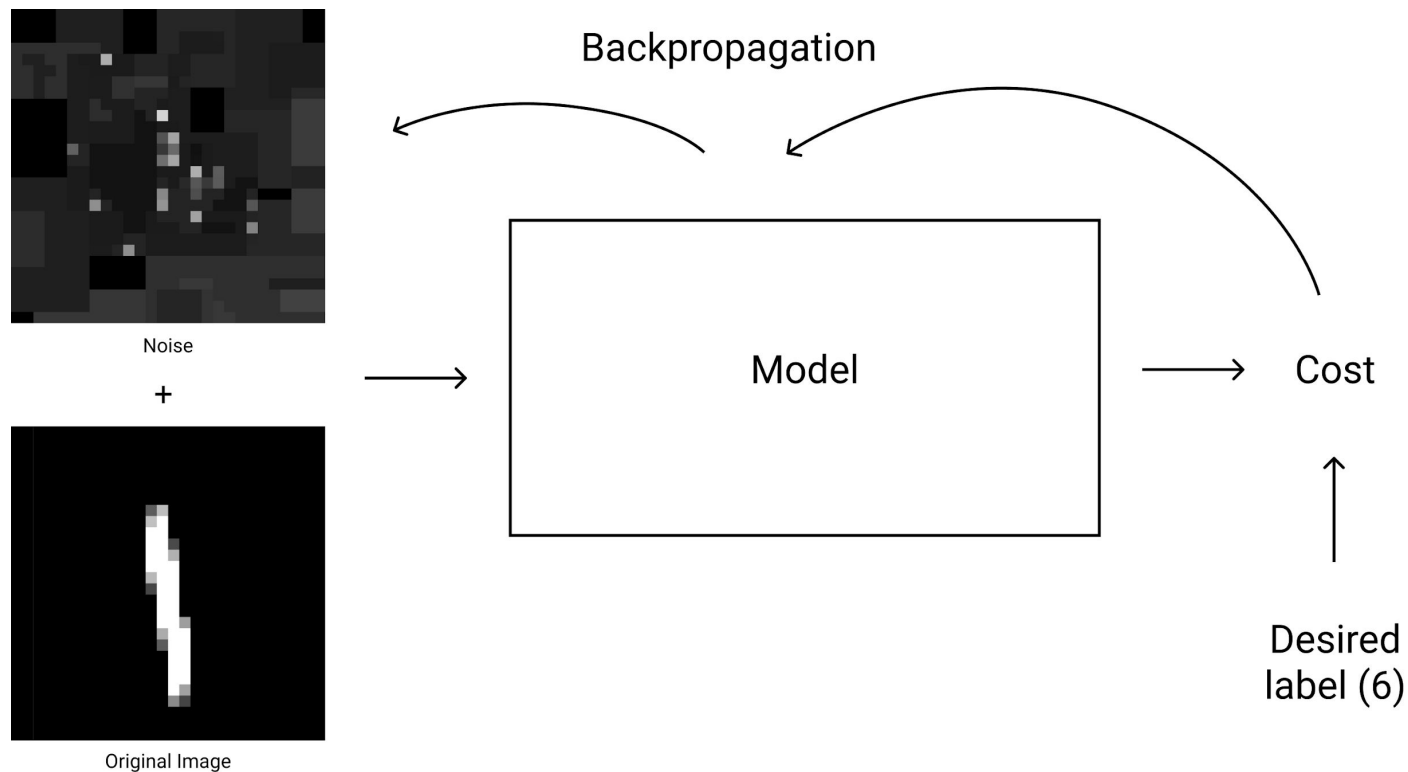


Class: 8



Class: 8

Прављење шумова



ЛеНет

Слојева: 8

Број параметара: 60,000

Прецизност: 97.9%

Слој	Филтери	Вел. слоја	Вел. филт.	Померај	Акт. фун.
Улаз	-	28x28	-	-	-
Конволу.	6	28x28	5x5	1	танх
Удруж.	6	14x14	2x2	2	танх
Конволу.	16	10x10	5x5	1	танх
Удружи.	16	5x5	2x2	2	танх
Конволу.	120	1x1	5x5	1	танх
Пот. конекнт.	-	84	-	-	танх
Излаз	-	10	-	-	софтмакс

АлексНет

Слојева: 13

Број параметара: 60,000,000

Прецизност: 97%

Слој	Филтери	Вел. слоја	Вел. филт.	Померај	Акт. функ.
Улаз	-	227x227	11x11	-	-
Конволу.	96	55x55	3x3	4	релу
Удруж.	96	27x27	5x5	2	релу
Конволу.	256	27x27	3x3	1	релу
Удруж.	256	13x13	3x3	2	релу
Конв. 2x	384	13x13	3x3	1	релу
Конволу.	256	13x13	3x3	1	релу
Удруж.	256	6x6	3x3	2	релу
Пот. кон.	-	9216	-	-	релу
Пот. Кон. 2x	-	4096	-	-	релу
Излаз	-	1000	-	-	софтмакс

ВГГ

Слојева: 22

Број параметара: 138,000,000

Прецизност: 98.4%

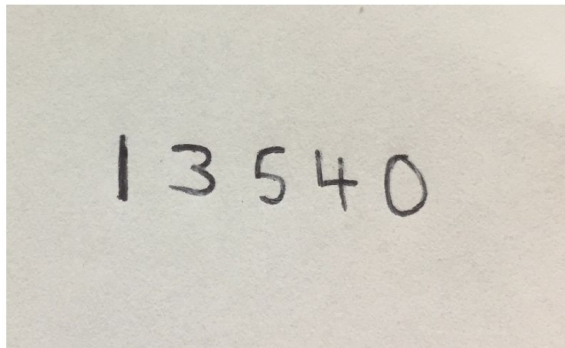
Слој	Филтери	Вел. слоја	Вел. филт.	Померај	Акт. функ.
Улаз	-	224x224	-	-	-
Конв. 2х	64	224x224	3х3	1	релу
Удруж.	64	112x112	3х3	2	релу
Конв. 2х	128	112x112	3х3	1	релу
Удруж.	128	56x56	3х3	2	релу
Конв. 2х	256	56x56	3х3	1	релу
Удруж.	256	28x28	3х3	2	релу
Конв. 3х	512	28x28	3х3	1	релу
Удруж.	512	14x14	3х3	2	релу
Конв. 3х	512	14x14	3х3	1	релу
Удруж.	512	7х7	3х3	2	релу
Пот. кон.	-	25088	-	-	релу
Пот. кон. 2х	-	4096	-	-	релу
Излаз	-	1000	-	-	софтмакс

Познате библиотеке

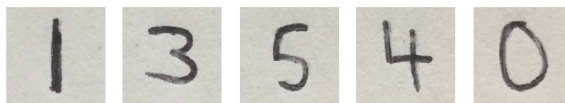
- Тенсорфлоу (енг. Tensorflow) - Направљен од стране Гугл-а за унутрашње потребе 2011 године. Изворни код је постао доступан 2015 године. Оптимизован за рад на графичким картицама.
- ПајТорч (енг. PyTorch) - Направљен од стране Фејзбук тима. Изворни код је постао доступан 2016 године. Оптимизован за рад на графичким картицама.

Коришћење у продукцији

Оригинална слика:



Издајање цифара:



Прерађивање слика:



Смањивање димензија слике:



Class: 1

Class: 3

Class: 5

Class: 4

Class: 0

Закључак

- Перформансе не зависе од величине модела и броја параметара, већ од архитектуре и квалитета података.
- Шумови могу представљати велики проблем ако модел није трениран да их препозна.
- Коришћењем метода обраде слика, модели се могу поједноставити.

Хвала на пажњи

Изворни код: <https://github.com/fioplazovic/filip.im>

Демо: <https://filip.im/demo>