

Podeanu Matei Alexandru, Schmidt Robert Eduard



Introducere

Convolutional Kernel Networks (CKNs) sunt un tip de rețele neuronale profunde care combină avantajele rețelelor neuronale convoluționale (CNNs) cu metode bazate pe kernel. Ele sunt concepute pentru a oferi o alternativă mai teoretic solidă la CNN-urile tradiționale, oferind un cadru bazat pe teoria kernelilor din învățarea automată.



Idei cheie ale implementarii CKN

- Kernel trick pentru reprezentarea patch-urilor
- Proiecția pe un subspațiu de dimensiune finită în RKHS
- Pooling liniar pentru obţinerea invariabilităţii la mici translaţii
- Construirea unei reprezentări ierarhice prin stivuirea de straturi





Dataset-uri



Set de date cu cifre scrise de mână.

1



SVHN

Numere stradale din imagini Google Street View.

CIFAR-10

Imagini color de 10 clase diferite.





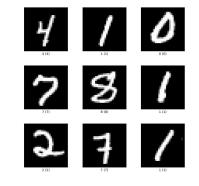
Fruits 360

Set de imagini cu fructe pentru clasificare.





MNIST







MNIST

Descriere

Cifre scrise de mână (0-9) utilizat pentru clasificarea imaginilor.

Training

60000 imagini de antrenare, 10000 imagini de testare

Rezolutie

28x28, alb-negru

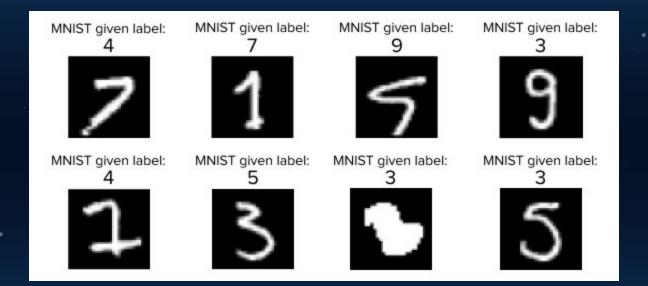
Clase

10, (cifrele 0-9)

Caracteristici

Set de date simplu, rezoluție redusă, ideal pentru testarea rețelelor neuronale de bază.

Este suficient?





SVHN





SVHN

Descriere

Un set de date cu numere de pe clădiri, preluate din imagini Google Street View.

Dimensiune

630.420 imagini

Rezolutie

32x32, RGB

Clase

10, (cifrele 0-9)

Caracteristici

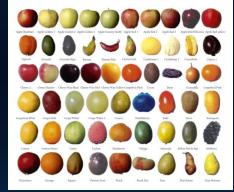
Mai complex decât MNIST, conținând cifre în diverse orientări, iluminări și fundaluri.

~83%

Pare cam mic. Incepem sa avem indoieli...

05

Fruits 360







Fruits 360

Descriere

Imagini color ale diferitelor tipuri de fructe și legume, utilizat pentru clasificarea imaginilor.

Classes

131

Dimensiune

90000 imagini

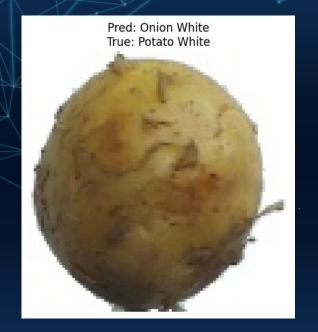
Rezolutie

100x100, RGB

Caracteristici

Imagini capturate într-un mediu controlat (fundal alb), ideal pentru antrenarea modelelor de recunoaștere a obiectelor

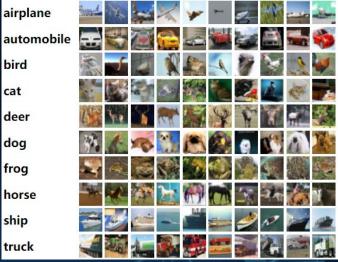
Si totusi, ce observam?



CLASS	CNN Acc(%)	CKN Acc(%)	Difference
Carambula	61.45	100.0	38.55
Corn Husk	55.19	66.23	11.04
Eggplant	62.82	67.95	5.13
Pear	63.36	71.95	8.59



CIFAR-10







CIFAR-10

Descriere

Un set de date cu imagini naturale, împărțit în 10 clase.

Dimensiune

50.000 imagini de antrenare, 10.000 imagini de testare

Rezolutie

32x32, RGB

Clase

10, (avion, automobil, pasăre, pisică, cerb, câine, broască, cal, navă, camion)

Caracteristici

Mai dificil decât MNIST datorită imaginilor color și diversității obiectelor.



Cum ar fi daca, in loc de numere si statistici, am putea vizualiza cum opereaza modelul? R: Ar fi <u>cam</u> bine...



Grad-CAM

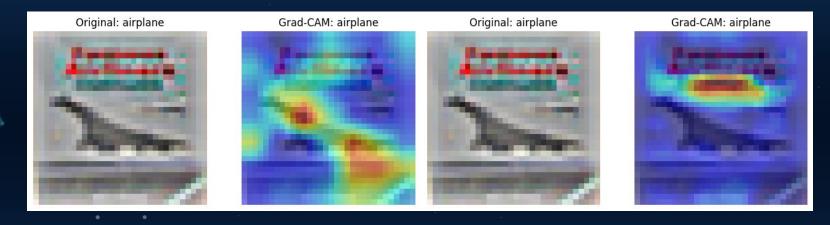
"a technique for producing 'visual explanations' for decisions from a large class of Convolutional Neural Network (CNN)-based models, making them more transparent and explainable."



Diferente

Accuracy for class 'airplane': 71.40%

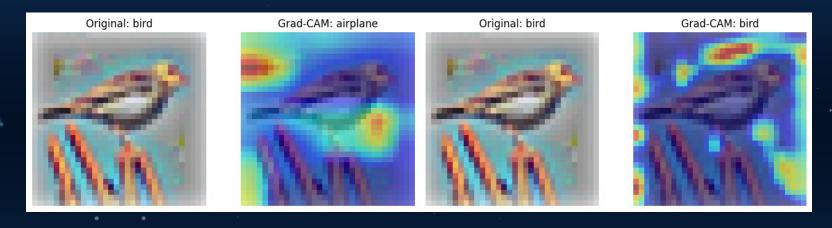
Accuracy for class 'airplane': 81.60%



Diferente

Accuracy for class 'bird': 49.30%

Accuracy for class 'bird': 65.80%





Referinte

- Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization
- Kernel Trick I Deep Convolutional Representations in RKHS Starting paper
- <u>End-to-End Kernel Learning with Supervised Convolutional Kernel Networks Original paper</u>
- Github Repository CKN