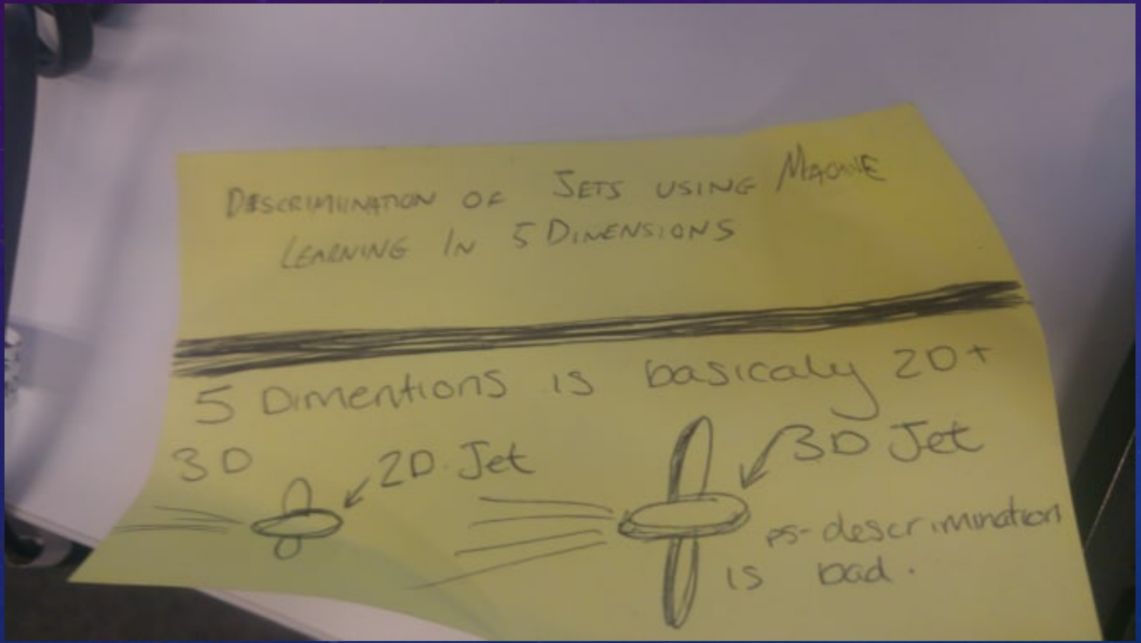
PRELIMINARY ADVENTURES IN NEURAL NETWORKS FOR PARTICLE PHYSICS

JACK ROLPH

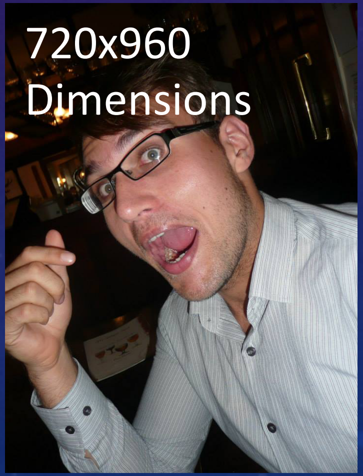


EXPLORING THE PROBLEM OF NEURAL NETS IN PARTICLE PHYSICS

- THERE ARE COMPLICATED NON-LINEAR PROBLEMS IN DETECTOR PHYSICS THAT COULD BE SOLVED USING NEURAL NETWORKS:
  - PARTICLE IDENTIFICATION
  - ENERGY RECONSTRUCTION

- FROM PHYSICS STANDPOINT, NEURAL NETWORK ACTS AS A FUNCTION THAT SOLVES A PARTICULAR HIGH-DIMENSIONAL PROBLEM BY TRIAL, ERROR AND IMPROVEMENT.

- EXAMPLE:



JACK, THE BRITISH POKEMON

Jack is the Pokemon in the picture (151 dimensions)

EXPLORING THE PROBLEM OF NEURAL NETS IN PARTICLE PHYSICS

- WHY DON'T WE DO THIS ALREADY? – NEW FIELD; MAIN BODY OF EXISTING RESEARCH NOT GEARED TOWARDS PARTICLE PHYSICS
- PROBLEMS TO SOLVE FOR NEURAL NETS IN PARTICLE PHYSICS:
  - **OFTEN TOO MANY DIMENSIONS FOR QUICK TRAINING**
  - TECHNIQUES THAT WORK WELL IN 2-DIMENSIONS MUST BE EXTENDED TO N-DIMENSIONS
  - **WHAT INPUT DATA SHOULD I USE? (HUMANS ARE BIASED)**
  - CAN I LEARN SOMETHING PHYSICAL FROM HOW MY NEURAL NET SOLVED THE PROBLEM

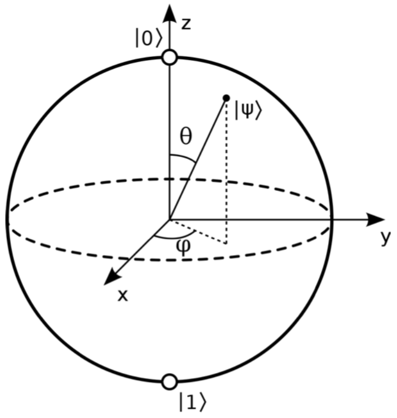
DIMENSIONALITY REDUCTION USING PRINCIPLE COMPONENT ANALYSIS (PCA)

- PCA VERY USEFUL TOOL IN DATA SCIENCE TO:
  - UNDERSTAND THE COVARIANCES BETWEEN THE VARIABLES OF A FIT
  - REDUCE THE DIMENSIONALITY OF MULTIVARIATE DATA

- ABSTRACT TOOL: QUITE DIFFICULT TO UNDERSTAND THE FIRST TIME AROUND!

EIGENDECOMPOSITION:

$$Av = \sum_{i=0}^{\dim(v)} \lambda_i v_i$$



BLOCH SPHERE:

- PSI CAN BE DECOMPOSED INTO BASIS STATES

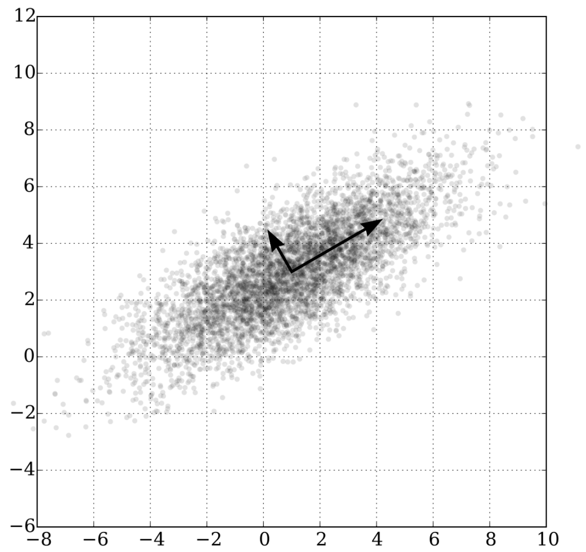
- LENGTH OF EIGENVALUE IS ‘IMPORTANCE’ OF BASIS VECTOR IN SPACE

DIMENSIONALITY REDUCTION USING PRINCIPLE COMPONENT ANALYSIS (PCA)

- EIGENDECOMPOSITION OF COVARIANCE (CORRELATION) MATRIX YIELDS 'DIRECTIONS OF MOST STATISTICAL VARIATION AND SIGNIFICANCE

EIGENDECOMPOSITION:

$$Av = \sum_{i=0}^{\dim(v)} \lambda_i v_i$$



BIVARIATE GAUSSIAN:

NEEDS A COVARIANCE MATRIX TO EXPLAIN

EIGENVECTORS -> DIRECTION OF VARIANCE

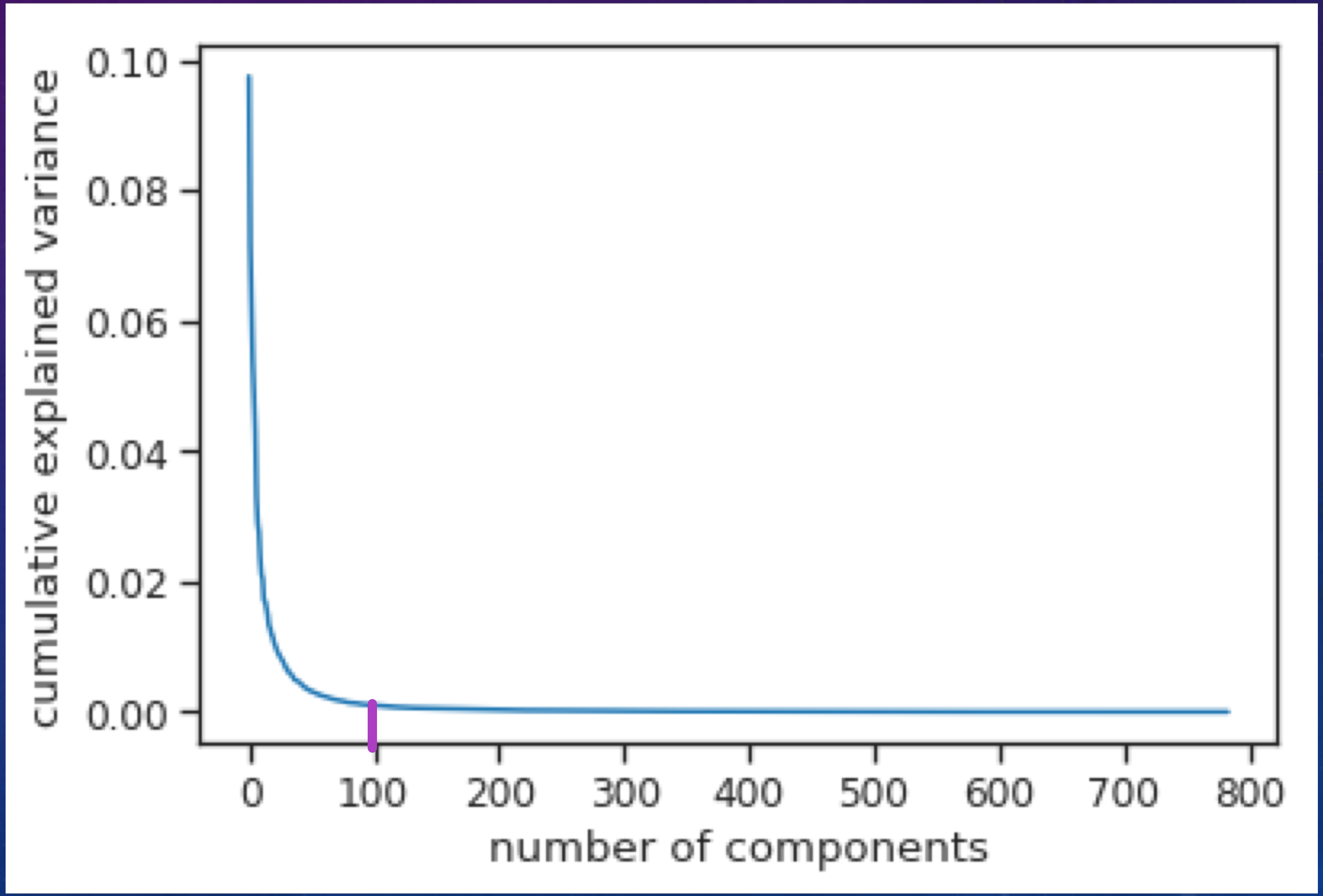
EIGENVALUES -> SIGNIFICANCE OF DIRECTION OF VARIANCE

BETTER TO PROJECT IN EIGENVECTORS

EXAMPLE OF PCA: MNIST DATASET

60,000 images of 28x28 pixel handwritten numbers (784 degrees of freedom)

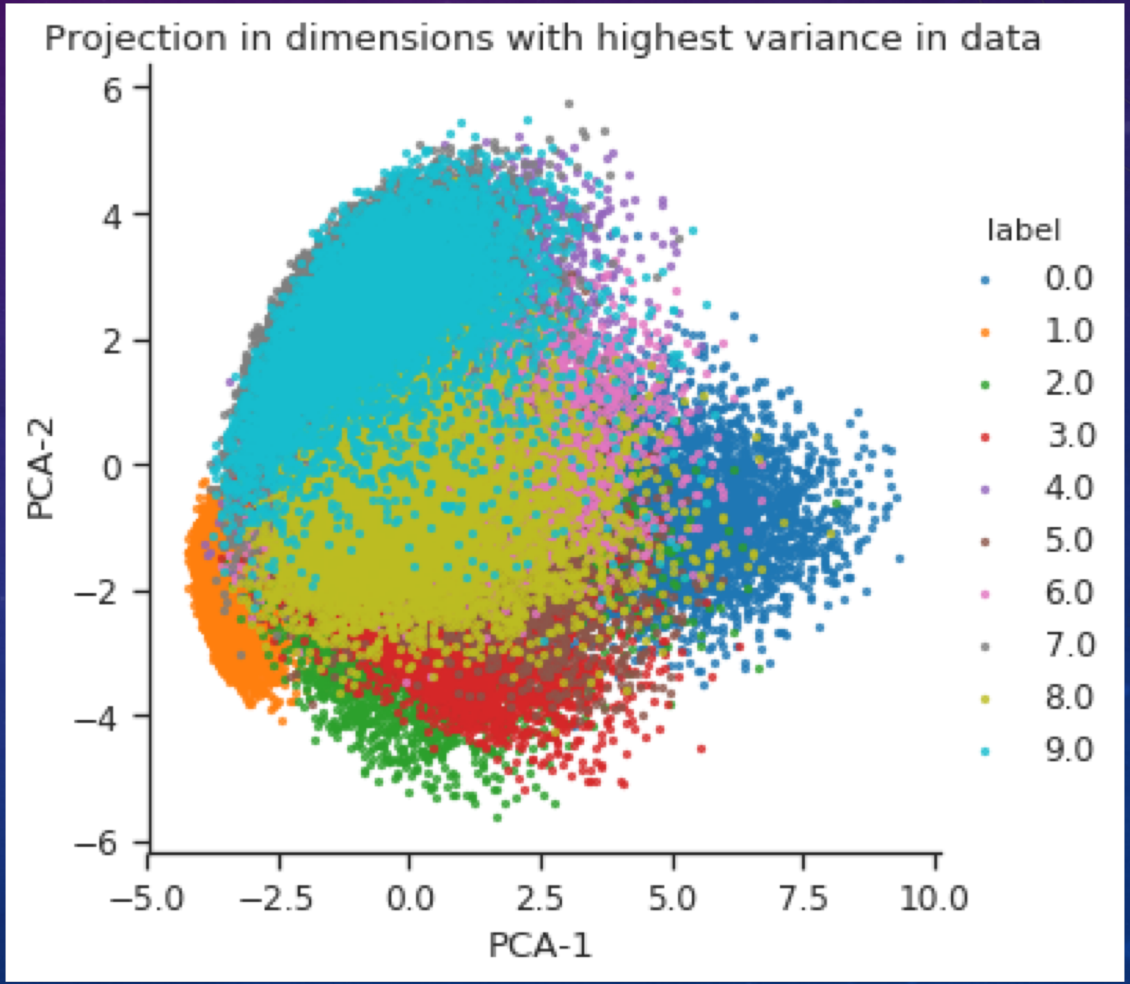
Explained variance
(MOST OF VARIANCE EXPLAINED BY 100 DEGREES OF FREEDOM)



EXAMPLE OF PCA: MNIST DATASET

60,000 images of 28x28 pixel handwritten numbers (784 degrees of freedom)

Plot in PCA-1 (15% variance) and PCA-2 (8% Variance)



DIMENSIONALITY REDUCTION USING PRINCIPLE COMPONENT ANALYSIS (PCA)

- WE HAVE **EVEN MORE** DIMENSIONS IN A ’CALORIMETER IMAGE’ I.E. N DEPTH, TIME
- PLAN:
  - CREATE BASIS VECTORS OF CALORIMETER IMAGES FROM MONTE CARLO SIMULATIONS OF ELECTRONS, MUONS, PIONS ETC
  - RUN THEM THROUGH A PCA TO OBTAIN A SET OF BASIS VECTORS, WHICH IS THEN FED TO NEURAL NET.
  - TYPICALLY THIS REDUCES PROCESSING TIME DRASTICALLY (30S -> 8S FOR MNIST NEURAL NET

- CANNOT CAPTURE NON-LINEARITY WITH PCA, SO FAIRLY USELESS IN JET DISCRIMINATION ALONE.
- USED TO REDUCE THE DIMENSIONS ONLY! NEURAL NETWORK HANDLES NON-LINEARITY

EXPLORING THE PROBLEM OF NEURAL NETS IN PARTICLE PHYSICS

- INITIAL PROBLEMS TO SOLVE FOR NEURAL NETS IN PARTICLE PHYSICS:
  - OFTEN TOO MANY DIMENSIONS FOR FEASIBLE TRAINING
  - **TECHNIQUES THAT WORK WELL IN 2-DIMENSIONS MUST BE EXTENDED TO HIGHER INPUT DIMENSIONS**
  - WHAT INPUT DATA SHOULD I USE? (HUMANS ARE BIASED)
  - **CAN I LEARN SOMETHING PHYSICAL FROM HOW MY NEURAL NET SOLVED THE PROBLEM**

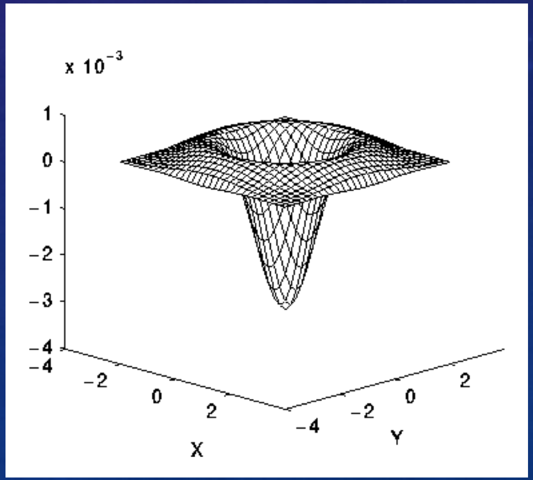
N-DIMENSIONAL CONVOLUTIONS IN CNNS

- CONVOLUTIONAL NEURAL NETWORKS VERY GOOD AT IMAGE RECOGNITION
- CNN 'LEARNS' BY FINDING PASSING A MATRIX KERNEL (FILTER) OVER A PICTURE TO HIGHLIGHT AN OPTIMAL FEATURE FOR A PARTICULAR PROBLEM

LAPLACIAN OF GAUSSIAN =
SMOOTHING (GAUS) + 2ND DERIVATIVE HIGHLIGHTING (LAPLACIAN) = EDGE FILTER

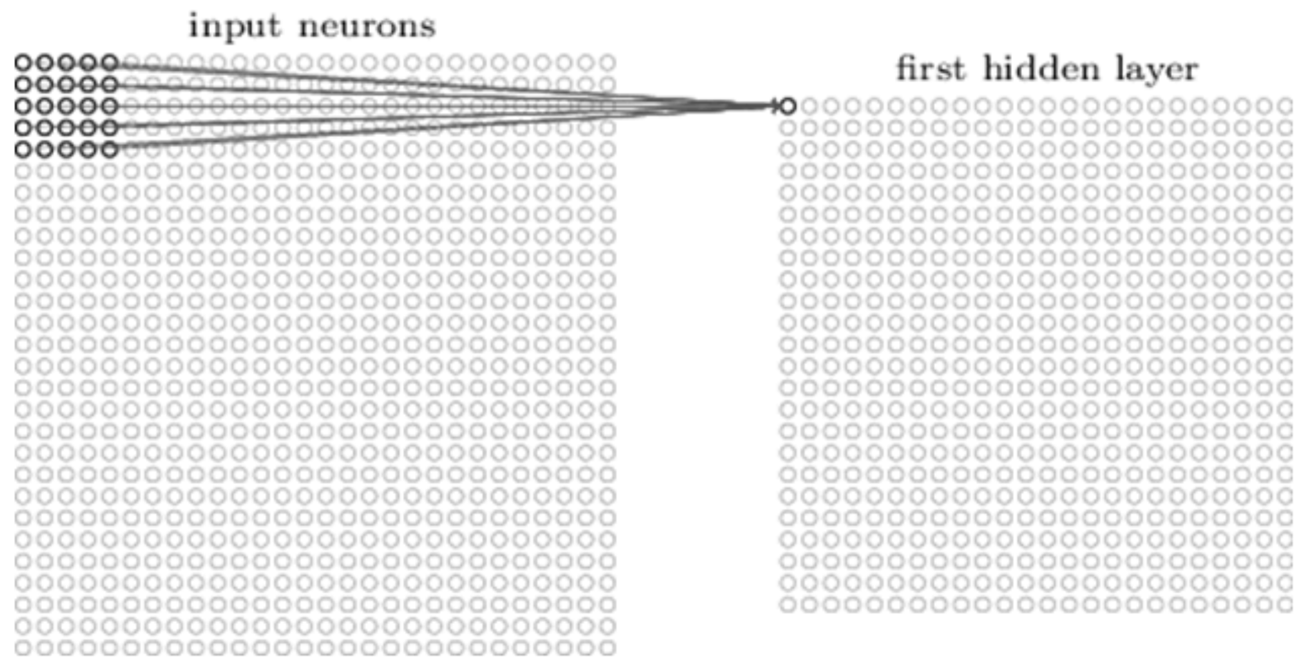
0	1	1	2	2	2	1	1	0
1	2	4	5	5	5	4	2	1
1	4	5	3	0	3	5	4	1
2	5	3	-12	-24	-12	3	5	2
2	5	0	-24	-40	-24	0	5	2
2	5	3	-12	-24	-12	3	5	2
1	4	5	3	0	3	5	4	1
1	2	4	5	5	5	4	2	1
0	1	1	2	2	2	1	1	0

=



N-DIMENSIONAL CONVOLUTIONS IN CNNS

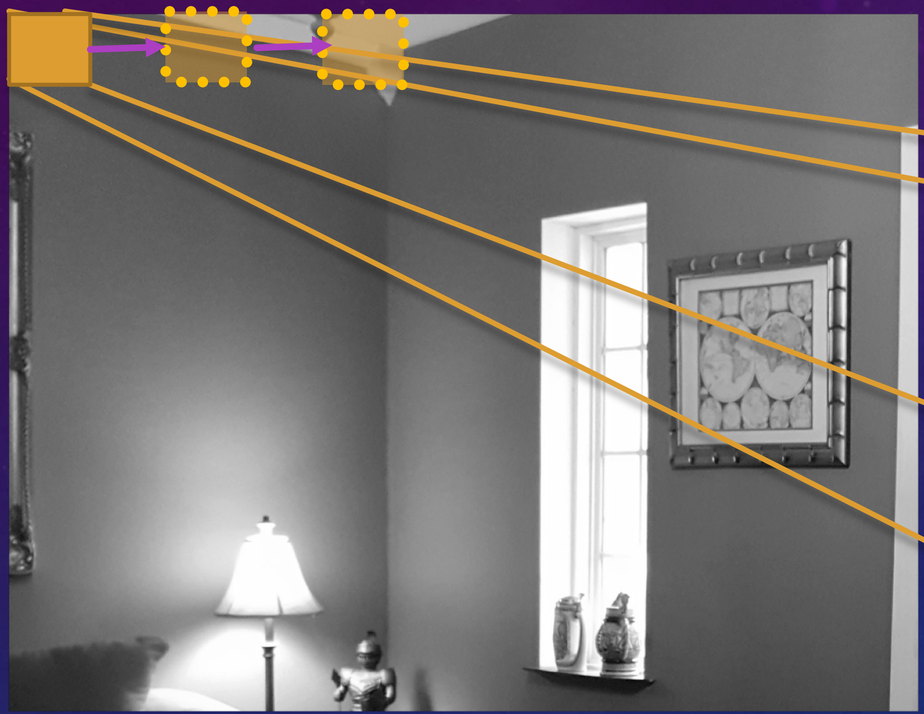
- CONVOLUTIONAL NEURAL NETWORKS VERY GOOD AT IMAGE RECOGNITION
- WITH HIGHLIGHTED FEATURES, THEN CONDENSED TO A SMALLER MATRIX WITH RELEVANT INFO (POOLING)



Visualization of 5 x 5 filter convolving around an input volume and producing an activation map

N-DIMENSIONAL CONVOLUTIONS IN CNNS

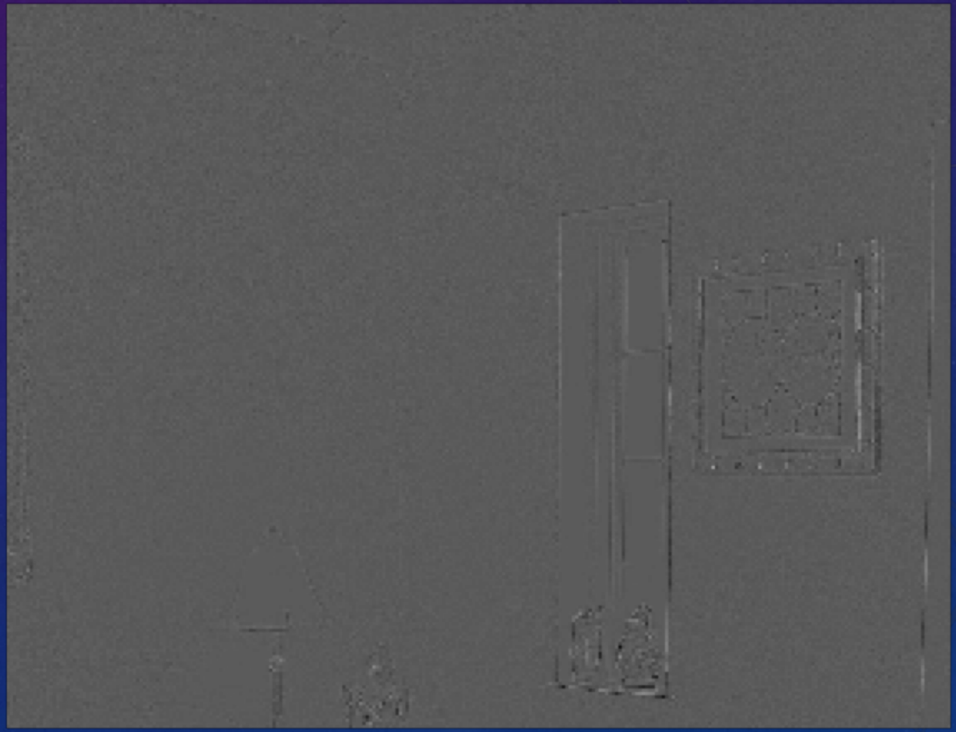
ORIGINAL IMAGE



X

0	1	1	2	2	2	1	1	0
1	2	4	5	5	5	4	2	1
1	4	5	3	0	3	5	4	1
2	5	3	-12	-24	-12	3	5	2
2	5	0	-24	-40	-24	0	5	2
2	5	3	-12	-24	-12	3	5	2
1	4	5	3	0	3	5	4	1
1	2	4	5	5	5	4	2	1
0	1	1	2	2	2	1	1	0

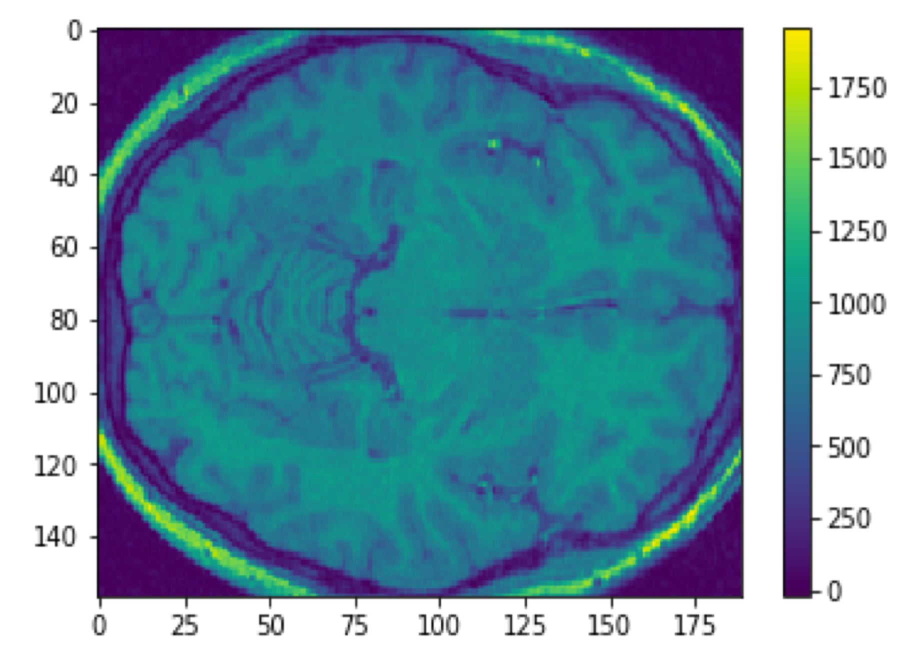
EDGE FILTER APPLIED



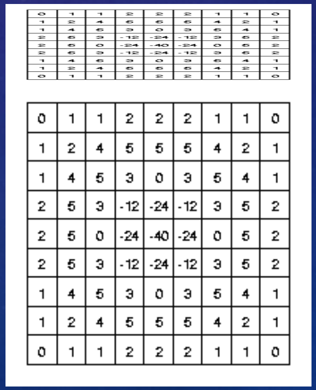
N-DIMENSIONAL CONVOLUTIONS IN CNNS

- PHYSICS ’PICTURES’ HAVE FAR MORE DIMENSIONS THAN MOST APPLICATIONS AT PRESENT (COMPUTER VISION ONLY NEEDS 2, KERAS GOES UP TO 4 DIMENSIONS)
- IN ORDER TO HIGHLIGHT FEATURES IN N-DIMENSIONS, WE (PROBABLY) NEED N-DIMENSIONAL KERNELS
- EXTENDED THIS TO N-DIMENSIONS BY SIMPLY EXTENDING THE PROCESS USED TO CONVOLVE FILTERS

ORIGINAL IMAGE

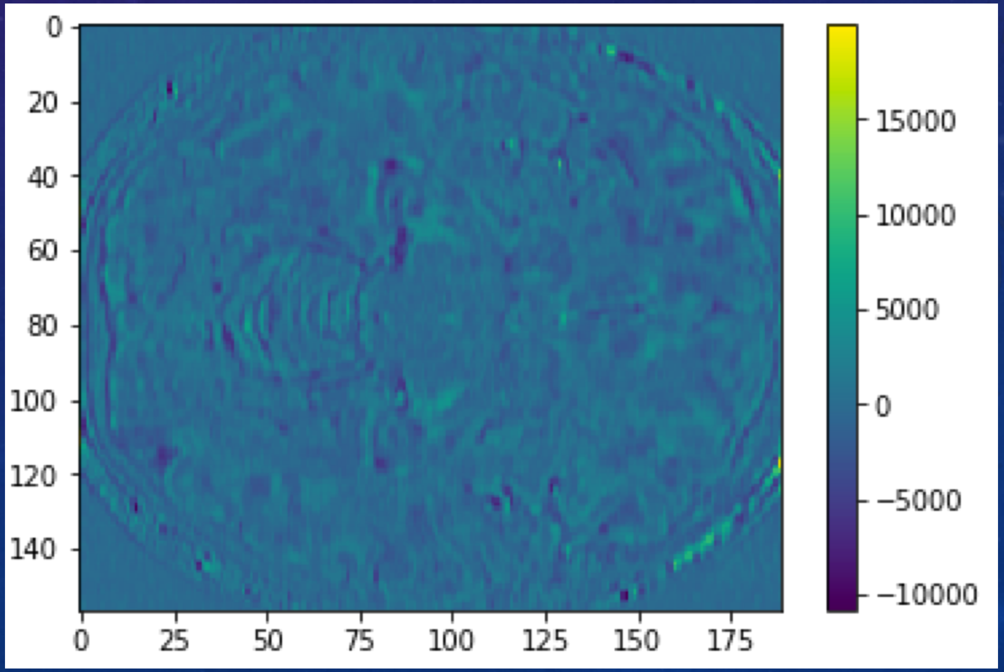


X



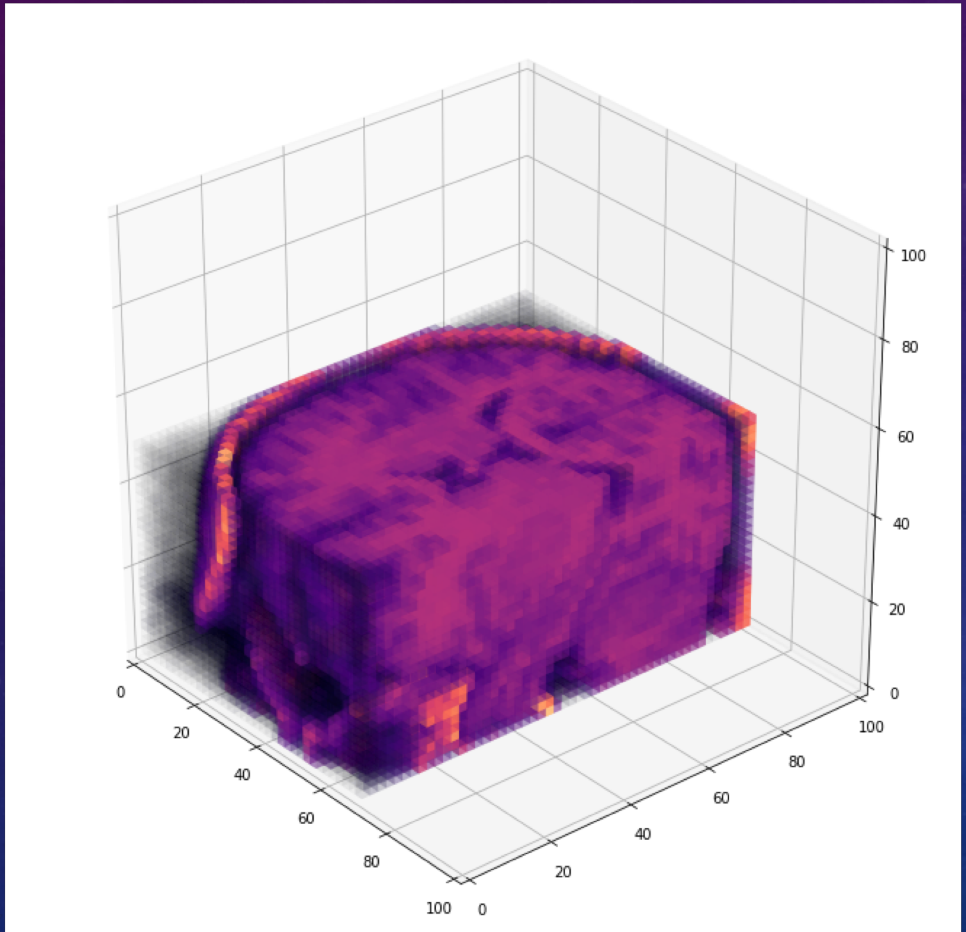
0	1	1	2	2	2	1	1	0
1	2	4	5	5	5	4	2	1
1	4	5	3	0	3	5	4	1
2	5	3	-12	-24	-12	3	5	2
2	5	0	-24	-40	-24	0	5	2
2	5	3	-12	-24	-12	3	5	2
1	4	5	3	0	3	5	4	1
1	2	4	5	5	5	4	2	1
0	1	1	2	2	2	1	1	0

EDGE FILTER APPLIED (SOBEL X IN 3D)



N-DIMENSIONAL CONVOLUTIONS IN CNNS

ORIGINAL IMAGE



EDGE FILTER APPLIED (SOBEL X IN 3D)