Object Detection, CNN based Networks Ganesh Ramakrishnan, Dept of CSE, IITB

Acknowledgement: Anurag Mundhada, Arjun Jain

Practically:

The task of assigning a label and a bounding box to all objects in an image





Data: The oil driving object detection research

Object Detection datasets



IM GENET

200,000 images and 80 object categories

Detection: 500,000 images and 200 object categories

Google Open Images v4

15,440,132 boxes on 600 categories

Computer Vision Tasks



CS231 Lecture Notes, Stanford Liningr, Sitty eptive Code LLC

Classification



Classification + Localization



Classification + Localization



IIT Bombay, Perceptive Code LLC

Classification + Localization

How would you do it for multiple objects?



IIT Bombay, Perceptive Code LLC

Correct Class: Dog

Classification + bounding box regression for **multiple boxes**



Brute force approach - run a classifier for **every possible box**



Brute force approach - run a classifier for **every possible box** with a **sliding window**



Histograms of Oriented Gradients for Human Detection Dalal and Triggs 2005

Brute force approach - run a classifier for **every possible box** with a **sliding window at every scale on an image pyramid**





Histograms of Oriented Gradients for Human Detection Dalal and Triggs 2005

Brute force approach - run a classifier for **every possible box** with a **sliding window at every scale on an image pyramid**





Histograms of Oriented Gradients for Human Detection Dalal and Triggs 2005

Very slow, unfeasible if your classifier is heavy

Smarter approach?

Smarter approach - run classifier for region proposals or boxes likely to contain objects

Smarter approach - run classifier for region proposals or boxes likely to contain objects

• How to reduce number of boxes?

Smarter approach - run classifier for region proposals or boxes likely to contain objects

- How to reduce number of boxes?
- Find 'blobby' image regions which are likely to contain objects

Smarter approach - run classifier for region proposals or boxes likely to contain objects

- How to reduce number of boxes?
- Find 'blobby' image regions which are likely to contain objects
- Class-agnostic object detector "Region Proposals"

Region Proposals

Greedily combine sub-segmentation to produce larger candidate object locations



Selective Search for Object Recognition, Uijlings et al (2013)

Region Proposals

Greedily combine sub-segmentation to produce larger candidate object locations



Pre deep learning



Post deep learning



81.3%

R-CNN

Rich feature hierarchies for accurate object detection and semantic segmentation, Girshick et al (2013)

1. Input image 2. Extract region proposals (~2k) 3. Compute CNN features 4. Classify regions SVM classifier

Exactly the same as Selective Search, except for CNN features instead of HOG https://arxiv.org/abs/1311.2524

R-CNN: Regions with CNN features

R-CNN: Problems

- Ad hoc training objectives
 - Fine-tune network with softmax classifier (log loss
 - Train post-hoc linear SVMs (hinge loss)
 - Train post-hoc bounding-box regressions (least squares)
- Training is slow (84h), takes a lot of disk space
- Inference (detection) is slow
 - 47s / image with VGG16 [Simonyan & Zisserman. ICLR15]
 - Fixed by SPP-net [He et al. ECCV14]



Fast R-CNN: Region Proposals in Feature Space



Fast R-CNN: ROI Pooling

input									
0.88	0.44	0.14	0.16	0.37	0.77	0.96	0.27		
0.19	0.45	0.57	0.16	0.63	0.29	0.71	0.70		
0.66	0.26	0.82	0.64	0.54	0.73	0.59	0.26		
0.85	0.34	0.76	0.84	0.29	0.75	0.62	0.25		
0.32	0.74	0.21	0.39	0.34	0.03	0.33	0.48		
0.20	0.14	0.16	0.13	0.73	0.65	0.96	0.32		
0.19	0.69	0.09	0.86	0.88	0.07	0.01	0.48		
0.83	0.24	0.97	0.04	0.24	0.35	0.50	0.91		

https://deepsense.ai/region-of-interest-pooling-explained/

IIT Bombay, Perceptive Code LLC

Fast R-CNN: ROI Pooling

input									
0.88	0.44	0.14	0.16	0.37	0.77	0.96	0.27		
0.19	0.45	0.57	0.16	0.63	0.29	0.71	0.70		
0.66	0.26	0.82	0.64	0.54	0.73	0.59	0.26		
0.85	0.34	0.76	0.84	0.29	0.75	0.62	0.25		
0.32	0.74	0.21	0.39	0.34	0.03	0.33	0.48		
0.20	0.14	0.16	0.13	0.73	0.65	0.96	0.32		
0.19	0.69	0.09	0.86	0.88	0.07	0.01	0.48		
0.83	0.24	0.97	0.04	0.24	0.35	0.50	0.91		

region proposal								
0.88	0.44	0.14	0.16	0.37	0.77	0.96	0.27	
0.19	0.45	0.57	0.16	0.63	0.29	0.71	0.70	
0.66	0.26	0.82	0.64	0.54	0.73	0.59	0.26	
0.85	0.34	0.76	0.84	0.29	0.75	0.62	0.25	
0.32	0.74	0.21	0.39	0.34	0.03	0.33	0.48	
0.20	0.14	0.16	0.13	0.73	0.65	0.96	0.32	
0.19	0.69	0.09	0.86	0.88	0.07	0.01	0.48	
0.83	0.24	0.97	0.04	0.24	0.35	0.50	0.91	



pooling sections 0.44 0.14 0.16 0.37 0.77 0.96 0.88 0.27 0.45 0.57 0.19 0.16 0.63 0.29 0.71 0.70 0.66 0.26 0.82 0.64 0.54 0.73 0.59 0.26 0.29 0.62 0.85 0.34 0.76 0.84 0.75 0.25 0.74 0.39 0.34 0.33 0.32 0.21 0.48 0.20 0.14 0.16 0.13 0.73 0.65 0.96 0.32 0.09 0.86 0.88 0.07 0.01 0.19 0.69 0.48 0.24 0.83 0.24 0.97 0.04 0.35 0.50 0.91

https://deepsense.ai/region-of-interest-pooling-explained/

IIT Bombay, Perceptive Code LLC



Non Maximum Suppression

Remove boxes with high overlap but lower score



- Discard boxes with softmax_output<confidence_threshold
- Sort boxes in descending order of confidence
- For box in sorted(box_list)
 - Discard all boxes with IoU>0.5 with box

Non-Max Suppression

Before non-max suppression



After non-max suppression



Region Proposals



Region Proposals

Maximum time still taken by Region proposal generation



Faster R-CNN: Use a CNN for region proposals



Region Proposals using CNN

At each uniformly sampled grid cell, set of anchor boxes with different aspect ratios and scales



For anchors, we use 3 scales with box areas of 1282, 2562, and 5122 pixels, and 3 aspect ratios of 1:1, 1:2, and 2:1.

For each anchor box

- Regress to a final box with 4 numbers: (*dx, dy, dh, dw*)
- Predict 2 objectness scores (including background as a separate class)

https://arxiv.org/abs/1506.01497





Can this be further simplified?



Compress into single stage -Remove region-based computation
Single Stage Object Detection

YOLO: You Only Look Once & SSD: Single Shot Multibox Detector

Directly estimate box coordinates and class scores



IIT Bombay, Perceptive Code LLC



- Split the image into a grid
- Each cell predicts boxes and confidences
 P(Object)



- Split the image into a grid
- Each cell predicts boxes and confidences
 P(Object)



- Split the image into a grid
- Each cell predicts boxes and confidences
 P(Object)



Bicycle

- Split the image into a grid
- Each cell predicts boxes and confidences
 P(Object)
- Each cell also predicts a class probability

Conditioned on object: P(Car | Object)

Dog

https://arxiv.org/abs/1506.02640



Car

- Split the image into a grid
- Each cell predicts boxes and confidences
 P(Object)
- Each cell also predicts a class probability
- Then box and class predictions are combined



- Split the image into a grid
- Each cell predicts boxes and confidences
 P(Object)
- Each cell also predicts a class probability
- Then box and class predictions are combined
- Followed by NMS



YOLO: You Only Look Once output space

Each cell predicts:

- For each bounding box:
- 4 coordinates (x, y, w, h)
- 1 confidence value
- Some number of class probabilities

For Pascal VOC:

- 7x7 grid

- 2 bounding boxes / cell
- 20 classes
- 7 x 7 x (2 x 5 + 20) = 7 x 7 x 30 tensor = **1470 outputs**



SSD: Single Shot Multibox Detector

Like YOLO, but predicts using a multi-scale pyramidal feature hierarchy



- IoU
- Precision and Recall
- Mean Average Precision (mAP)

IoU (intersection over union): Measure of box similarity



Precision: what percentage of your positive predictions are correct

Recall: what percentage of ground truth objects were found



Mean Average Precision (mAP) for object detection

Step 1: Sort predictions according to confidence (usually classifier's output after softmax)

Step 2: Calculate IoU of every predicted box with every ground truth box

Step 3: Match predictions to ground truth using IoU, correct predictions are those with IoU > threshold (typically 0.5)

	_	
Confidence	Rank	Correct
0.91	1	TRUE
0.87	2	TRUE
0.83	3	FALSE
0.81	4	TRUE
0.77	5	FALSE
0.65	6	TRUE
0.56	7	TRUE
0.40	8	FALSE
0.32	9	FALSE
0.31	10	TRUE

Mean Average Precision (mAP) for object detection

Total number of ground truth boxes = 6

Step 4: Calculate precision and recall at every row

Step 5: Take the mean of maximum precision at 11 recall values {0.0, 0.1, ... 1.0) to get AP

Step 6: Average across all classes to get the mAP score

Confidence	Rank	Correct	Precision	Recall
0.91	1	TRUE	1.00	0.17
0.87	2	TRUE	1.00	0.33
0.83	3	FALSE	0.67	0.33
0.81	4	TRUE	0.75	0.50
0.77	5	FALSE	0.60	0.50
0.65	6	TRUE	0.67	0.67
0.56	7	TRUE	0.71	0.83
0.40	8	FALSE	0.63	0.83
0.32	9	FALSE	0.56	0.83
0.31	10	TRUE	0.67	1.00

Precision and Recall curve

Precision vs Recall



Recall

Precision and Recall curve



Anchor free box prediction

- Directly predict (dx, dy, w, h) at each grid cell
 - \circ \rightarrow Like YOLO (which does this 2x)
- Center of grid cell is center of bbox
- Harder problem, as output space has higher variance



Anchored box prediction

At each uniformly sampled grid cell, set of anchor boxes with different aspect ratios and scales



Faster R-CNN: For anchors, we use 3 scales with box areas of 128, 256, and 512 pixels, and 3 aspect ratios of 1:1, 1:2, and 2:1.

For each anchor box

- Regress to a final box with 4 numbers: (*dx, dy, dw, dh*)
- Predict 2 objectness scores (including background as a separate class)

https://arxiv.org/abs/1506.01497 IIT Bombay, Perceptive Code LLC

SSD: Single Shot Multibox Detector

Like YOLO, but predicts using a multi-scale pyramidal feature hierarchy



Why still use a two-stage object detector?

- Better recall of RPN as compared to SSD/YOLO
 - Trained with all object instances
 - Generic first stage, usable for multitask

- Finer control over training classifier
 - Custom minibatch (sampling 3:1 negative samples)

• Instance-level multitask (Mask-RCNN)

Mask R-CNN – towards instance-level understanding



https://arxiv.org/abs/1703.06870

IIT Bombay, Perceptive Code LLC

Mask R-CNN – towards instance-level understanding



Zoom in on instances

Mask R-CNN

Preserves pixel-to-pixel



Fast R-CNN: ROI Pooling

input								
0.88	0.44	0.14	0.16	0.37	0.77	0.96	0.27	
0.19	0.45	0.57	0.16	0.63	0.29	0.71	0.70	
0.66	0.26	0.82	0.64	0.54	0.73	0.59	0.26	
0.85	0.34	0.76	0.84	0.29	0.75	0.62	0.25	
0.32	0.74	0.21	0.39	0.34	0.03	0.33	0.48	
0.20	0.14	0.16	0.13	0.73	0.65	0.96	0.32	
0.19	0.69	0.09	0.86	0.88	0.07	0.01	0.48	
0.83	0.24	0.97	0,04	0.24	0.35	0.50	0.91	

		re	egion	propos	sal		
0.88	0.44	0.14	0.16	0.37	0.77	0.96	0.27
0.19	0.45	0.57	0.16	0.63	0.29	0.71	0.70
0.50	0.20	0.82	0.64	0.54	0.75	0.39	26
0. 35	0.34	0.76	0.84	0.29	0.75	0.62	0 25
0.32	0.74	0.21	0.39	0.34	0.03	0.33	0 48
0.20	0.14	0.16	0.13	0.73	0.65	0.96	0 32
0.19	0.69	0.09	0.86	0.88	0.07	0.01	0 48
0.83	0.24	0.97	0.04	0.24	0.35	0.50	0 91



pooling sections 0.88 0.44 0.14 0.16 0.37 0.77 0.96 0.27 0.19 0.45 0.57 0.16 0.63 0.29 0.71 0.70 0.54 0.73 0.66 0.26 0.82 0.64 0.59 0.26 0.34 0.29 0.75 0.62 0.76 0.84 0.85 0.25 0.32 0.74 0.21 0.34 0.03 0.33 0.48 0.32 0.73 0.05 0.90 0.10 0.69 0.09 0.86 0.88 0.48 0.19 0.07 0.01 0.97 0.24 0.35 0.50 0.91 0.83 0.24 0.04

Double Quantization – loss of pixel-topixel alignment

https://deepsense.ai/region-of-interestappelingrexplained/

ROI Align

Improvement on ROI Pooling



- Input: Feature map (5x5 here) and region proposal (normalized float coordinates)
- Output: 2x2 'pooled' bins
- Sample 4 points in every bin uniformly
- Compute value at each bin using bilinear interpolation
- Max or average the 4 bins

Effective Receptive Field Calculation

- Let $R_{k,j}$ be the effective receptive field of a neuron in layer k projected on layer j
- *f_j*, *s_j* respectively be filter size and stride of layer *j*

Writing $R_{k,j}$ in terms of $R_{k,j+1}$:

$$R_{k,j} = (R_{k,j+1} - 1) s_{j+1} + f_{j+1}$$

Finally, $R_{k,k} = 1$



https://arxiv.org/pdf/1705.07049.pdf

Classification

Retrieval



[Krizhevsky 2012]

Detection



Segmentation



Detection



Instance Segmentation



[Taigman et al. 2014]



input video

[Goodfellow 2014]

IIT Bombay, Perceptive Code LLC

[Simonyan et al. 2014]

stride 1

stride 1

stride 2

pool 2x2

stride 1

pool 2x2

dropout

dropout

stride 2

norm.

pool 2x2

multi-frame

optical flow



[Ciresan et al. 2013]

α_{0}	攤	哲	蟣	赃	赃	葬	灗
择	ĸJ	泽	贼	点	憎	樒	窗
诉	摘	寄	宅	冤	债	雾	瞻
湛	娖	樟	章	彰	漳	驮	李
SE M	罩	K	摩	B	遞	折	扔
针	棙	枕	疠	咨	震	뜞	镇
郛	₹£	艺	枝	支	吱	购	名の
Ł	趾	Ŷ	NG6	狨	まの	势	挪



[Sermanet et al. 2011] [Ciresan et al.]





NVIDIA Tegra X1

self-driving cars



[Turaga et al., 2010]



I caught this movie on the Sci-Fi channel recently. It actually turned out to be pretty decent as far as B-list horror/suspense films go. [We guys (one naive and one loud mouthed a**) take a road trip to stop a welding but have the worst possible lack when a maniac in a freaky, make-shift tank/truck hybrid decides to play cat-and-mouse with them. Things are further complicated when they pick up a ridiculously whorish hitchhiker. What makes this film unique is that the combination of comedy and terore actually work in this movie, unlike so many others. The two guys are likable enough and there are some good chase/suspense scenes. Nice pacing and comic timing make this movie more than passable for the horror/slasher buff. Definitely worth checking out.

I just saw this on a local independent station in the New York City area. The cast showed promise but when I saw the director, George Cosmotos, I became suspicious. And sure enough, it was every bit as bad, every bit as pointless and stupid as every George Cosmotos movie I ever saw. He's like a stupid man's Michael Bey – with all the avfulness that accolade promises. There's no point to the conspiracy, no burning issues that urge the conspirators on. We are left to ourselves to connect the dots from one bit of graffiti on various walls in the film to the next. Thus, the current budget crisis, the war in Iraq. Islamic extremism, the fate of social security, 47 million Americans without health care, stagnating wages, and the death of the middle class are all subsumed by the sheer terror of graffiti. A truly, stumingly idiotic film.

Graphics is far from the best part of the game. **This is the number one best TH game in the series**. Next to Underground. If deserves strong love, It is an instant game. There are massive levels, massive unlockable characters... it's just a massive game. **Maste your money on this game**. This is the kind of money that is waited property. And even though graphics suck, thats doesn't make a game good. Actually, the graphics were good at the time. Today the graphics are crap. WHO CARES? As they say in Canada, This is the fung game, aye. (You get to go to Canada in THPS3) Well, I don't know if they say that, but they might, who knows. Well Canadian people do. Wait a minute, I'm getting off topic. This game crocks up it, play it, play it, play it, play it, play it, play It.REE BRILLANCE.

The first was good and original. I was a not bad horror/comedy movie. So I heard a second one was made and I had to watch it. What really makes this movie work is Judd Nelson's character and the sometimes clever script. A pretty good script for a person who wrote the Final Destination films and the direction was okay. Sometimes there's scenes where it looks like it was filmed using a hone video camera with a grainy - look. Great made - for - TV movie. If was worth the rental and probably worth buying just to get that nice ceric feeling and watch Judd Nelson's Stanley doing what he does best. I suggest newcomers to watch the first one before watching the sequel, just so you'll have an idea what Stanley is like and get a little history background.

[Denil et al. 2014]


Whale recognition, Kaggle Challenge



Mnih and Hinton, 2010



Juggling Sequence Very Complex motion (30 FPS)



Elhayek et al., 2016

David Eigen and Rob Fergus, 2015





Molchanov et al., 2016

Kendall et al., 2016

Reinforcement Learning



Automatic Speech Recognition



AlphaGo, 2016 (Go Board treated as an image)

Wavenet, 2016 (Dilated 1D Convolutions)

Object Tracking



Optical Flow



FlowNet, 2015



(a) Predicted 2D translations



(b) Generated vote-map



(c) Final extracted points

[Doiphode et al.]

Case Study Bonus: DeepMind's AlphaGo



The input to the policy network is a $19 \times 19 \times 48$ image stack consisting of 48 feature planes. The first hidden layer zero pads the input into a 23×23 image, then convolves *k* filters of kernel size 5 × 5 with stride 1 with the input image and applies a rectifier nonlinearity. Each of the subsequent hidden layers 2 to 12 zero pads the respective previous hidden layer into a 21×21 image, then convolves *k* filters of kernel size 3×3 with stride 1, again followed by a rectifier nonlinearity. The final layer convolves 1 filter of kernel size 1×1 with stride 1, with a different bias for each position, and applies a softmax function. The match version of AlphaGo used *k* = 192 filters; Fig. 2b and Extended Data Table 3 additionally show the results of training with *k* = 128, 256 and 384 filters.

policy network:

[19x19x48] Input CONV1: 192 5x5 filters , stride 1, pad 2 => [19x19x192] CONV2..12: 192 3x3 filters, stride 1, pad 1 => [19x19x192] CONV: 1 1x1 filter, stride 1, pad 0 => [19x19] (probability map of promising moves)



reddit.com/r/deepdream

TLDR



Follow

 \sim

 \sim

 \sim

Flying to #<u>CVPR17</u> later tonight! ConvNets ConvNets

11:58 AM - 21 Jul 2017



A Few CNN Case Studies

Case Study: LeNet-5

[LeCun et al., 1998]



Conv filters were 5x5, applied at stride 1 Subsampling (Pooling) layers were 2x2 applied at stride 2 i.e. architecture is [CONV-POOL-CONV-POOL-CONV-FC]





First layer (CONV1): 96 11x11 filters applied at stride 4 =>

Q: what is the output volume size?







First layer (CONV1): 96 11x11 filters applied at stride 4
=>
Q: what is the output volume size? (227-11)/4+1 = 55





First layer (CONV1): 96 11x11 filters applied at stride 4 =>

Output volume [55x55x96]

Q: What is the total number of parameters in this layer?





First layer (CONV1): 96 11x11 filters applied at stride 4 =>

Output volume **[55x55x96]** Parameters: (11*11*3)*96 = **35K**





Input: 227x227x3 images After CONV1: 55x55x96

Second layer (POOL1): 3x3 filters applied at stride 2

Q: what is the output volume size? Hint: (55-3)/2+1 = 27





Input: 227x227x3 images After CONV1: 55x55x96

Second layer (POOL1): 3x3 filters applied at stride 2 Output volume: 27x27x96

Q: what is the number of parameters in this layer?





Input: 227x227x3 images After CONV1: 55x55x96

Second layer (POOL1): 3x3 filters applied at stride 2 Output volume: 27x27x96 Parameters: 0!





Input: 227x227x3 images After CONV1: 55x55x96 After POOL1: 27x27x96

• • •

Case Study: AlexNet

[Krizhevsky et al. 2012]

Full (simplified) AlexNet architecture: [227x227x3] INPUT [55x55x96] CONV1: 96 11x11 filters at stride 4, pad 0 [27x27x96] MAX POOL1: 3x3 filters at stride 2 [27x27x96] NORM1: Normalization layer [27x27x256] CONV2: 256 5x5 filters at stride 1, pad 2 [13x13x256] MAX POOL2: 3x3 filters at stride 2 [13x13x256] NORM2: Normalization layer [13x13x384] CONV3: 384 3x3 filters at stride 1, pad 1 [13x13x384] CONV4: 384 3x3 filters at stride 1, pad 1 [13x13x256] CONV5: 256 3x3 filters at stride 1, pad 1 [6x6x256] MAX POOL3: 3x3 filters at stride 2 [4096] FC6: 4096 neurons [4096] FC7: 4096 neurons [1000] FC8: 1000 neurons (class scores)



Case Study: AlexNet

[Krizhevsky et al. 2012]

Full (simplified) AlexNet architecture: [227x227x3] INPUT [55x55x96] CONV1: 96 11x11 filters at stride 4, pad 0 [27x27x96] MAX POOL1: 3x3 filters at stride 2 [27x27x96] NORM1: Normalization layer [27x27x256] CONV2: 256 5x5 filters at stride 1, pad 2 [13x13x256] MAX POOL2: 3x3 filters at stride 2 [13x13x256] NORM2: Normalization layer [13x13x384] CONV3: 384 3x3 filters at stride 1, pad 1 [13x13x384] CONV4: 384 3x3 filters at stride 1, pad 1 [13x13x256] CONV5: 256 3x3 filters at stride 1, pad 1 [6x6x256] MAX POOL3: 3x3 filters at stride 2 [4096] FC6: 4096 neurons [4096] FC7: 4096 neurons [1000] FC8: 1000 neurons (class scores)



Details/Retrospectives:

- Popularized use of ReLU in Vision
- used Norm layers (not common anymore)
- heavy data augmentation
- dropout 0.5 in only last few fully-connected
- batch size 128
- SGD Momentum 0.9
- Learning rate 1e-2, reduced by 10 manually when val accuracy plateaus
- L2 weight decay 5e-4
- 7 CNN ensemble: 18.2% -> 15.4%

Aside: Useful Tool

http://dgschwend.github.io/netscope/#/editor

Chrome File Edit View History Bookmarks People Window	Help 🛇 😝 🛞 🗣 🎅 100% 🖼 Wed 11:31 AM 🔍 😑 Ielp 🛇 😂 🐇 🕚 🛡 🎓	🤶 100% E/D We
AlexNet — Netscope CNN Anal ×	Arjun	
\leftarrow \rightarrow C (i) dgschwend.github.io/netscope/#/editor	Q 🛧 🐠 🐻 🖸 🗄	Q
1 name: "AlexNet" 2 layer { 3 name: "data" 4 type: "Data" 5 top: "data" 6 top: "label" 7 include {	AlexNet (edit) AlexNet (edit)	4K 4k
8 phase: TRAIN 9 } 10 transform_param { 11 mirror: true 12 crop size: 227	image: space of the space	8k 62k 4k
13 mean_file: "data/ilsvrc12/imagenet_mean.binaryproto" 14 } 15 data_param { 16 source: "examples/imagenet/ilsvrc12_train_lmdb" 17 batch_size: 256 18 backend: LMDB 19 }	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	52k 62k
20} 21layer { 22 name: "data" 23 type: "Data" 24 type: "data"	Image: Note of the strength 9 pool2 Pooling 256 13x13 comp389.88k activation 64.9 256 13x13 256 13x13 384 13x13 macc 149.52M activation 64.9 256 13x13 256 13x13 macc 149.52M activation 64.9	6k k .74k
2 ²⁴ top: "dabet" 26 include { 27 phase: TEST 28 }	11 relu3 ReLU 384 13x13 384 13x13 comp 64.9k activation 64.9 12 conv4 Convolution 384 13x13 384 13x13 macc 224.28M activation 64.9 12 conv4 Convolution 384 13x13 macc 224.28M activation 64.9	k ik
29 transtorm_param (0 mirror: false 31 crop_size: 227 32 mean_file: "data/ilsvrc12/imagenet_mean.binaryproto" 33 }	add 13 relu4 ReLU 384 13x13 384 13x13 comp 64.9 k activation 64.9 add add 14 conv5 Convolution 384 13x13 256 13x13 macc 149.52M activation 43.2 add add </td <td>6k 74k</td>	6k 74k
34 data_param { 55 source: "examples/imagenet/ilsvrc12_val_lmdb" 36 batch_size: 50 37 backend: LMDB 38 }	2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2560 13013 2500 13013 2500 13013 2500 13013 2500 13013 2500 13013 2500 13013 2500 13013 2500 13013	6k .k
39} 40layer { 41 name: "conv1" 42 type: "Convolution" 43 bottom: "data"	add add <td>5M</td>	5M
44 top:"conv1" 45 param { 46 lr_mult: 1 47 decay_mult: 1	Image: Comparison of the second se	'8M
40 / 9 param { 50 lr_mult: 2 51 decay_mult: 0 52 }	21 relu7 ReLU 4096 1x1 4096 1x1 comp4.1k activation 4.1k 22 drop7 Dropout 4096 1x1 4096 1x1 comp4.1k activation 4.1k 23 fc8 InnerProduct 4096 1x1 1000 1x1 mace 4.1M activation 1000	<u>.</u>
53 convolution_param 1 54 num_output: 96 55 kernel_size: 11 56 stride: 4 57 weight_filter { 58 type: "gaussian"	Diversion Control Contro Control Control <	<u>1</u>
3 5 6 } 61 bias_filler { 62 type: "constant" 63 value: 0 64 } 65 > 67 >	2 conv1 Convolution 3 227x227 96 55x5 macc10542M activation 290.4k param 34.8k activation 290.4k activation 290.4k </td <td>м 7м</td>	м 7м
68 name: "relu" 69 type: "RAU" 70 bottom: "cony1"	div 58.0.8k exp 290.4k Details:	

Case Study: ZFNet

[Zeiler and Fergus, 2013]



ILVRC 2013

AlexNet but:

CONV1: change from (11x11 stride 4) to (7x7 stride 2) CONV3,4,5: instead of 384, 384, 256 filters use 512, 1024, 512

ImageNet top 5 error: 15.4% -> 14.8 (and later 11.2)%

		0.0104 - 1.02.0003 - D.04			
A	A-LRN	B	С	D	E
11 weight	11 weight	13 weight	16 weight	16 weight	19 weight
layers	layers	layers	layers	layers	layers
	i	:)			
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64
	LRN	conv3-64	conv3-64	conv3-64	conv3-64
	a	max	pool	a salah salah salah sa	
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
		conv3-128	conv3-128	conv3-128	conv3-128
		max	pool		
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
			conv1-256	conv3-256	conv3-256
					conv3-256
		max	pool	·	
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512 conv3-512		conv3-512	conv3-512	conv3-512
			conv1-512	conv3-512	conv3-512
	-				conv3-512
		max	pool		
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
			conv1-512	conv3-512	conv3-512
					conv3-512
		max	pool		
		FC-	4096]
		FC-	4096		
		FC-	1000		
		soft	-max		

Case Study: VGGNet [Simonyan and Zisserman, 2014]

Only 3x3 CONV stride 1, pad 1 and 2x2 MAX POOL stride 2

best model

11.2% top 5 error in ILSVRC 2013 ->

7.3% top 5 error (did not win!)

Table 2: Number of parameters (in millions).

Network	A,A-LRN	B	С	D	E
Number of parameters	133	133	134	138	144

INPUT: [224x224x3] memory: 224*224*3=150K params: 0 CONV3-64: [224x224x64] memory: 224*224*64=3.2M params: (3*3*3)*64 = 1,728 CONV3-64: [224x224x64] memory: 224*224*64=3.2M params: (3*3*64)*64 = 36,864 POOL2: [112x112x64] memory: 112*112*64=800K params: 0 CONV3-128: [112x112x128] memory: 112*112*128=1.6M params: (3*3*64)*128 = 73,728 CONV3-128: [112x112x128] memory: 112*112*128=1.6M params: (3*3*128)*128 = 147,456 POOL2: [56x56x128] memory: 56*56*128=400K params: 0 CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*128)*256 = 294,912 CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*256)*256 = 589,824 CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*256)*256 = 589,824 POOL2: [28x28x256] memory: 28*28*256=200K params: 0 CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*256)*512 = 1,179,648 CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*512)*512 = 2,359,296 CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*512)*512 = 2,359,296 POOL2: [14x14x512] memory: 14*14*512=100K params: 0 CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296 CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296 CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296 POOL2: [7x7x512] memory: 7*7*512=25K params: 0 FC: [1x1x4096] memory: 4096 params: 7*7*512*4096 = 102,760,448 FC: [1x1x4096] memory: 4096 params: 4096*4096 = 16,777,216 (params not counting biases) FC: [1x1x1000] memory: 1000 params: 4096*1000 = 4,096,000

ConvNet C	onfiguration		
В	С	D	
13 weight	16 weight	16 weight	19
layers	layers	layers	
put (224×2	24 RGB image	e	
conv3-64	conv3-64	conv3-64	cc
conv3-64	conv3-64	conv3-64	cc
max	pool		_
conv3-128	conv3-128	conv3-128	co
conv3-128	conv3-128	conv3-128	co
max	pool		
conv3-256	conv3-256	conv3-256	co
conv3-256	conv3-256	conv3-256	co
	conv1-256	conv3-256	co
			co
max	pool		
conv3-512	conv3-512	conv3-512	co
conv3-512	conv3-512	conv3-512	co
	conv1-512	conv3-512	co
			co
max	pool		
conv3-512	conv3-512	conv3-512	co
conv3-512	conv3-512	conv3-512	co
	conv1-512	conv3-512	co
			co
max	pool		
FC-	4096		
FC-	4096		
FC-	1000		
soft	-max		

INPUT: [224x224x3] memory: 224*224*3=150K params: 0 CONV3-64: [224x224x64] memory: 224*224*64=3.2M params: (3*3*3)*64 = 1,728 CONV3-64: [224x224x64] memory: 224*224*64=3.2M params: (3*3*64)*64 = 36,864 POOL2: [112x112x64] memory: 112*112*64=800K params: 0 CONV3-128: [112x112x128] memory: 112*112*128=1.6M params: (3*3*64)*128 = 73,728 CONV3-128: [112x112x128] memory: 112*112*128=1.6M params: (3*3*128)*128 = 147,456 POOL2: [56x56x128] memory: 56*56*128=400K params: 0 CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*128)*256 = 294,912 CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*256)*256 = 589,824 CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*256)*256 = 589,824 POOL2: [28x28x256] memory: 28*28*256=200K params: 0 CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*256)*512 = 1,179,648 CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*512)*512 = 2,359,296 CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*512)*512 = 2,359,296 POOL2: [14x14x512] memory: 14*14*512=100K params: 0 CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296 CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296 CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296 POOL2: [7x7x512] memory: 7*7*512=25K params: 0 FC: [1x1x4096] memory: 4096 params: 7*7*512*4096 = 102,760,448 FC: [1x1x4096] memory: 4096 params: 4096*4096 = 16,777,216 (params not counting biases) FC: [1x1x1000] memory: 1000 params: 4096*1000 = 4,096,000

TOTAL memory: 24M * 4 bytes ~= 93MB / image (only forward! Backward?) TOTAL params: 138M parameters

ConvNet C	onfiguration		
В	С	D	
13 weight	16 weight	16 weight	19
layers	layers	layers	
out (224×2)	24 RGB image	e)	
conv3-64	conv3-64	conv3-64	cc
conv3-64	conv3-64	conv3-64	cc
max	pool		
conv3-128	conv3-128	conv3-128	co
conv3-128	conv3-128	conv3-128	co
max	pool		
conv3-256	conv3-256	conv3-256	co
conv3-256	conv3-256	conv3-256	co
	conv1-256	conv3-256	co
			co
max	pool		
conv3-512	conv3-512	conv3-512	co
conv3-512	conv3-512	conv3-512	co
	conv1-512	conv3-512	co
9			co
max	pool		
conv3-512	conv3-512	conv3-512	co
conv3-512	conv3-512	conv3-512	co
	conv1-512	conv3-512	co
			co
max	pool		
FC-	4096		
FC-	4096		
FC-	1000		
soft	-max		

INPUT: [224x224x3] memory: 224*224*3=150K params: 0	
CONV3-64: [224x224x64] memory: 224*224*64=3.2M params: (3*3*3)*64 = 1,728	Noto
CONV3-64: [224x224x64] memory: 224*224*64=3.2M params: (3*3*64)*64 = 36,864	Note:
POOL2: [112x112x64] memory: 112*112*64=800K params: 0	Most momory
CONV3-128: [112x112x128] memory: 112*112*128=1.6M params: (3*3*64)*128 = 73,728	wost memory
CONV3-128: [112x112x128] memory: 112*112*128=1.6M params: (3*3*128)*128 = 147,456	is in parly
POOL2: [56x56x128] memory: 56*56*128=400K params: 0	15 III Carry
CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*128)*256 = 294,912	CONV
CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*256)*256 = 589,824	CONV
CONV3-256: [56x56x256] memory: 56*56*256=800K params: (3*3*256)*256 = 589,824	
POOL2: [28x28x256] memory: 28*28*256=200K params: 0	
CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*256)*512 = 1,179,648	
CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*512)*512 = 2,359,296	
CONV3-512: [28x28x512] memory: 28*28*512=400K params: (3*3*512)*512 = 2,359,296	
POOL2: [14x14x512] memory: 14*14*512=100K params: 0	
CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296	Most params
CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296	
CONV3-512: [14x14x512] memory: 14*14*512=100K params: (3*3*512)*512 = 2,359,296	are
POOL2: [7x7x512] memory: 7*7*512=25K params: 0	
FC: [1x1x4096] memory: 4096 params: 7*7*512*4096 = 102,760,448	in late FC
FC: [1x1x4096] memory: 4096 params: 4096*4096 = 16,777,216 (params not counting biases)	
FC: [1x1x1000] memory: 1000 params: 4096*1000 = 4,096,000	(Avg. Pool)
TOTAL memory: 24M * 4 bytes ~= 93MB / image	

(only forward! ~*2 for bwd)

TOTAL params: 138M parameters

Case Study: GoogLeNet

[Szegedy et al., 2014]



In this paper, we will focus on an efficient deep neural network architecture for computer vision, codenamed Inception, which derives its name from the Network in network paper by Lin et al [12] in conjunction with the famous "we need to go deeper" internet meme [1]. In our case, the word

[1] Know your meme: We need to go deeper. http://knowyourmeme.com/memes/ we-need-to-go-deeper. Accessed: 2014-09-15.





Case Study: GoogLeNet



Case Study: GoogLeNet

[Szegedy et al., 2014]

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	$112 \times 112 \times 64$	1							2.7K	34M
max pool	3×3/2	$56 \times 56 \times 64$	0								
convolution	3×3/1	$56 \times 56 \times 192$	2		64	192				112K	360M
max pool	3×3/2	$28 \times 28 \times 192$	0								
inception (3a)		$28 \times 28 \times 256$	2	64	96	128	16	32	32	159K	128M
inception (3b)		$28 \times 28 \times 480$	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	$14 \times 14 \times 480$	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		$14 \times 14 \times 512$	2	160	112	224	24	64	64	437K	88M
inception (4c)		$14 \times 14 \times 512$	2	128	128	256	24	64	64	463K	100M
inception (4d)		$14 \times 14 \times 528$	2	112	144	288	32	64	64	580K	119M
inception (4e)		$14 \times 14 \times 832$	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	$7 \times 7 \times 832$	0				8				
inception (5a)	8	$7 \times 7 \times 832$	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	$1 \times 1 \times 1024$	0								
dropout (40%)		$1 \times 1 \times 1024$	0	1		2					
linear		$1 \times 1 \times 1000$	1			3 2				1000K	1M
softmax		$1 \times 1 \times 1000$	0								6

Fun features:

- Only 5 million params! (Removes FC layers completely)

Compared to AlexNet:

- 12X less params
- 2x more compute
- 6.67% (vs. 16.4%)

Case Study: ResNet

[*He et al., 2015*] ILSVRC 2015 winner (3.6% top 5 error)



Slide from Kaiming He's recent presentation <u>https://www.youtube.com/watch?v=1PGLj-uKT1w</u>



(slide from Kaiming He's recent presentation)

CIFAR-10 experiments



Case Study: ResNet

[He et al., 2015] ILSVRC 2015 winner (3.6% top 5 error)



(slide from Kaiming He's recent presentation)


[He et al., 2015]



Case Study: ResNet [He et al., 2015]





Case Study: ResNet [He et al., 2015]

- Batch Normalization after every CONV layer
- Xavier/2 initialization from He et al.
- SGD + Momentum (0.9)
- Learning rate: 0.1, divided by 10 when validation error plateaus
- Mini-batch size 256
- Weight decay of 1e-5
- No dropout used

ILVRC 2016

- Trimps-Soushen was the winner with an ensemble approach.
 - Jie Shao, Xiaoteng Zhang, Zhengyan Ding, Yixin Zhao, Yanjun Chen, Jianying Zhou, Wenfei Wang, Lin Mei, Chuanping Hu, The Third Research Institute of the Ministry of Public Security, P.R. China.
- Classification error is down to 3.0% from 3.6% last year.
- A bit boring:
 - <u>https://www.reddit.com/r/MachineLearning/comments/54jiyy/lar</u> <u>ge_scale_visual_recognition_challenge_2016/</u>
 - <u>http://image-net.org/challenges/LSVRC/2016/results#loc</u>

ILVRC 2017, Squeeze & Excitation Network

- Squeeze and Excitation block that can be added to a ConvLayer
- Distributed deep learning training suing "ROCS". Our system trains SE-ResNet152 with a minibatch size of 2048 on 64 Nvidia Pascal Titan X GPUs in 20 hours using synchronous SGD
- Winning entry ensemble of SENets that obtain a 2.251% top-5 error on the test set. This result represents a ~25% relative improvement on the winning entry of 2016 (2.99% top-5 error).

ILVRC 2017, Squeeze & Excitation Network

- Squeeze and Excitation block that can be added to a ConvLayer
- Let's add parameters to each channel of a convolutional block so that the network can adaptively adjust the weighting of each feature map.



ILVRC 2012 - 2017



ILVRC 2017, Squeeze & Excitation Network



Why ConvNets?



Summary

- ConvNets stack CONV, POOL, FC layers
- Trend towards smaller filters and deeper architectures
- Trend towards getting rid of POOL/FC layers (just CONV)
- Typical architectures look like [(CONV-RELU)*N-POOL?]*M-(FC-RELU)*K-SOFTMAX

where N is usually up to ~5, M is large, 0 <= K <= 2.

 but recent advances such as ResNet/GoogLeNet challenge this paradigm

Things you should know:

"ConvNets need a lot of data to train"

Things you should know:

"ConvNets need a lot of data to train"



finetuning! we don't always need to train ConvNets from scratch.





conv-64 conv-64 maxpool	1. Train on Imagenet
conv-128 conv-128 maxpool	
conv-256 conv-256 maxpool	
conv-512 conv-512 maxpool	
conv-512 conv-512 maxpool	
FC-4096 FC-4096 FC-1000 softmax	



2. Small dataset: feature extractor





Interne	2. Small dataset:
Image	fastura axtractor
onv-64	
onv-64	
naxpool	
onv-128	
onv-128	
naxpool	
onv-256	
onv-256	
naxpool	Freeze these
onv-512	
onv-512	
naxpool	
onv-512	
onv-512	
naxpool	
C-4096	
C-4096	<u></u>
C-1000	
oftmax	Irain this



3. Medium dataset: finetuning

more data = retrain more of the network (or all of it)

Freeze these

tip: use only ~1/10th of the original learning rate in finetuning top layer, and ~1/100th on intermediate layers

Train this

CNN Features off-the-shelf: an Astounding Baseline for Recognition [Razavian et al, 2014]

DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition [Donahue*, Jia*, et al., 2013]

	DeCAF ₆	DeCAF ₇	
LogReg	40.94 ± 0.3	40.84 ± 0.3	
SVM	39.36 ± 0.3	40.66 ± 0.3	
Xiao et al. (2010)	38.0		



image				
conv-64				
conv-64			vory similar	vorv difforent
maxpool				
conv-128	more generic		dataset	dataset
conv-128				
maxpool				
conv-256		very little data	?	?
conv-256				
maxpool	more specific			
conv-512	/			
conv-512				
maxpool				
conv-512				
conv-512		quite a lot of	?	?
maxpool		data		
FC-4096		uala		
FC-4096	-			
FC-1000				
softmax				

image			
conv-64 conv-64 maxpool conv-128 conv-128 more generic		very similar dataset	very different dataset
maxpool conv-256 conv-256 maxpool conv-512 conv-512 maxpool	very little data	Use Linear Classifier on top layer	?
conv-512 conv-512 maxpool FC-4096 FC-1000 softmax	quite a lot of data	Finetune a few layers	?

image			
conv-64 maxpool conv-128 conv-128		very similar dataset	very different dataset
maxpool conv-256 conv-256 maxpool conv-512 conv-512 maxpool	very little data	Use Linear Classifier on top layer	You're in trouble Try linear classifier from different stages
conv-512 conv-512 maxpool FC-4096 FC-4096 FC-1000 softmax	quite a lot of data	Finetune a few layers	Finetune a larger number of layers

Transfer learning with CNNs is pervasive... (it's the norm, not an exception)



Object Detection (Faster R-CNN) Image Captioning: CNN + RNN



Transfer learning with CNNs is pervasive... (it's the norm, not an exception)



Object Detection (Faster R-CNN)

Transfer learning with CNNs is pervasive... (it's the norm, not an exception)



E.g. Caffe Model Zoo: Lots of pretrained ConvNets https://github.com/BVLC/caffe/wiki/Model-Zoo https://github.com/szagoruyko/loadcaffe

Model Zoo

ELM edited this page 21 days ago - 56 revisions

Check out the model zoo documentation for details.

To acquire a model:

- download the model gist by ./scripts/download_model_from_gist.sh <gist_id> <dirmame> to load the model metadata_architecture, solver configuration, and so on. (<dirmame> is optional and defaults to caffeinnedes).
- download the model weights by ./scripts/download_model_binary.py <model_dir> where <model_dir> is the gist directory from the first step.

or visit the model zoo documentation for complete instructions.

Berkeley-trained models

- Finetuning on Flickr Style: same as provided in models/, but listed here as a Gist for an example.
- BVLC GoogleNet: models/bvlc_googlenet

Network in Network model

The Network in Network model is described in the following ICLR-2014 paper:

Network In Network M. Lin, Q. Chem, S. Yan International Conference on Learning Representations, 2014 (arXiv:1409.1556)

please cite the paper if you use the models.

Models:

 NIN-Imagenet: a small(29MB) model for imagenet, yet performs slightly better than AlexNet, and fast to train. (Note: a more caffe-compatible version with correct convolutional weights shape: https://drive.google.com/foiderview?

Id=080ledYUunOQINEFIUI1QNWVhVVU&usp=drive_web)
NINLCIFAR10: NIN model on CIFAR10, originally published in the paper Network in Network

The error rate of this model is 10.4% on CIFAR10.

Models from the BMVC-2014 paper "Return of the Devil in the Details: Delving Deep into Convolutional Nets"

The models are trained on the ILSVRC-2012 dataset. The details can be found on the project page or in the following BMVC-2014 page:

Return of the Devil in the Details: Delving Deep into Convolutional Nets K. Chatfield, K. Simonyan, A. Vedaldi, A. Zisserman British Machine Vision Conference, 2014 (arXiv ref. cs1495.3531)

Please cite the paper if you use the models.

Models

- VGG_CNN_S: 13.1% top-5 error on ILSVRC-2012-val
- VGG_CNN_M: 13.7% top-5 error on ILSVRC-2012-val
- VGG_CNN_M_2046: 13.5% top-5 error on ILSVRC-2012-val
- VGG_CNN_M_1024: 13.7% top-5 error on ILSVRC-2012-val
 VGG_CNN_M_128: 15.6% top-5 error on ILSVRC-2012-val
- VGG_CNN_F: 18.7% top-5 error on ILSVRC-2012-val
- voo_onid_P; to.r % top-a error on iL3 VRC-2012-Val

Models used by the VGG team in ILSVRC-2014

Places-CNN model from MIT. Places CNN is described in the following NIPS 2014 page

B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva Learning Deep Features for Scene Recognition using Flaces Database. Advances in Neural Information Processing Systems 27 (NIPS) spotlight, 2014.

The project page is here

- Placebild-sharehold in 205 scene categories of Placeb Database (and in NPT 14) and ~ 2.5 million mapses. The architecture is the same as Cafe inference network, and the state of the state of the same as Cafe inference inference and and PD database (adaptives from the bank and et al.UNC202) (mapshift) with ~ 2.6 million mapses. The architecture is the same as Cafe inference network. Placebild-comparel: Compared for Manated on 20 million (2016)
- Placesco-occipterie: Cooperies one same on so scene categories of Places Database. It is used by Google In the deep dream visualization

GoogLeNet GPU implementation from Princeton.

We implemented GoogLeNet using a single GPU. Our main contribution is an effective way to initialize the network and a trick to overcome the GPU memory constraint by accumulating gradients over two training iterations.

models on ImageNet and Places, and the training code are available for download. • Make sure ds2_fc2 and ds3_fc have num_output = 1000 in the problet. Otherwise, the trained model would crash on test. Fully Convolutional Semantic Segmentation Models (FCN)

WGoodLeNet/ for more information. Pre-trained

Xs) These models are described in the paper

Fully Convolutional Models for Semantic Segmentation Jonathan Long, Evan Shelhaner, Trevor Darrell CVPR 205 arXiv:1411.4038

Dieses check http://deion.pdpcatps.ed.

They are available under the same license as the Caffe-bundled models (i.e., for unrestricted use see http://caffe.berkeleyvision.org/model_zoo.htmi#bw/c-model-license). These are pre-release models. They do not run in any current version of BVLC/caffe, as they

require unmerged PRs. They should run in the preview branch provided at https://github.com/longion/uafte/tee/huture. The FCN-335 PA3CAL-Context model is the most complete example including network definitions, solver configuration, and Python scripts for solving and interence.

Models trained on PASCAL (using extra data from Hariharan et al. and finetuned from the ILSVRC-trained VGG-16 model above):

FCN-32s PASCAL: single stream, 32 pixel prediction stride version
 FCN-35s PASCAL: two stream, 16 pixel prediction stride version
 FCN-8s PASCAL: three stream, 6 pixel prediction stride version

 FCN-AlexNet PASCAL: AlexNet (CaffeNet) single stream, 32 pixel prediction sindle version To reproduce the validation scores, use the seg11valid split defined by the paper in footnote 7. Since BBD train and PASCAL VOC 11 segval intersect, we only evaluate on the non-intersecting well for validation numbers.

- Models trained on SIFT Flow (also finetuned from VGG-16):
- FCN-10s SIFT Flow: two stream, 16 pixel prediction stride version
- Models trained on NYUDV2 (also finetuned from VGG-16, and using HHA features from Gupta et al. https://github.com/s-gupta/rcnn-depth); FGN-32s NYUDV2: single stream. 32 pixel prediction stride version
- FCN-16s NYLDV2: two stream, 10 pixel prediction stride version
 Models trained on PASCAL-Context including training model definition, solver configuration, and
- barebones solving script (finetuned from the ILSVRC-trained VGG-16 model): FCN-325 PASCAL-Context: single stream, 32 pixel prediction stride version
- FCN-16s PASCAL-Context: two stream, 16 pixel prediction stride version
 FCN-8s PASCAL-Context: three stream, 8 pixel prediction stride version

CaffeNet fine-tuned for Oxford flowers dataset

The is the reference Catholic (modified Alsolve) fine-luned for the Oxford 100 calegory flower dataset. The number of outputs in the inner product layer has been set to 100 to reflect the number of homor calegories. Theoretainment of the set of the set of the layer Recognition on "Flow? Sight" Data. The goods learning rate is reduced which the learning rate for the flaw layer.

After 50,000 iterations, the top-1 error is 7% on the test set of 1,020 images.

CNN Models for Salient Object Subitizing.

CNN models described in the following CVPR'15 paper "Salent Object Sublizing":

Salient Object Subilizing J. Zhang, S. Ma, M. Sameki, S. Sclaroff, M. Betke, Z. Lin, X. Shen, B. Price and R. CVMP. 245.

.∢ _____

- AlexNet: CNN model finetuned on the Salient Object Sublizing dataset (~5500 images). The architecture is the same as the Caffe reference network.
- VGG18: CNN model finebuned on the Satient Object Sublitzing dataset (~5500 images). The architecture is the same as the VGG16 network. This model gives better performance than the AexNet model: but is solver for training and testing.

Deep Learning of Binary Hash Codes for Fast Image Retrieval

We present an effective deep learning framework to create the hash-like binary codes for fast image retrieval. The details can be found in the following "CVPRW15 paper":

Deep Learning of Binary Hash Codes for Fast Image Retrieval K. Lin, H.-F. Yang, J.-H. Hsiao, C.-S. Chen CVPR 2815, DeepVision workshop

please cite the paper if you use the model:

- caffe-cvprw15: See our code release on Github, which allows you to train your own deep backling model and create bloary back codes.
- CIFAR10-46bit: Proposed 48-bits CNN model trained on CIFAR10.

Places CNDS models on Scene Recognition

 Places-CNDS-5 is a "8conv3tc layer" deep Convolutional neural Networks model trained on MIT Places Dataset with Deep Supervision.

The details of training this model are described in the following report. Please cite this work if the model is useful for you.

Training Deeper Convolutional Networks with Deep Supervision L.Wang, C.Lee, Z.Tu, S. Lazebnik, arXiv:1585.82496, 2015

Models for Age and Gender Classification.

 AgerGender.net are models for age and gender classification trained on the Adience-OUI dataset. See the Project page.

The models are described in the following paper.

Age and Gender Classification using Convolutional Neural Networks Gil Levi and Tal Hassner IEEE Workshop on Analysis and Nodeling of Faces and Gestures (AMFG), at the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Boston, June 28

If you find our models useful, please add suitable reference to our paper in your work

GoogLeNet_cars on car model classification

GoogLeNet_carr is the GoogLeNet model pre-trained on ImageNet classification task and finetuned on 431 car models in CompCars dataset. It is described in the technical report. Please cite the following work if the model is useful for you.

A Large-Scale Car Dataset for Fine-Grained Categorization and Verification L. Yang, P. Luo, C. C. Loy, X. Tang, arXiv:1586.08050, 2015

Holistically-Nested Edge Detection

The model and code provided are described in the ICCV 2015 paper

Holistically-Nested Edge Detection Saining Xie and Zhuowen Tu ICCV 2015

For details about training/evaluating HED, please take a look at http://github.com/s9xle/hed

Model trained on BSDS-500 Dataset (finetuned from the VGGNet

· HED BSDS-500

Translating Videos to Natural Language

These models are described in this NAACL-HLT 2015 paper.

Translating Videos to Natural Language Using Deep Recurrent Neural Networks 5. Venugopalan, H. Xu, J. Donahue, M. Rohrbach, R. Mooney, K. Seenko MACL-HLT 2015

More details can be found on this project page.

Node:: Video2Text_VGG_mean_pool: This model is an improved version of the mean pooled model described in the NAACL-HLT 2015 paper. It uses video frame features from the VGG-16 layer model. This is trained only on the Youtube video dataset.

Compatibility: These are pre-release models. They do not run in any current version of BVLC/caffe, as they require unmerged PRs. The models are currently supported by the recurrent branch of the Caffe fork provided at https://github.com/jeffdonahue/zaffe/hee/recurrent and https://github.com/sethose/https://diffe/recurrent.

VGG Face CNN descriptor

These models are described in this BMVC 2015 pape

Deep Face Recognition Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman BMVC 2015

More details can be found on this project page

Model: VGG Face: This is the very deep architecture based model trained from scratch using 2.6 Million images of celebrities collected from the web. The model has been imported to work with Caffe from the original model trained using MacConvNet literary.

If you find our models useful, please add sultable reference to our paper in your work.

Yearbook Photo Dating

Model and prototxt files: Yearbook

ICCV 2815

arXiv:1586.83648

Model from the ICCV 2015 Extreme Imaging Workshop paper:

Weakly Supervised Segmentation

These models are described in the ICCV 2015 paper

Deepak Pathak, Philipp Krähenbühl, Trevor Darrell

A Century of Portraits: Exploring the Visual Historical Record of American High Scho Shiry Ginosar, Kate Rakelly, Brian Yin, Sarah Sachs, Alyosha Efros ICCV Workshop 2015

. . .

IIT Bombay, Perceptive Code LLC

CCNN: Constrained Convolutional Neural Networks for

Constrained Convolutional Neural Networks for Weakly Supervised Segmentation

These are pre-release models. They do not run in any current version of BVLC/caffe, as they

require unmerged PRs. Full details, source code, models, prototxts are available here: CCNN

Thank you!

Data Preprocessing

Data Preprocessing

TLDR: In practice for Images: center only

e.g. consider CIFAR-10 example with [32,32,3] images

- Subtract the mean image (e.g. AlexNet) (mean image = [32,32,3] array)
- Subtract per-channel mean (e.g. VGGNet) (mean along each channel = 3 numbers)

Not common to normalize variance, to do PCA or whitening

Data Augmentation

Data Augmentation

4. Get creative!

Random mix/combinations of :

- translation (what about a pure ConvNet?)
- rotation
- stretching
- shearing,
- lens distortions, ... (go crazy)

Data Augmentation: Takeaway

Simple to implement, use it

Especially useful for small datasets, helps a lot with fixing overfitting

Fits into framework of noise / marginalization

Weight Initialization

Lets look at

some activation statistics



self.W = torch.randn(fan_out, fan_in) * 1.0 self.b = torch.randn(fan_out)* 1.0

input layer had mean 0.001800 and std 1.001311 hidden layer 1 had mean -0.000430 and std 0.981879 hidden layer 2 had mean -0.000849 and std 0.981649 hidden layer 3 had mean 0.000566 and std 0.981601 hidden layer 4 had mean 0.000483 and std 0.981755 hidden layer 5 had mean -0.000682 and std 0.981614 hidden layer 6 had mean -0.000401 and std 0.981560 hidden layer 7 had mean -0.000237 and std 0.981520 hidden layer 8 had mean -0.000448 and std 0.981728 hidden layer 9 had mean -0.000889 and std 0.981728 hidden layer 10 had mean 0.000584 and std 0.981736

*1.0 instead of *0.01



Almost all neurons completely saturated, either -1 and 1. Gradients will be all zero. self.W = torch.randn(fan_out, fan_in) / math.sqrt(fan_in)
self.b = torch.randn(fan_out) / math.sqrt(fan_in)

"Xavier initialization" [Glorot et al., 2010]

Reasonable initialization. (Mathematical derivation assumes linear activations)



input layer had mean 0.001800 and std 1.001311 hidden layer 1 had mean 0.001198 and std 0.627953 hidden layer 2 had mean -0.000175 and std 0.486051 hidden layer 3 had mean 0.000055 and std 0.407723 hidden layer 4 had mean -0.000306 and std 0.357108 hidden layer 5 had mean 0.000142 and std 0.320917 hidden layer 6 had mean -0.000389 and std 0.292116 hidden layer 7 had mean -0.000228 and std 0.273387 hidden layer 8 had mean -0.000291 and std 0.254935 hidden layer 9 had mean 0.000139 and std 0.228008



IIT Bombay, Perceptive Code LLC
self.W = torch.randn(fan_out, fan_in) / math.sqrt(fan_in/2)
self.b = torch.randn(fan_out) / math.sqrt(fan_in/2)

He et al., 2015 (note additional /2)

input layer had mean 0.000501 and std 0.999444 hidden layer 1 had mean 0.562488 and std 0.825232 hidden layer 2 had mean 0.553614 and std 0.827835 hidden layer 3 had mean 0.545867 and std 0.813855 hidden layer 4 had mean 0.565396 and std 0.826902 hidden layer 5 had mean 0.547678 and std 0.834092 hidden layer 6 had mean 0.587103 and std 0.860035 hidden layer 7 had mean 0.596867 and std 0.870610 hidden layer 8 had mean 0.623214 and std 0.889348 hidden layer 9 had mean 0.567498 and std 0.844523





Batch Normalization

"you want unit gaussian activations? just make them so."

consider a batch of activations at some layer. To make each feature dimension unit gaussian, apply:

$$\widehat{x}^{(k)} = \frac{x^{(k)} - \mathbf{E}[x^{(k)}]}{\sqrt{\mathrm{Var}[x^{(k)}]}}$$

this is a vanilla differentiable function...

Batch Normalization

[loffe and Szegedy, 2015]

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1m}\}$; Parameters to be learned: γ, β Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$									
$\mu_{\mathcal{B}} \leftarrow \frac{1}{m} \sum_{i=1}^{m} x_i$	// mini-batch mean								
$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\mathcal{B}})^2$	// mini-batch variance								
$\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}}$	// normalize								
$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathbf{BN}_{\gamma,\beta}(x_i)$	// scale and shift								

- Improves gradient flow through the network
- Allows higher learning rates
- Reduces the strong dependence on initialization
- Acts as a form of regularization in a funny way, and slightly reduces the need for dropout, maybe

Babysitting the Learning Process

Notebook

https://github.com/stencilman/CS763_Spring2017/blob/master/Lec3%2C4/CrossEntropy-Linear.ipynb

2. Data Preprocessing: We compute the mean and standard deviation 'images' and then subtract and divide by the same respectively (like AlexNet). We also visualize them.

```
In [3]: x_mean = torch.mean(tr_x:float(), 1)
x_std = torch.std(tr_x:float(), 1)
itorch.image(x_mean)
itorch.image(x_std)

In [7]: function get_xi(data_x, idx)
xi = (data_x[idx]:float() - x_mean)
xi = xi:cdiv(x_std)
xi = xi:reshape(3*32*32)
return xi
end
```

Step 2: Choose the architecture: Say we start with single layer network:



1. Data Loading: Let us load the training and the test data and check the size of the tensors. Let us also display the first few images from the training set.

```
In [1]: -- load trainin images
tr_x = torch.load('cifar10/tr_data.bin')
-- load trainin labels
tr_y = torch.load('cifar10/tr_labels.bin'):double() + 1
-- load test images
te_x = torch.load('cifar10/te_data.bin')
-- load test labels
te_y = torch.load('cifar10/te_labels.bin'):double() + 1
print(tr_x:size())
print(tr_y:size())
```

Out[1]: 50000 3 32 [torch.LongStorage of size 4] 50000 [torch.LongStorage of size 1]

```
In [2]: -- display the first 36 training set images
require 'image';
itorch.image(tr_x[{{1,36},{},{},}])
```



Double check that the loss is reasonable:

```
op = model:forward(xi)
loss_tr = criterion:forward(op, ti)
print(loss_tr)
```



Double check that the loss is reasonable:

```
op = model:forward(xi)
loss_tr = criterion:forward(op, ti,
print(loss_tr)
```



Lets try to train now...

Tip: Make sure that you can overfit very small portion of the training data

```
tr_x = tr_x[{{1,20},{},{},{},{}]
te_x = tr_x[{{1,20},{},{},{}]
tr_y = tr_y[{{1,20}}]
print(tr_x:size())
print(tr_y:size())
Out[14]: 20
3
32
32
[torch.LongStorage of size 4]
20
[torch.LongStorage of size 1]
```

```
-- run it
lr = 0.0001
lambda = 0
train_and_test_loop(100000, lr, lambda)
```

The above code:

- take the first 20 examples from CIFAR-10
- turn off regularization (reg = 0.0)
 - use simple vanilla 'sgd'

Random Search vs. Grid Search



Random Search for Hyper-Parameter Optimization Bergstra and Bengio, 2012

Hyperparameters to play with:

- network architecture
- learning rate, its multiplier schedule
- regularization (L2/Dropout strength)

neural networks practitioner music = loss function



Karpathy's crossvalidation "command center"

Instant States

auther 1	workey 2	Section 1	sector 4	Server 1	anytes &	airkey 1	perter 1	atcher 3	webo 30	2	essiler 12	sorters 13
And process, Stat-Do-	service and the control of the service of the servi	And and a street and	Law, appendix 2015-20-	and apprent 201-02-	Anti-upland; (122-50) family from the county	Last spikken 2017-D-	Contraction 2022-01-	han wrene mil-th-	And Address Statistics		Los grinted 1004-12-	Lass method illu-D-
Free MMAGE/M	time LEPTEVILIDAE	View MURANON	The DITHOUGH	ation: 1201/Molett	New SPREAD (19)	The distriction	Line 200000-00	Vier TTMATERIA	Alex NORTHERE		the dorwing the	The STREETING
Anno philo domantati y A. Machanitta Romin, Tantan estimitina al setti 2011, "Anto base thata?" (1)	Proceeding of the second secon	Prop. years produced in Pagessor's Process, Takap seconding score"s.	Personal Annual Control Control (Control Control Contr	Price role sponsters in Parson?". Pricework?.	Print only decident's in "manager", Print, "deeps moreling of all,	Poly of a blocker's to "assest's bond", "seas wording stat", 201, "seas an other state",	Party proce provident's 11, "Manager's Putter, Transporter American proof's	Party (with provident's in the second to function, foreign extending of \$4") with, foreign but thereign at \$4".	Paris relationships for the framework of the first of the first second s		Proc. Mills Provability 11, "Annual to Provide, Transport estimation (COV), Phys. Science and Constants, 12,	Party, prior president's all "Reception" of Transfer Transfer encoding with the Party, "Same and the second of the
New area decider's t. s.	Party prof. Storter's Lt., Party etcology (1997) 405.	Name your promited a him.	Paragraphic annually 1.5. Perceptional and 1.505	Press, and Jonato's (New year month? Lt.	Party area accessive to the	Name and successful 2.5, Never receiving scient (10),	NAME AND DESCRIPTION AND DESCRIPTICANO AND DESCRIPTICANO AND DESCRIPTICANO AND DESCRIPTICANO A	Paragentic Income 1 1.34		National actions, and a state of the state o	Para prop markets 2, 15 Para receipting and 2, 20
the second to be the transmitter	Paral Internet, St. Paral, Internet Call, Manager, P. Paral, Approximation, Parallel	Press, Sports, St. W., Preprint's Press, Sandy Scient's US, Preprint's Print, Approximate Control.	Pour decrary (n. Presi, secon marks (n. Tesamori, Time, "Second of Fiscal	Party and the second se	Frank international and a set fragment	The state ways of a based	Part Start, Co. Start,	Part South? ' W. Prof. Star, South etc., "Southaft, Florid, Spectrum?, Tom?	Tent, barrier in, finnerit,		Frail Mary starts Mil. Tesamit's	Print, Intelligence 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997, 1997
Non-Berlin, and the Net-	Press, places, and press in a	Print Str. 17 and 191 mil	The second secon	Parameters and a spectrum of the second	Print, Mar. 10, America An.	Party provide and the real	Part and a second secon	Print and the matching on	Print provide and the state		Paris property and the second se	Post pho an among the dis-
Parolis manapulat par Minescol ⁴ 1 21, Philay marks 1, 605 Para Yang marks 1, Paratas scatte	 Perfile Deligned pp. Greens P1 91, "Sing party 1, 405, Nam (her) warfs 1, 5, Station starfs. 	President and the second secon	Petty (refution and trefs. http://www.ueu/contents. Para street anothe 1, Names anaph.	Partic Heads in the state of	Production and the second seco	Party checked by there of a term	Performance and and an and a series of the s	Particle Constantial and Constant Pro- late "Minageneral" multi- Parts Into Sector 1, Physics Const.	Party Original Attantic Pt		Parties discription and interest PL Provident month in the Provident month in Allanee conthe	Particle compared and investments
PROPERTY L.	and, "possible") in the second	ant, "sub-traff to some china propin manufity	PL Selecter 1. Normer, result (Marin, Branneth)	This real down is "prime provi-	10, Nobel of L. Sector P.	The Solid of the American Street of the Solid	Notes and and another	State state and processors	ert, "whethe" is preserve.		 Netwiden, S. Netwiden, Strategie, Strategie, C. 	100, "Anterpre") () Sector stars initial himstellers
Antonipal Internet.	Section Contractor	Anno miles miles 44,	(Articles (Deriver) Party (Mark (Deriver)	Tangton Repetite Provident	Construction Construction	Paint man man 14	Paint man, marks we	Provident of Aug. 10.	State of the state		Party name states in	Northerstein topy, in.
Converse storts of 2014.	Phanipping sign's 1, Mills, "Simplify of per American".	Constant of other distances	"Description of all statistics,"	Paris mail http://www.	Conception and 1,2000,	Conversion of the second secon	"Dampelie wrate (hereite)	Contract of the Contract of th	Conception, and a conception of the		"Design of a start from the "	The Special Adaption (Contract)
decertary on " the goal second is	Second second a final state of the line of the second seco	Services Inf., No. and Service C.	manufactors", "no and determines,	" seal of high diversion of the	Section 107, No and Amil's S.	interference", but and starth in-	management of the goad started of the	invalues or C. Nor and onset (in-	and state by C. No. and Street's Approximate Street, Street, Street, Street, St. N.		annulaester, No published. No	annaites in C. Sai anai sheaft in the
A American and a second state	Page 1, Jacob 10, Nught 1, John 1, Tana 1, Jacob 1, Tanjar 1, Tana 1, Tanjar 1, Tanjar 1, T	Participanty My Supply Links, Resides apply 14-1, forcesty	Paper, 1997 In. Super commun. Research and State, Super-	Special segment of the second of the second	Pravid, your's lab," hugh's hubble	Constant and head head in	And start set, taget roles	Participant's the "super- commu-	Practic county and the fourth to the		Partic (conf) m. "dap" (contact, Percent apple det, folger)	Plant, send in Approximation
the set product of a	wante programming to the plane per lange 1, 200 per 1, 200 per 1, 200 per 1, 200 per	Para propagation and and a second second	Party and products (3.5)	Finances, from persiste Life.	Name and Annual States of	Patricipani, consulation (1)	Freeze genie percentrality's factor of	Public Line, annual Park, Science, State	Alter and March/L.Mix		Fairle programmer's Cr. (the are accounted as a second
12, and	12,082	12 442	12 026	tion per sector and	11.14	12. 478	12 288	12.878	11.903		12.182	18. Mg
-1 11				-	1 11	-1 D D		120 T T T T T			and the second s	20 000
	12.1	1		- Harris				. et .	3		1.0	
-		-		a manufacture description	ALL		13				2	
Strikterry	A CONTRACTOR OF A CONTRACTOR OFTA A	A Present and a state	Constant and a second	Tanta and the second second	A Service merel	Continuentian	- Porchardered	Concernant and	1		for section in many	C. C. Strategie and Strategies and
autors 14	Marchine 11	Barrier 18	Martin 17	Marchen 18	factor 11	Mainter 31	Sector 11	Marchine 11	(Linna z)	Barahar 24	August 25	Jacober 25
TAN MINING JULY-Do-	tore arginal lot-	ANT. BRANK, TEP-D-	Los appears 112-21-	and address 201-0-	Anty spiked) (102-50)	Last spinster, 1917-D-	Carl spatest 202-0-	Date manager (Co-f)-	new press	Carry and the Pile Ci-	Los mintes pu-to-	Dev addeds (00)-D-
month 2. 24.760	month La Jacking	manufe 20. 66/00	maint 22. W5/00	mench 20. Te/mi	manufacture and the little	second in 42/80	manuk 12, PR/M	manch 20. 82/76/	Summer A. 16784	Marcola, N. 140/74	mainth of Amilia	mente da fecter
the play store in.	Print relations of the Advanced	responses to the second	S "platt, cop", its	Not the product of the second	Personal and an entropy of the Page and the	April 110" ().	The manufacture is the second	Aglad, 2710-7114, Pate phile minester's its fragments	"And right in the Starger's	The second secon	Personal Production in Pressoning	the standard is harring
North Annual South Provide Street Str	Phone and memory car	Print, "state states", d.	Purport, Charge estimation, etc.etc. and, "setta tool, thereast", etc. Purport article another y 1276.	Name your and an and the	Party "With the bary of the bar	Print and another's c.b.	Party Science Annual and Annual State	Print Party and Statistics and	First, "white an install "1 day	Print for control of	Policie da Carta da C	Party Composition (1971)
Next Association of 10% (201), Next Association (201)	Parts withing tracks (b), was been at the	Part and a faith of a	Providential state") 405	Providence and the second seco	Party and a cart of	Peers working tracks bit, Page provider: the	Part Annulus and Cally	Party and Strategy and Art	Party studiets, starts, 45, 45, 46, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10	Non-pice and a second of the	Compared Anisotry 2.5, Compared and Anisotry 10,	These recently sounds and
Call, "property fight".	Planet, SpeecelsPi Planet, Planet, St. M., Marth, By	Tunin", "geomyrun": "Labet; New yes, Le amartes mis	Tarrey, Springer, Splat, Splat	Titlef, "provided" (fortef)	TINE, American Transf.	PLANE, Sprender (1998) Part (1993), Martin (1998)	Planet, Specification Codets	Frend, "general", "cont",	Parchesters", "prairies") "en",	Paul Anto American, 16, Topper's	Paral Parat Aren 18, Trapperty	Turnet, Spencerart Annet, Statute, M., Annet St.
Participante an American	Partic response in the second of the	Practic research is Professional and and the second (E)	NALL PARTY IN AN INTERVIEW	Plant, married and factors.	Contra contactor and thereefer	Sector revealed and thread its	Party designed in the start of	Party devices and been of	Forthe Conception, and Chineses (P).	Print and resident in an annual of	Party and the descent of the second s	Parts sectory is mentality
Non-José acedia L. Nolatini, Kan't RE, Natastad'i L.	The life with the fight start	Total Area (artafic S. Nordan, 1987) 201, "Nordonal", S.	Park and seen's 1. Name you's	Anial deals arriefly C., Russian, J. Saffy, ADIC: "Researcing?", SL	Part part conf. 1. "Some conf.	Toy and short it. Sciency out	The one and C. Sciences	Page State and St. L. Academican's All. Sciences 1	The loss cost, in frame, and	 Mean cards 1.405 Pers, field access 1, "Soldier, cont"). 	 Manay Sandy C.HL. Taking and Analysis of Taking Solid". 	The second second to the Northen station of the
Annual Johnson Science States	Pointer phone to be many	Product and a state of the second sec	Prevale prevalence in the second seco	Partial (Dece, Saltar, Builder, P.)	Comment of a company of a compa	Constant Plana (separ activation) Constant internet (separ-	Contract of the second se	Provide provide section provide and the section of	Porter (Artic School and School a	Provine atoms reveal investory?.	Print Parallel and a faither of the	Partial Maria Jacks' Articles (***
Constant Security March 11	Paragraphic property and	Antonio antigati antigati anti-	Print John (1947) 45. Concern (1977) Collector	Party solar professor	Country and the state	Conversion (Sector Conversion)	Print State States	Consultation of the second sec	"staming said" is more.	Ganerical entropy, C. Stollers,	Planning sear" 1, 100000.	Point Management of Contraction
Private of the part of the second	 Institution for the state of th	A second of the party into the party of the second of the party of	* constant or shapping (non-board)	 A standard and an analysis (and start)	Construction on the second sec	Proving the control of the control o	Presidents of providents of a second se	Providence and a second	Party many manufactory is	President of compart, development in the state of the series and the series of the ser	Principal of content, hubblican's Principal of distances (monormalis) when have not a first peak time of a first	 Construction of the second seco
Party chain interesting " An	Party real distribution in the second	Party man channel Physics, Summer	Panet many descention?" In Panet, and the standard status.	Party man, thread Pr. 1. Stand, marks and family results	Contraction Contraction (1)	Part out blackers, but the	Party real branchers to	Parit out that it's happy the	tracili starfe 10, frequi con-	Party cost, Strand Mr. N.	Party and Protection, S.	Party costs (Costal Party Part
"sepped", "set parts", 1.0,	Supplie, well print to the	Paragent and parts and failed	"Shapping" - Shan partial", 1.0.	Property April 1975, "Second participation of the	"Daniel", "see, property 10,	"Shared says had a set of the	"Service Ball's larth, Second 1 "Service Ball, Second Jackson", 1, 5, "Ball, String Secondar", 1, 4, 1	Property And Address of Arts	Print pre analytic (, 1)	Company's New address for the	"source", Source and Street, S	Strain P., Neur Jacob (C.),
Compare and a state of the stat	NAME AND ADDRESS OF AD	to any pre-harder Of-Office many set to any and the set of the set	sum are involved with the	Store are bench 17,30%	section reacts have very set	Adam and Adams and Adams	Low per Jacob 26, 201 marine 10,000 (per: 07,000	training and Advantage of Contract of Cont	94.401-	and an area included in the second se	a per per land-de, the	the particular and the
14, 954	100,004	12 439	1.00	14 088		-1.2.9	1.1	1. TH	- 1 I P	12.00	14.142	
				걸린 네 집	그렇게 다 나라 다	21	-1	1.2	- Man	24	- Contraction	
1		1. A.			24	1.00	8	2	The state of the s			
					1 Martin					2000		
WITHE L	SCOME AT	DATE OF	NOCASE 10	PACEART AND PACEART	Section 1.1	TANGTAL OR	ACCULT OF	Contract of the second se	NOCHELL 201 NOCH	Becchic of	No. 100 Tool of the local sector of the local	CONT OF CONTRACT
append to the first	SPECIA MALEY, PERSON applications, and man	apart, an atom	MANY TRANSPORT	agoods 15: 85/100	apore an rares	apare a second	appendix of tables	aposts 1 ALCON	again in sales	ADDRESS OF THE REPORT OF THE R	ACCOUNT OF ALL AND A	apoint to Barts
L Tarrel (1) and 1. The second	Compared and Array and Array	Di Table riterio in Di Table riterio in Prist table sementerio in Thomas	and Alternations	Contraction of the second seco	1 Agent (Carry C.	D Area of a line of a line of a second secon	Canal Ampeointee D. April (1997) 1. First rate another Print, "Management	in Age of the second se	A AND DESCRIPTION OF THE ADDRESS OF	Contract and the second states	I VERTICATION	Aller Anderson and a second se
Preside, Viewp, Mandred Marthi, D.L. Need, Jan Happer V. Hu-	ProsiP, Gauge scenario starki-	Policies, Names, Association, 62,0071 1975, Noval and Despiritures,	Priorit, "Datas incoming analysis", "Disc. News, and comparis, eds.	Party, Then, moving starts	Print, Small Monthly and C.	Printerial's print of the start of the	Pressil, Casar providing poor's - DOL, Territ, non-Despertivty	Pressilly, "Group processing start"1.	Prototi, Campo averages, et al. 1 1915, "Send year (respective);	Percept, Compromotion prophy-	Price's, "large enclose staffs".	Frends, Finner, monante parente -
Party and an eren the	Panel straking or opt that	Patrick montain start's hits	Parrie strating suggly hits,	Parent manhane score"; 101; They exclude the	Party incollege at pr ¹ (10%)	"and sold another"; i.b.	Personal and a start of the	Parameters are to an	"Provi provinsi score" y Alia. "Sant special" y Ba	Parry monthly ready may	Party Structure structure with	Party manage storts will.
Group, Appendiately on, Happenry, Group, Appendiately Const.	Filler, American States, American, Ameri American, American, Ameri	Paral and party on the second states	Resail has a poly or you, Manually Resail, "galaxies, also family the	Access, Accession and Accessio	Print Infor, 10147, 101, Parameter Particle, Sprendard's Salary,	Trul and and bit. Manual	Caref, Springer, Hu, Proparty, Caref, Springer, Volume, Name and Strategies, No.	First Later, and a try "regular", First, "pendant", forst,	Paral Married State, Vol. Parameter, Paral, Parameter's State, Non-Sec. in state 197	Provid addr. 10071 101, "repair", Party C. Tarracian's Hala?,	Forst, Sector States and Passanets	Frink Safet Lines, Inc. Supports
"taily depict" 1.	Party designed put strasted?	Party Second in Association	Party Medical II.	Practic (Lasting),	Party depicts in the second	NAT BROAD AND A STREET	Party distant, s.	Party depicted in the second of the	Praise skellars, h. Perite, Ambrid, pil, meningels	Print School in Sectors	Party designs to personal designs and the second se	Persia destares a
 Therapy Social's Lister, which young Accuracy, n. Weindoor, page 44 (31). Nonembookly, 1. 	The first water in the second	All there and a street room	Philip space and the second states	The Association of the Association of the	Post Automatic Colling	Perfly checkpilled halt. Unessed of - pre, "Milling provides to 1995. Press cancel entrol 1, Recommendation	Party State (1997) 1, 50,000, 1997.	Pain State Land, 1, 197	The state of the s	Proc. Sant Date: 1, Winner, proc.	The second state of the second state	The Application of Product starts
Technel (Marcal Aphillio), il rection "1-	"solar shout oracle states?"	Penter main simply him in a	Percent status volume, di tertoratti-	Patrice at \$1.00 crimes, \$1.0000007-	"come many prior, manyor,"	11. Notestari 6. Notas plana antas anaran-	Petter Mana -Han Averter/"-	Percent planas compartmentality	Product Annual Strength (Strength Pro-	People many speet howing?	Network States on the State of the	Petiter status scipis (cherror") Protector ambiguited and maps
No.1 Admit Charles Th.	Parts Arial Dole's Ta	Party works (calify Tr)	Pault antici (Call) *1. Princhase (Call) 1. Antices	Phone Armed Scopel Pt.	Pault acted study re-	Contractor Inspection Contractor	Para and Augo to	Party and August 19	Party Area Starts Pro	Proof and they re-	Table and Alary, the	Plat habe black the
Advertantial polaris distances of a	Polyage and an and a second second	Advisories volume streets of	Adverturation options, statement of a	Admitpublic proper developer	Contraction operation of the second s	Polycolarity, update, size cardinate and a second s	Tanonal P. Salari, Spectry's	Anterior Angeland, Streetungf 1 1 August - Cherger Der Balle Heigen	Portogette appartation destroyers	Construction, one pair of the state of the	Construction, and party statements of the	Strengthen, stager, skieserserf), Strengthen, strengthen, strengthen,
Annual control of the state of	Practic State Contraction of Con-	Parent count history 201 h.	Parent preservation of the second sec	Court start stores of the start of the	Pand scale, on "heat", in the l	Parent start shorters after A	Panel contraction of the second secon	Part cost descial to a	Party rear discovery's to	Press (real literated) Product (President discontraction in the	Parent result meaning \$7 - 5. Parent value 1 100, Aug/71 1, 100-
Company, party has been and	Particle apply in the Source's Company, Sweet, percently from	h, "adox, gar" int, "anale" Perfective, "ent, press?" U.L.	Support and state Supports	Company, and and their star-	Constrainty Social Constrainty	Tenning", "end period v 1.25.	reniera", "real provide the	Constant, Sector Sciences's	President, State and States	Paners, and such discourts	Press of the second second	response, was preserved.
terre par advector of a	train fint advice in , 1980	tren for advance.com	tradi for talentilitation	tion for admitting links	from per science of the	Long for advantitions	NAME AND ADDRESS OF A DESCRIPTION ADDRESS OF	term per alter of the	warm part spinst tot, away	trans part balant of calls	And her added to the late	tion for all \$75,2000
12-324	13/199	13.413	L1.344	55.MEL	Still HAR	54.922	117-495	EL 7/1	57.040	11.945	32.543	51.432
	3111	5	1.00	196 T			-		1911			8 C
5	121			5		Market Law	32	1.1	- 31	21	22	2
1	-					-	-				5	-
Construction of the	C. THEIR CONCERNESS	- Providence and	10-00 10 10 10 10 10 10 10 10 10 10 10 10 1	200000000	11211111111111111	- 19030000-000-000	1. S. A. A. A. A. A. A. A. A.	1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.	1.000 10000000	Sector and showing		100000000000000000000000000000000000000
states of a	division di	survey di -	peratum d3.	autor di-	and the state	station 44	Section 11	Sector 11	airtar D	Rights M.	Contraction of the	Matter El
and 17. BAR	Line over dept.	Providence and the second	Contract Annual Contractor	resolution in contract	Automatical States	Contraction of Contraction	Part and a constants	Party Loader St., On Maller	Sector in the sector in	Provide Landson Provide Landson	Print and the Party of	Automotive reality
r Server Collect, Springer	T THE CONTRACTOR	Part Massimon	And WEREINFORM	and analysis of the	Alast Statistic Works	A State State (Same Same Same Same Same Same Same Same	Charles and Charle	A San Press Area	A Service Service	Const control of	A Say Annual Table	Characteria de
Colle, "Shap, and the state	and street women of the la function has	"THE AND A DESCRIPTION OF THE ADDRESS OF THE ADDRES	Proof. Charles and and a start	Process Compared to the Constant of States	Print, "Print and internet, in	Proc.Pt Single, and String at arts	Profile, "Chain an orange of the	Print Print Hard States and States	"Toth, "into the share"; h. "modes"; "toth, "into the share"; in	Private and a start of the	Print and an arrival and a state of the	"CONTRACTOR CONTRACTOR INCOMENTS
	still, "stall and respondy its	ATTING THE REPORT OF AN ADDRESS	The Party and Manual Provide State									
the print territory in the	and, "man and managers" its. Prince and resident's to be specific annihility stracts and	New consisting concerning	New you avoid to be	Compared and the second of the state	Conclusion of the second secon	New party second of a log	Contraction are intervented.	Non-Josh Marine, C.C.	Sortheodory, end of the	Print and second reaction	Table and supply the	True pairs and the rate
New york Antibal's 1.1, New Yorking Anton Maj New Yorking Anton State New York Anton's 2011 Magnetic Space State Antonio State Space State Antonio State Space State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State State Sta	 Bill, "Proc. and Property St. Prints and eventher," I. S. K. Peret Section States," and Proc. Section Sci. Proc. Proc. Proc. Sci. Sci. Proc. Proc. Proc. Sci. Sci. Proc. Proc. 	 Construction of the second seco	Pure and events's to a subcarding units' total has more to be here to a subch total subch, typestory total,	Company and a control of the second s	Clear Lord Account, 1, 1, 1, - And Account of States of States - State States of States - States of	 New party and doing to be recommended and the base processing starter way base processing to be been based to be based on the record, replacement record, 	Close and American's vice Sector constant of the Next American score's state Next American's State Sector Score's State Score's Spectration's States	Concernit Anima's C.S. New Anima's Anima's May Part Anima's Anima's May Part Anima's Anima's May Provide Part Anima's Anima's Market Part Anima's Anima's Anima's Part Anima's Anima's Anima's Part Anima's Anima's Anima's Anima's Part Anima's Anima's Anima's Anima's Anima's Part Anima's	Comparison and the Company of the Co	Production (control of the control o	Color, Josepher Y., Yu., Anna Dirich Markatari, Y.A., Neuroperstation, and the Second Science of Sciences Science of Sciences and Sciences Science of Sciences and Sciences and Sciences Sciences of Sciences and Sciences and Sciences Sciences of Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences and Sciences	Proj ne user's r.e. Proj (ral sector's r.e. Peroj estato daris (m) Peroj estato"s (n) Peroj estato"s (n) Peroj sector's (n)
New york: Animar's 1, 1, New York: Animar's 1, 2015 New York: Animar's 1, 2	H. Shi, and manifer S. Programs and models in a second second state of an based and second sec	Construction of the second sec	Comparison and Section 2.1 (A) Provide and Section 2.1	And a second sec	Characterization of the second	Ling and evolution of the based operating strategy way from the second of the based operating operating strategy based on the second of the based operating of the based operating of the based operating operating operating the second operating operating operating the second operating operating operating the second operating operating operating the second operating operating operating operating the second operating operating operating operating the second operating operating operating operating operating the second operating operating operating operating operating operating the second operating operating operating operating operating operating the second operating op	Constant, and Articlet's via Anti-South and Articlet and Net American Constant and Net Articlet and Articlet Articlet and Articlet's Antipic Articlet and Articlet's Antipic Articlet and Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articlet Articl	Print, and Arrison's Con- tract Annual and Con- tract Annual Con- Con- Con- Con- Con- Con- Con- Con-	Characterization and the state State and the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the	 Principano, José Marcin, M., C., C., Constanti, S., Constanti, S. (2010). Principano, C. (2010). Pr	 Data and second to the product of the	Terral Jan, Mathell's H., Berral (2014) Mathelline T. et al. Perrol (2014) Mathelline T. et al. Perrol (2014) Annual Annual Terral Perrol (2014) Annual Annual Terral Science (2014) Physics (2014) Physics (2014) Physics (2014) Phys. (2014) Physics (2014) Physics (2014) Phys. (2014) Phys. (2014) Phys. (2014) Phys. (2014) Phys. (2014) Phys. (2014) Phys. (2014
New years, Annexe S. L., San Annexe S. S. S. San A. San Annexe S. S. San A. San Annexe S. S. San Annexe S. San A. San A. San A.	Bill, Star, Star, San, Sangari, G., Yang, Jan Kamada, Yu, S. Y. Yang, Kanang Yu, S. Wang, Kanang Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu, Yu,	Construction of the constr	 C. S. Wei, M. S. Martin, Y. L. S. Wei, S. Wei, S. Wei, S. Wei, S. W. S. Wei, S. W. S. Wei, S. Wei	Calify and Sector's 1.1.2. Sector and an end of the sector and th	Proceeding and the second seco	Language and an and the state and particle of the state of the based of the state of the state of the based of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state o	The parts and the second second second second secon	Array and an end of the production party of the backback of the backback backback of the backbackback of the backbackback of the backbackback of the backbackbackback backbackbackbackbackbackback backbackbackbackbackbackbackbackback backbackbackbackbackbackbackbackbackback	Comparing services of the Comparing service of	Production and Production of the Production of t	Data Jan, ang Ph., Yu., Yang Jini, anitan'i Yu., Yang Yu., anitan'i Yu., Yu., anitan'i S., Yuli Jan, Yu., Yu., "Spanner Core: Construction (Contr.) Nicas, anitan'i Yu., "Nicas, anitan'i Yu., "Anito, Anitana (Jan, Amatonia') "Anito, Anitana (Jan, Amatonia') "Science, Anitana (Jan, Amatonia')	The part of the second
Name and Antonio C. C. A. San Chronic and S. Mariana Name and S. S. San	Bill, Strat. Sci. Sci. 2010. 11 (1) Thing, and scientific 3 (1) (1) Sci. 2010. 11 (1) (1) (1) (1) (1) (1) (1) (1) (1)	Construction of the second second second second sec	 Development of the second secon	Calls your benefit of the Anti-Antoneous and the Network and Network and N	Press Arm (with the first sec- which and the press sec- band press of the sec- temport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the sectemport of the	Note and products of the second and design products of the second and design products of the second and design products of the second second second second second research second second second second second research second second second second second second research second second second second second second second research second second second second second second second research second second second second second second second second research second second second second second second second research second second second second second second second second research second secon	Dec arts, sender (1993) - And Canada (2014) - And	Decemperation (accessed on the procession (accession	Determination of the second se	Device and second structures and second performance of a structure and second structures and structures between a structures and structures between a structures and structures of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structure of the structu	Mathaine and Sangara and Sangar Sangara and Sangara	Processing and Processing States Processing States and Processing States Processing Sta

My cross-validation "command center"



My cross-validation "command center"





My cross-validation "command center"



Monitor and visualize the loss curve







Monitor and visualize the accuracy:



Track the ratio of weight updates / weight magnitudes:

```
function gradient_descent(model, lr)
    w_scale = torch.norm(model.W:view(model.W:nElement()), 2, 1)
    update_scale = torch.norm(lr * model.gradW:view(model.gradW:nElement()), 2, 1)
    model.W = model.W + lr * model.gradW
    model.b = model.b + lr * model.gradb
    print(update_scale/w_scale) -- Want ~1e-3
end
```

ratio between the values and updates: ~ 0.0002 / 0.02 = 0.01 (about okay) want this to be somewhere around 0.001 or so

Visualize Activations

• Visualize features (feature maps need to be uncorrelated) and have high variance.



hidden unit

Good training: hidden units are sparse across samples and across features.

hidden unit

Bad training: many hidden units ignore the input and/or exhibit strong correlations.

Visualize (initial) Convolution Layer Weights

• Visualize features (feature maps need to be uncorrelated) and have high variance.



Good training: learned filters exhibit structure and are uncorrelated.

Visualize Linear Layer (Fully-Connected) Weights

- Visualization of Linear layer weights for some networks
- It has a banded structure repeated 28 times (Why?!) Hint: Images are 28x28
- Thus, looking at the weights we get some intuition

