CloudCV:
Large-Scale Computer Vision on the Cloud

http://cloudcv.org/

Dhruv Batra

Virginia Tech

Lead: Machine Learning & Perception Group @ VT

CloudCV Team:

Harsh Agrawal

Clint Solomon

Yash Goyal

Outline

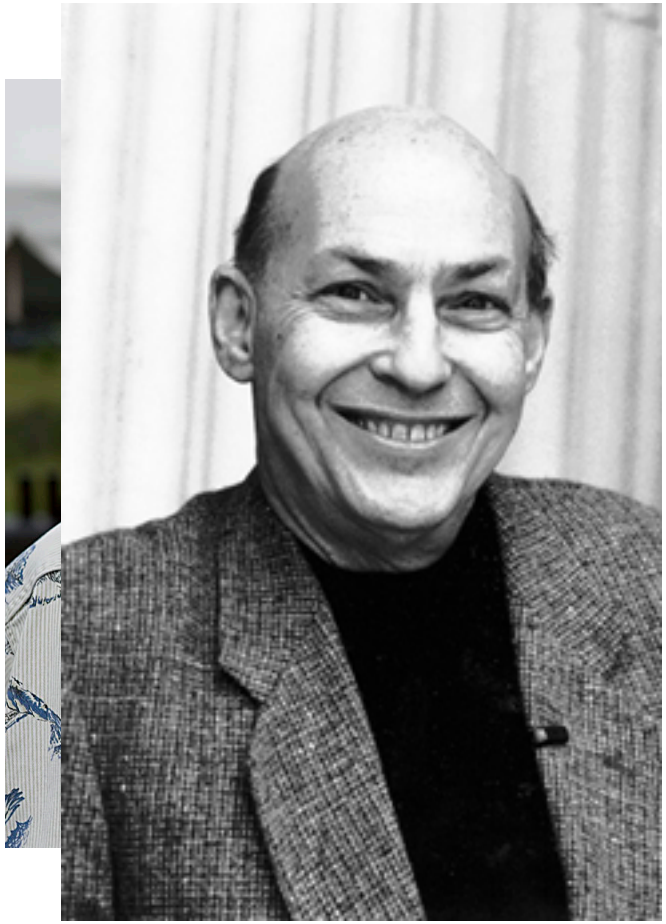
- Historical context about Computer Vision
- CloudCV
  - A mix of
    - Research in my group
    - Deployment and demos at cloudcv.org

Computer Vision: Making Machines See

"Color College Avenue", Blacksburg, VA, May 2012

Slide credit: Devi Parikh

Computer Vision

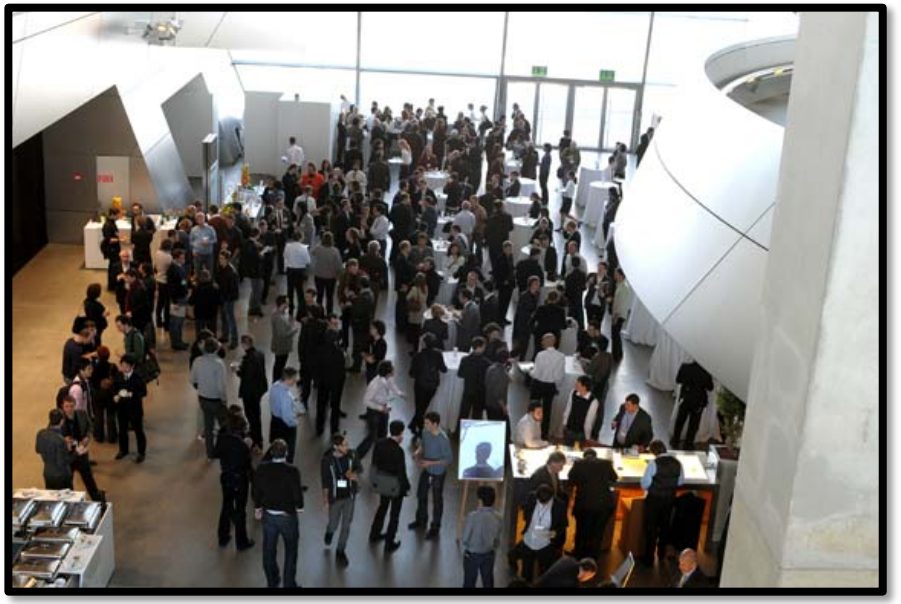


"spend the summer linking a camera to a computer and getting the computer to describe what it saw"

- Marvin Minsky (1966), MIT

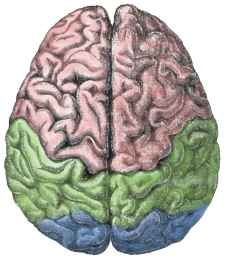
… 49 years later

Computer Vision



OR

Vision is HARD!



Datasets and computer vision

UIUC Cars (2004)

S. Agarwal, A. Awan, D. Roth

CMU/VASC Faces (1998)

H. Rowley, S. Baluja, T. Kanade

FERET Faces (1998)

P. Phillips, H. Wechsler, J. Huang, P. Raus

COIL Objects (1996)

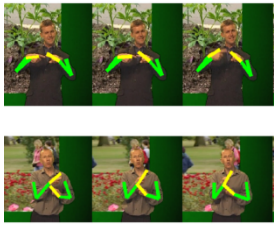
S. Nene, S. Nayar, H. Murase

MNIST digits (1998-10)

Y LeCun & C. Cortes

KTH human action (2004)

I. Leptev & B. Caputo

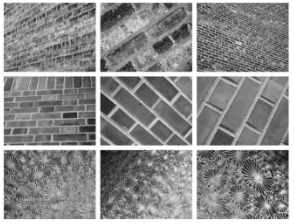


Sign Language (2008)

P. Buehler, M. Everingham, A. Zisserman

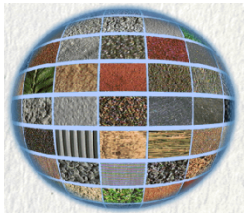
Segmentation (2001)

D. Martin, C. Fowlkes, D. Tal, J. Malik.



3D Textures (2005)

S. Lazebnik, C. Schmid, J. Ponce

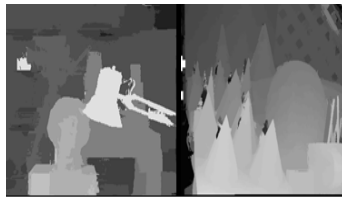


CuRRET Textures (1999)

K. Dana B. Van Ginneken S. Nayar J. Koenderink

CAVIAR Tracking (2005)

R. Fisher, J. Santos-Victor J. Crowley



Middlebury Stereo (2002)

D. Scharstein R. Szeliski

Backpack

Flute

Strawberry

Traffic light

Backpack

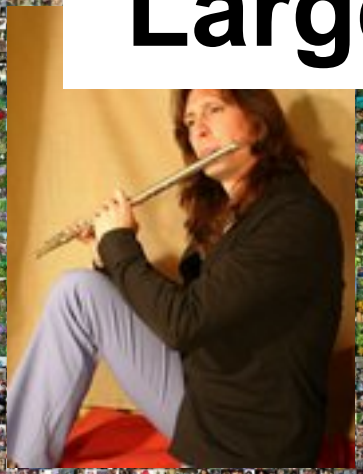
Matchstick

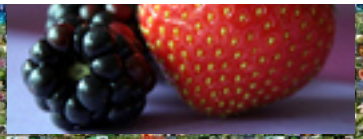
Bathing cap

Sea lion

Racket

Large-scale recognition





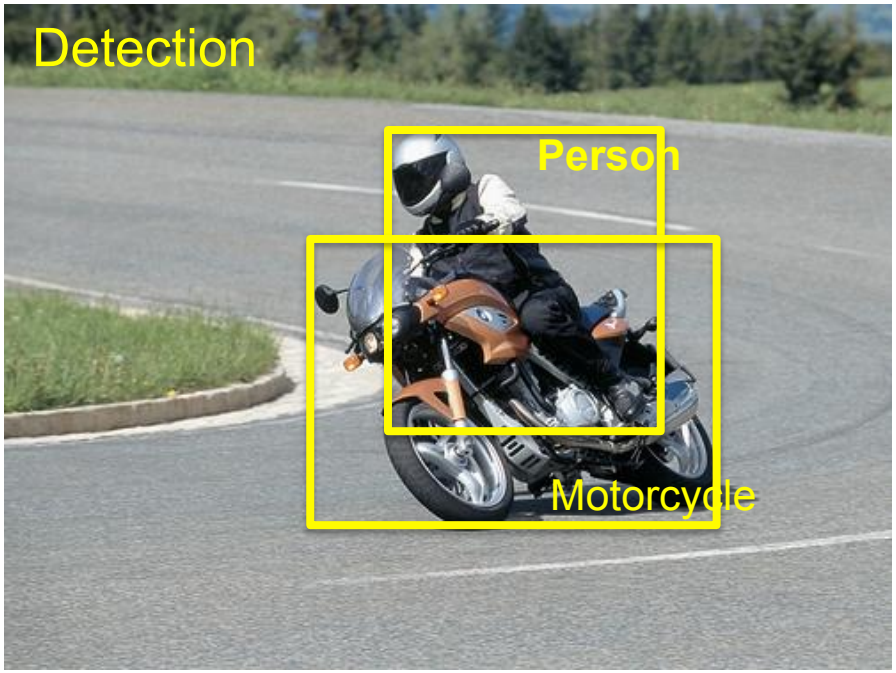
PASCAL VOC 2005-2012

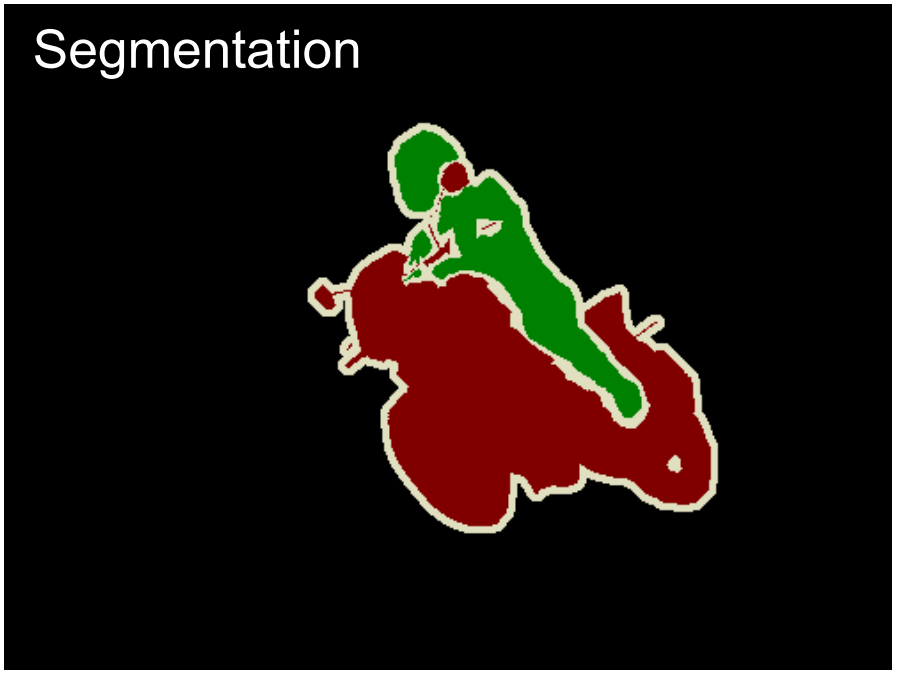
Everingham, Van Gool, Williams, Winn and Zisserman.

The PASCAL Visual Object Classes (VOC) Challenge. IJCV 2010.

20 object classes **22,591 images**

Classification: person, motorcycle





Action: riding bicycle

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

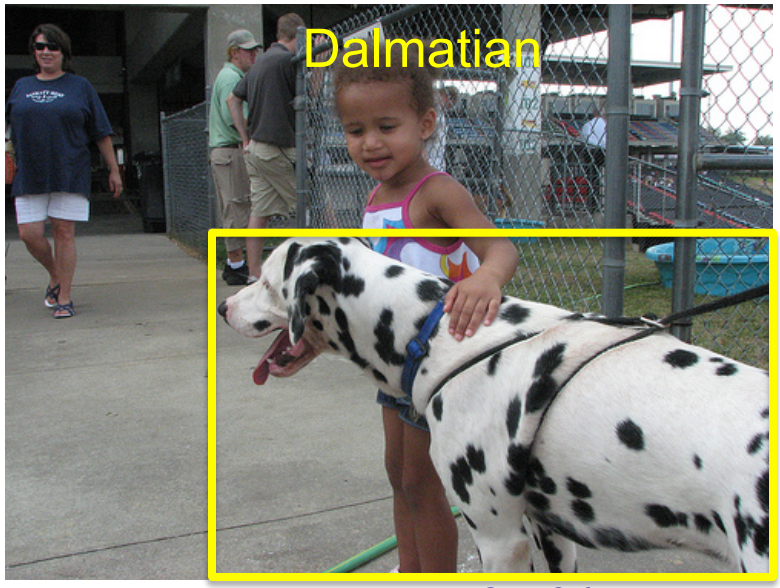
~~20 object classes 22,591 images~~

Classification:

1000 object classes 1.4M/50k/100k images

Detection:

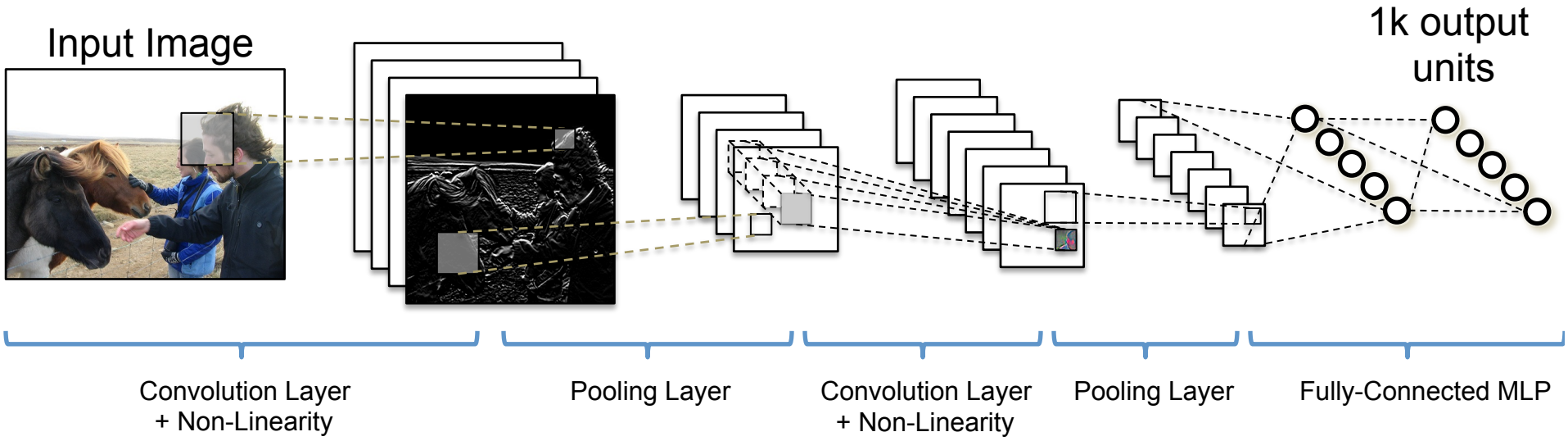
200 object classes 400k/20k/40k images



http://image-net.org/challenges/LSVRC/{2010,…,2014}

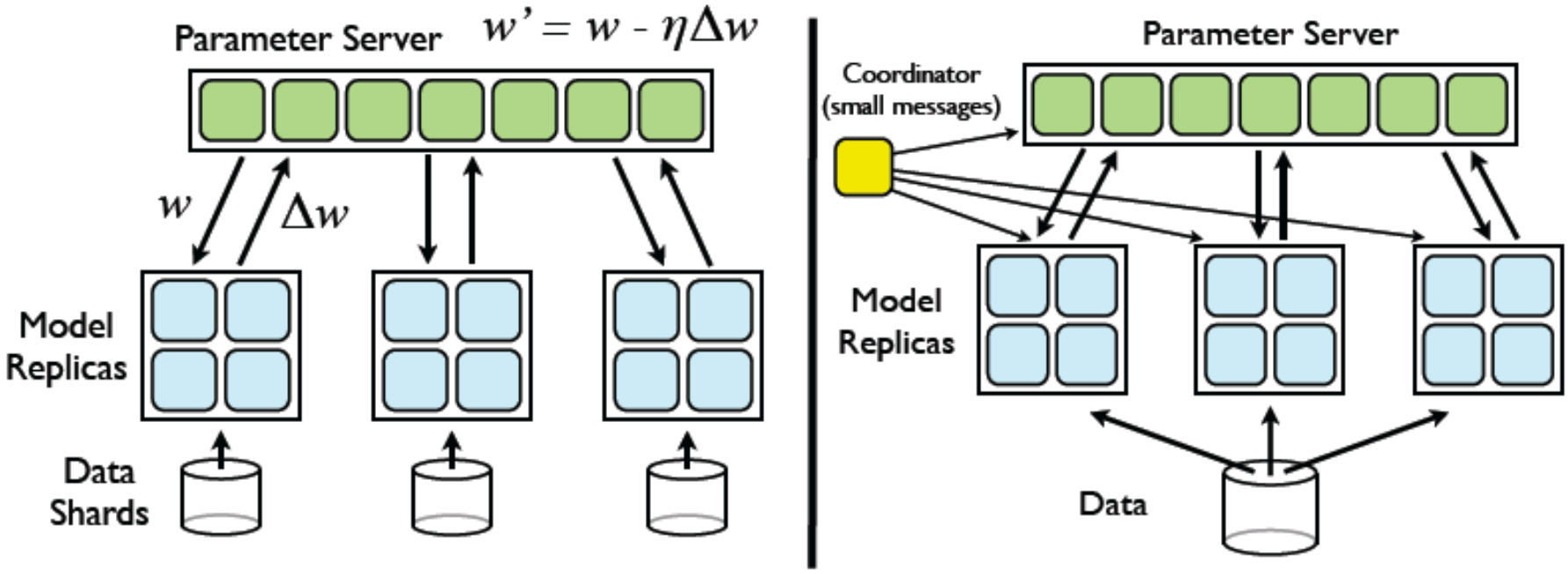
Data Enabling Richer Models

- [Krizhevsky et al. NIPS12, Donahue ICML14]
  - 54 million parameters
  - Trained on 1.4M images in ImageNet



Data Enabling Richer Models

- DistBelief [Dean et al. NIPS12]



Data Enabling Richer Models

- [Le et al. ICML12]
  - 2,000 machines / 32,000 cores for 1 week

- DistBelief [Dean et al. NIPS12]
  - 16 million images and 21k categories
  - 1.7 Billion parameters
  - 12,000 cores

Challenges

- Big data is an enabler and an isolator!
- All researchers repeatedly solving the same problems

  - Build and maintain a cluster
    - Job scheduler (PBS, Torque)
    - Distributed storage (Hadoop FS)

    → Logistical

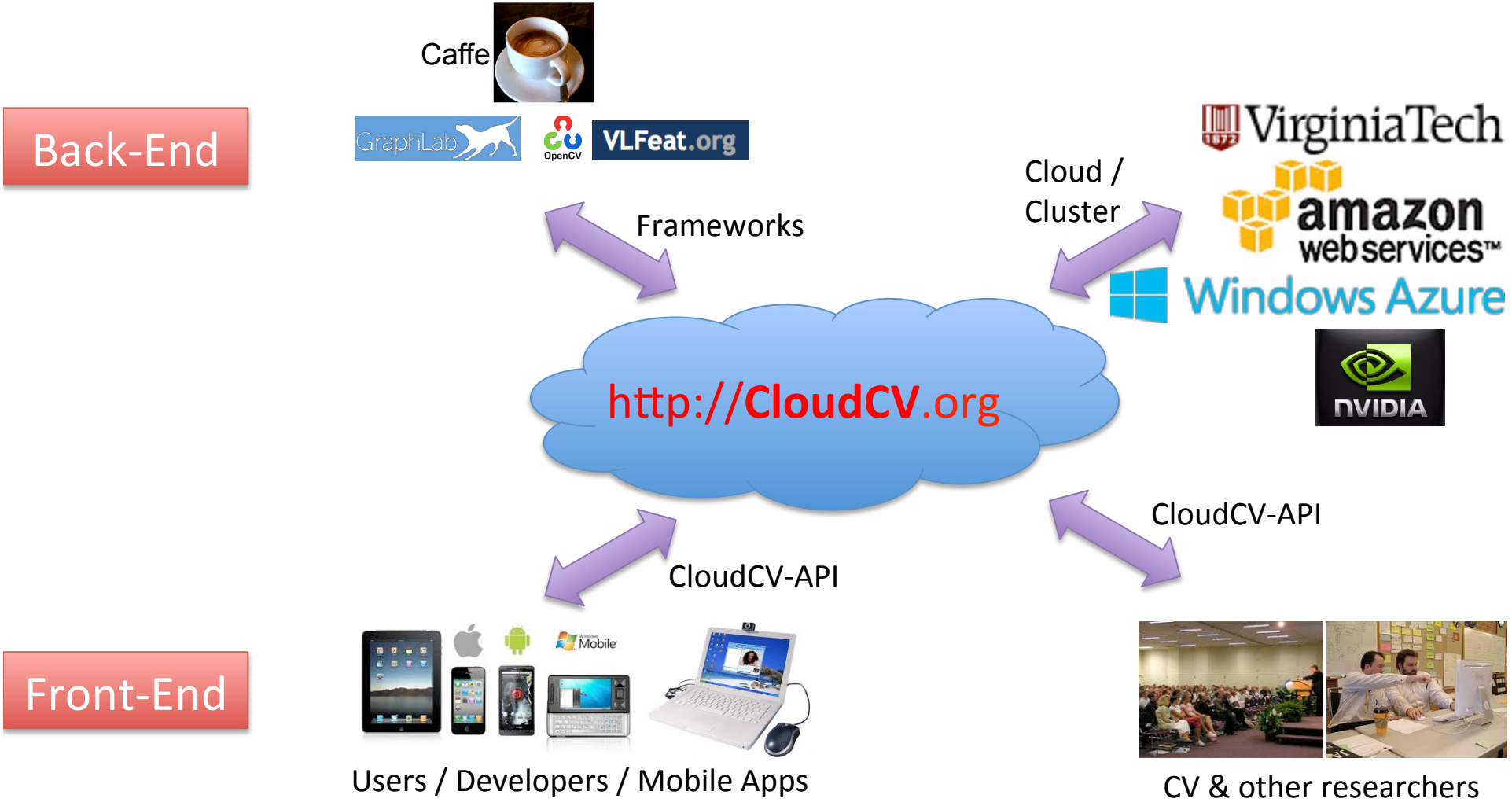
  - Scale vision algorithms
    - Identify model/data parallelism
    - Design & implement multi-threaded vision primitives

    → Computer Vision

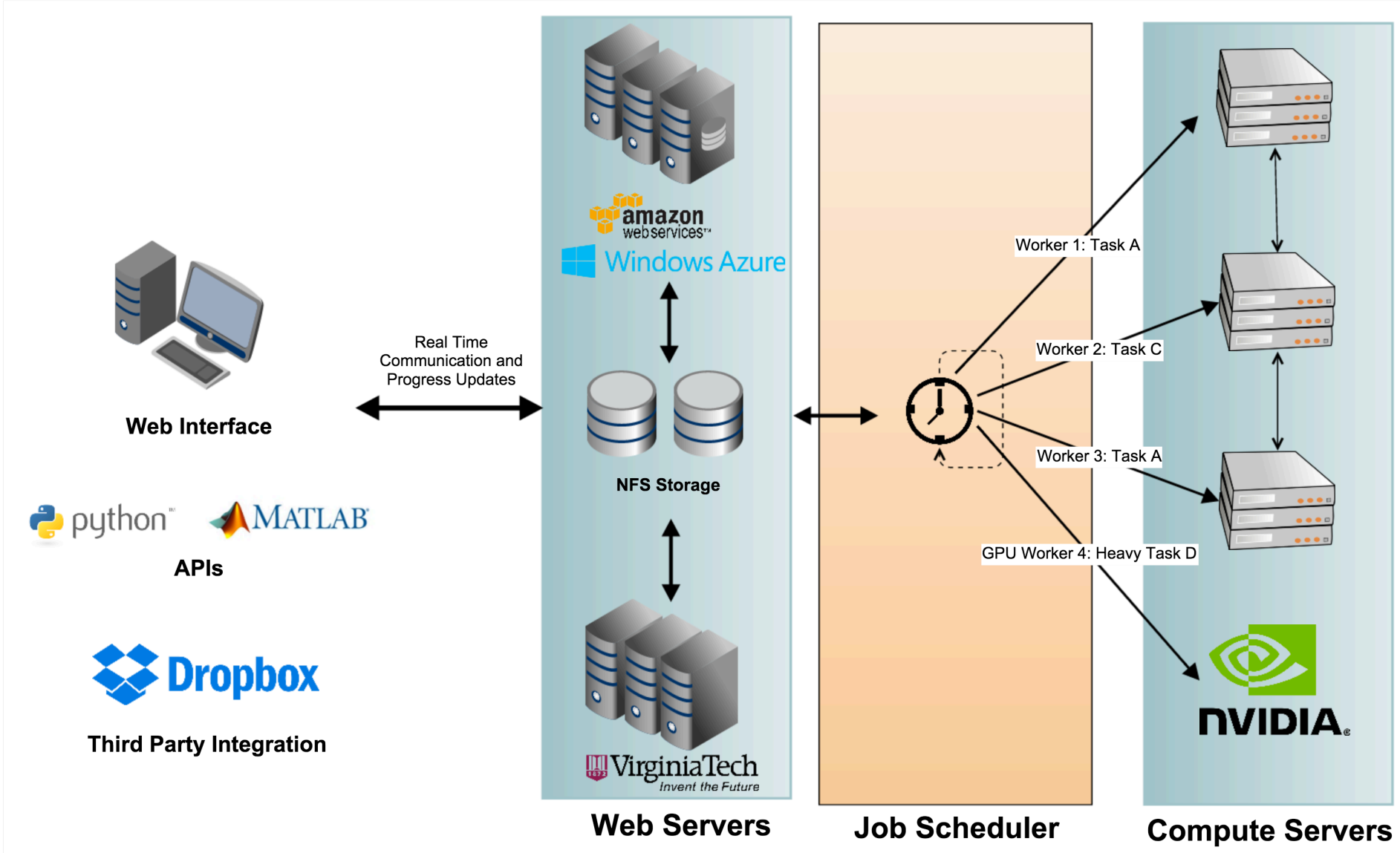
  - Distributed computing
    - Implement mechanisms to avoid race conditions & dead-locks
    - Ensure data consistency, locking, good scheduling

    → Distributed Computing

CloudCV


Back-End
Caffe
GraphLab
OpenCV
VLFeat.org
Frameworks
Cloud /
Cluster
VirginiaTech
amazon
web services™
Windows Azure
NVIDIA
http://CloudCV.org
CloudCV-API
CloudCV-API
Front-End
Mobile
Users / Developers / Mobile Apps
CV & other researchers

CloudCV: Architecture



CloudCV: Big Picture

- Goal: For developers
  - Reduced barrier to entry
  - Democratize Computer Vision
- Goal: For researchers
  - Easy comparison to baselines
  - Access to state-of-art techniques “off-the-shelf”
- Mini-steps
  - What we have today
  - A few algorithms
  - A few ways to reach CloudCV
  - Where we are headed

CloudCV

- Demo 1
  - Support for ImageNet Challenge
- Demo 2
  - Image Classification
- Demo 3
  - Training a new classifier for your categories
- Demo 4
  - Finding Important People in Images
- Demo 5
  - GigaPixel Image Stitching

“Demo” 1

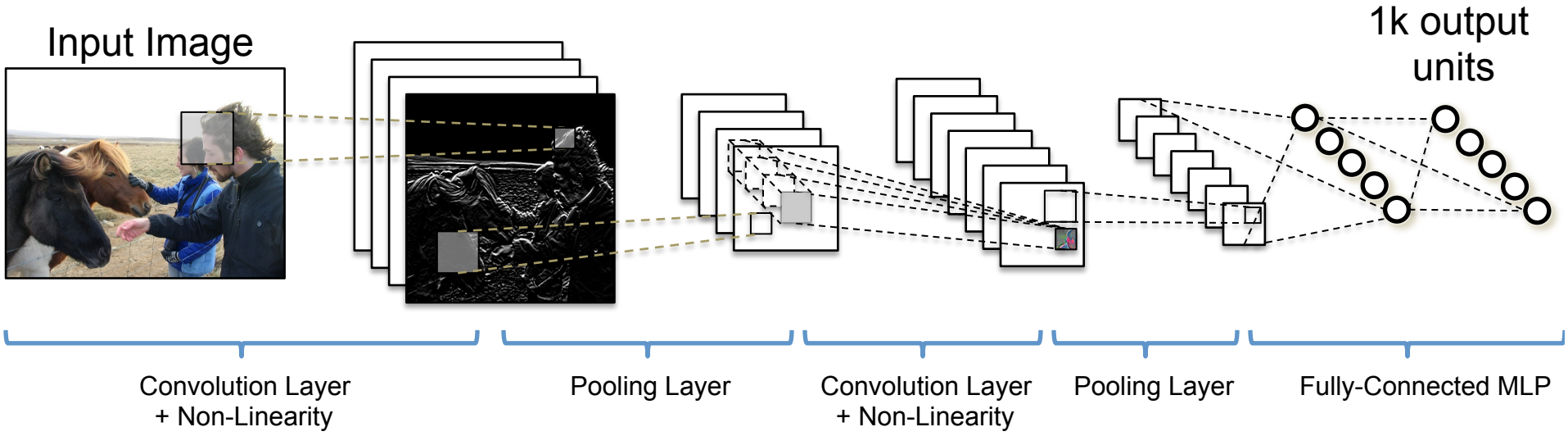
- ImageNet Challenge (ILSVRC13)
  - Training: 1.4 million
  - Val: 50k
  - Test: 100k
- Features
  - 16 “industry standard”
    - DeCAF, GIST, HOG2x2, Dense/Sparse SIFT, LBP, Self-Similarity …
- Webpage
  - http://cloudcv.org/objdetect/#features
- Total: 400 GB, *19 months or 1.5 years of CPU-time*

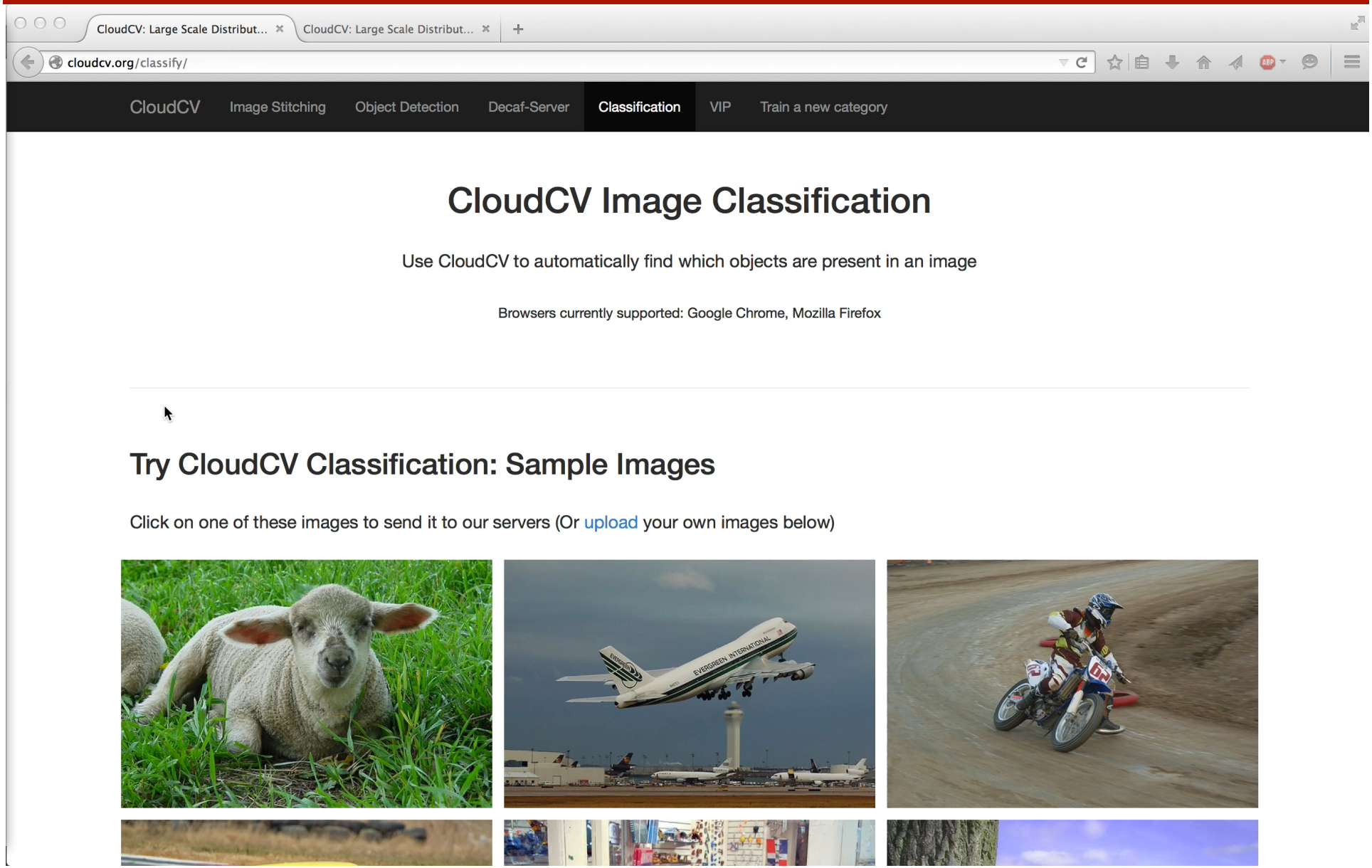
CloudCV

- Demo 1
  - Support for ImageNet Challenge
- Demo 2
  - Image Classification
- Demo 3
  - Training a new classifier for your categories
- Demo 4
  - Finding Important People in Images
- Demo 5
  - GigaPixel Image Stitching

Demo 2

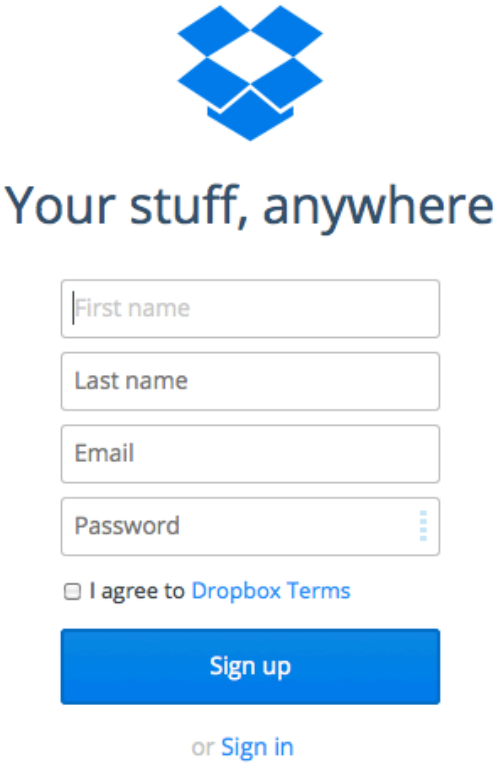
- [Krizhevsky et al. NIPS12, Donahue ICML14]
  - Trained on 1.4M images in ImageNet
  - 1000 categories
  - Available in Caffe framework from BVLC
  - http://cloudcv.org/classify/




CloudCV: Large Scale Distribut...
CloudCV: Large Scale Distribut...
cloudcv.org/classify/
CloudCV
Image Stitching
Object Detection
Decaf-Server
Classification
VIP
Train a new category
CloudCV Image Classification
Use CloudCV to automatically find which objects are present in an image
Browsers currently supported: Google Chrome, Mozilla Firefox
Try CloudCV Classification: Sample Images
Click on one of these images to send it to our servers (Or upload your own images below)
EVERGREEN INTERNATIONAL

Demo 2

- Drop-box integration
  - Files can live on dropbox
  - http://cloudcv.org/decaf-server/



Saves – Dropbox | CloudCV: Large Scale Distribut...

cloudcv.org/decaf-server/

CloudCV | Image Stitching | Object Detection | Decaf-Server | Classification | VIP | Train a new category

CloudCV Decaf Server

Need Decaf features? Don't have GPUs? Don't want to deal with installing Caffe?

CloudCV provides dedicated servers to extract Decaf features.

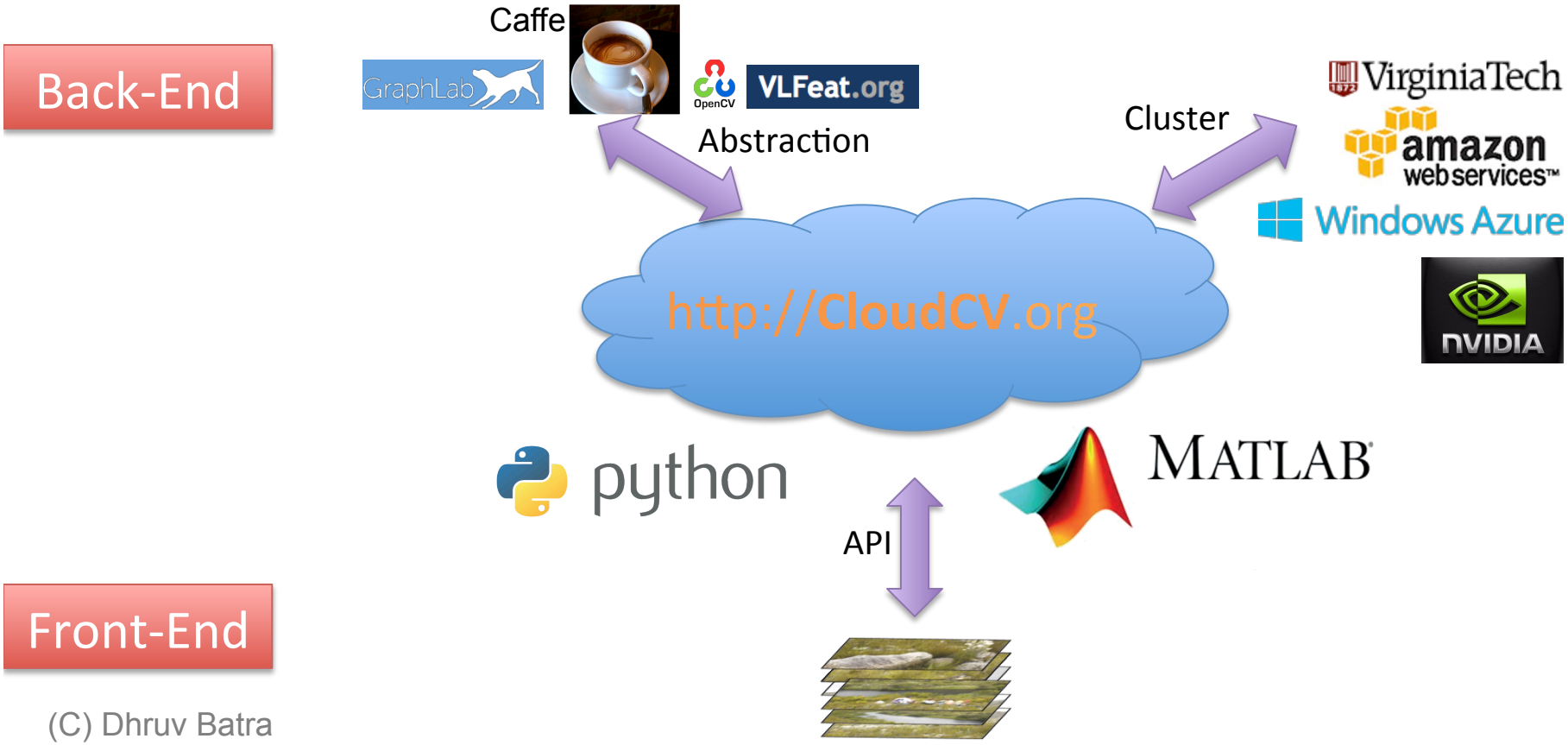
Browsers currently supported: Google Chrome, Mozilla Firefox

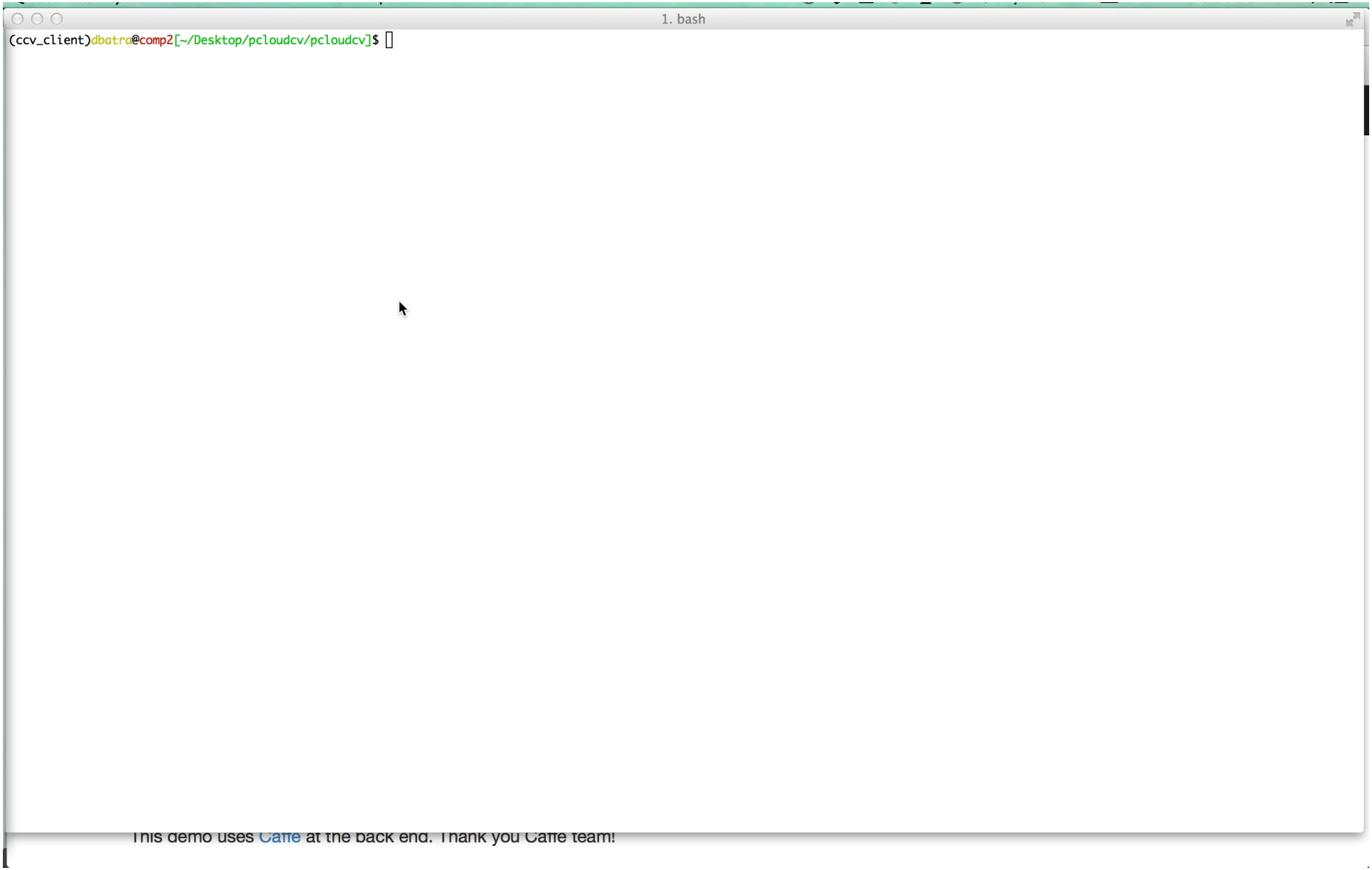
Try Decaf-Server: Sample Images

Click on one of these images to send it to our servers (Or upload your own images below)

Demo 2

- How about if you want to write code?
  - Python-API: https://github.com/batra-mlp-lab/pcloudcv
    - "python run.py myconfig.json –nologin"
  - Matlab-API: https://github.com/batra-mlp-lab/mcloudcv




1. bash
(ccv_client)dbatra@comp2[~/Desktop/pcloudcv/pcloudcv]$
This demo uses Caffe at the back end. Thank you Caffe team!

CloudCV

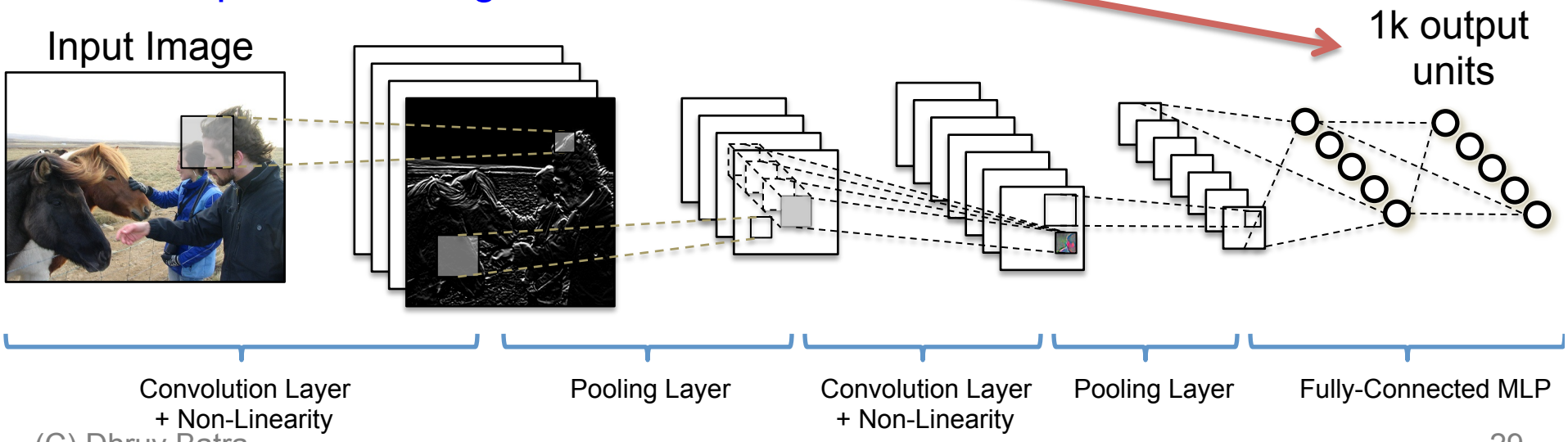
- Demo 1
  - Support for ImageNet Challenge
- Demo 2
  - Image Classification
- Demo 3
  - Training a new classifier for your categories
- Demo 4
  - Finding Important People in Images
- Demo 5
  - GigaPixel Image Stitching

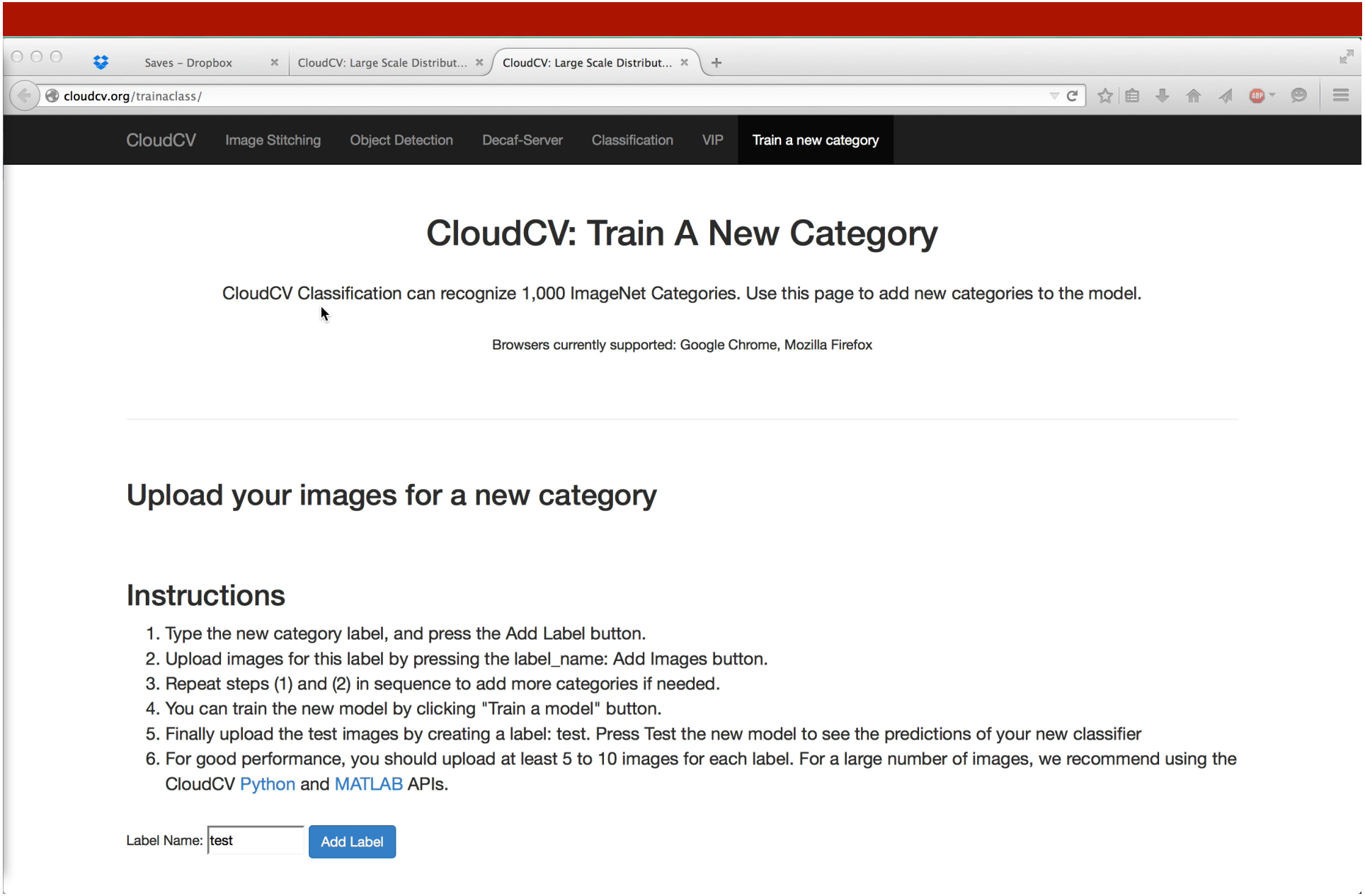
Demo 3

- [Krizhevsky et al. NIPS12, Donahue ICML14]
  - Trained on 1.4M images in ImageNet
  - 1000 categories
  - Available in Caffe framework from BVLC

How about adding a 1001th category?
Your company logo classifier?
In a few seconds, not weeks?

http://cloudcv.org/trainaclass/




Saves – Dropbox
CloudCV: Large Scale Distribut...
CloudCV: Large Scale Distribut...
cloudcv.org/trainaclass/
CloudCV
Image Stitching
Object Detection
Decaf-Server
Classification
VIP
Train a new category
CloudCV: Train A New Category
CloudCV Classification can recognize 1,000 ImageNet Categories. Use this page to add new categories to the model.
Browsers currently supported: Google Chrome, Mozilla Firefox
Upload your images for a new category
Instructions
1. Type the new category label, and press the Add Label button.
2. Upload images for this label by pressing the label_name: Add Images button.
3. Repeat steps (1) and (2) in sequence to add more categories if needed.
4. You can train the new model by clicking "Train a model" button.
5. Finally upload the test images by creating a label: test. Press Test the new model to see the predictions of your new classifier
6. For good performance, you should upload at least 5 to 10 images for each label. For a large number of images, we recommend using the CloudCV Python and MATLAB APIs.
Label Name: test
Add Label

CloudCV

- Demo 1
  - Support for ImageNet Challenge
- Demo 2
  - Image Classification
- Demo 3
  - Training a new classifier for your categories
- Demo 4
  - Finding Important People in Images
- Demo 5
  - GigaPixel Image Stitching

Who is the most important person in the photo?

Why is this useful?

- Better image descriptions
- Automatic photo cropping

Two people walking past a crowd

Why is this useful?

- Better image descriptions
- Automatic photo cropping
- Sort consumer photos

How do we do this?

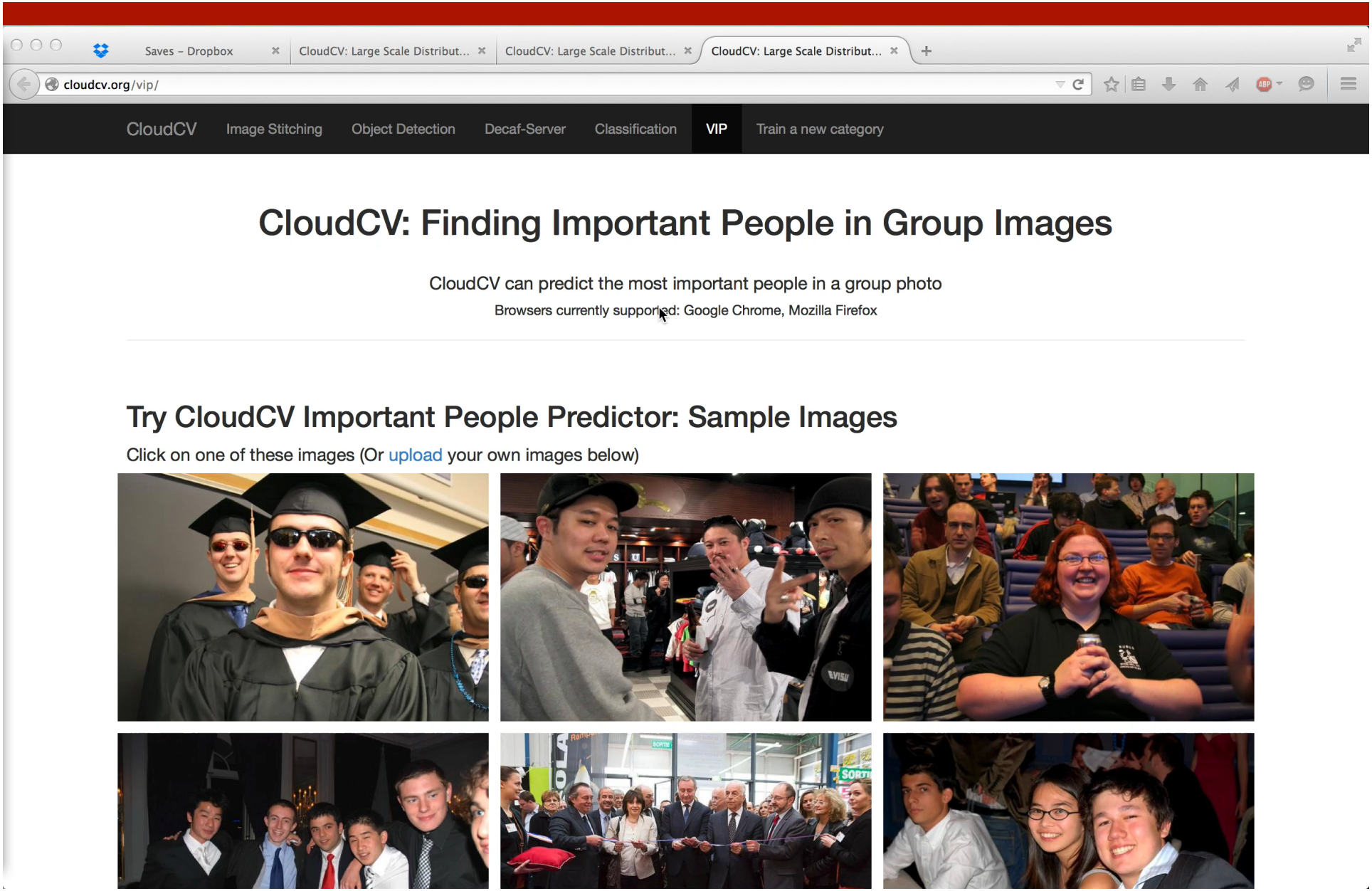
- Collect a large dataset
  - VT Person Importance Dataset
  - Images scraped from Flickr
  - Annotations using Mechanical Turk
- For each face measure:
  - Distance from center
  - Scale
  - Sharpness
  - Face Pose
  - Face Occlusion
- Train a relative importance predictor

Results

- http://cloudcv.org/vip/

Method	Accuracy
Our Approach	**78.91%**
Center Baseline	68.46%
Scale Baseline	67.86%
Sharpness Baseline	71.03%

- Technical Details:
  - VIP: Finding Important People in Images
  - Clint S. Mathialagan, Andrew C. Gallagher, Dhruv Batra
  - http://arxiv.org/abs/1502.05678


Saves - Dropbox
CloudCV: Large Scale Distribut...
CloudCV: Large Scale Distribut...
CloudCV: Large Scale Distribut...
cloudcv.org/vip/
CloudCV
Image Stitching
Object Detection
Decaf-Server
Classification
VIP
Train a new category
CloudCV: Finding Important People in Group Images
CloudCV can predict the most important people in a group photo
Browsers currently supported: Google Chrome, Mozilla Firefox
Try CloudCV Important People Predictor: Sample Images
Click on one of these images (Or upload your own images below)

CloudCV

- Demo 1
  - Support for ImageNet Challenge
- Demo 2
  - Image Classification
- Demo 3
  - Training a new classifier for your categories
- Demo 4
  - Finding Important People in Images
- Demo 5
  - GigaPixel Image Stitching

Where is CloudCV headed?

- Back-end
  - Open model for contributing code
- Dynamic Database
  - If “familiar” image, we can get you results without computing
  - If new image, we’ll cache the results for the next person
- Lots of challenges unsolved
  - Bandwidth, optimal compression
  - Computation on front end vs back end
  - Compressions on front end that bound performance?
    - Coresets, summarization, etc

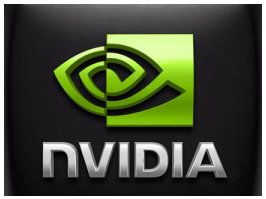
Where is CloudCV headed?

- Long way to go
  - But we think this is exciting!
- Think about the first APIs for
  - User authentication, Credit-card processing
  - Search, Maps, Twitter feeds, …
- We want to do that for the scientific research and development community.

Acknowledgements

- Collaborator and Mentor
  - Carlos Guestrin (UW / Graphlab / Dato)

- Sponsors



CloudCV Team:

Harsh Agrawal

Clint Solomon

Yash Goyal

Thanks!