

Agenda

- > Introduction to DIGITS
- > Install and start DIGITS sever
- > CREATING DATASETS
- > TRAIN A NETWORK
- > MODIFY YOUR NETWORK
- Testing/Classification with DIGITS
- > DIGITS Features at a Glance
- > References

DIGITS makes it way easier to design the best network for the job. The DIGITS interface makes it super easy to track key diagnostics during training. The field will definitely benefit from having tools like this for configuration and introspection"



Interactive Deep Learning GPU Training System

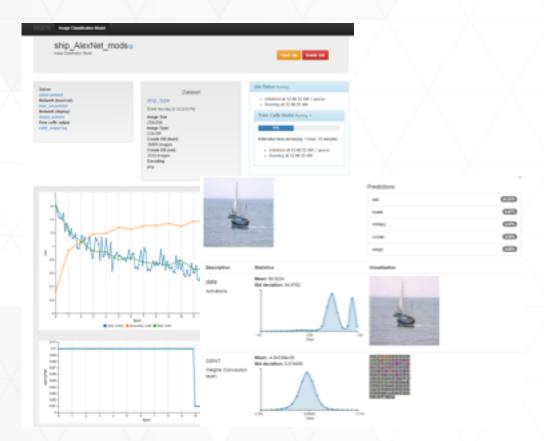
Process Data

Configure DNN

Monitor Progress

Visualization

Who is DIGITS for?



Data Scientists & Researchers:

- Quickly design the best deep neural network (DNN) for your data
- Monitor DNN training quality in realtime
- Manage training of many DNNs in parallel on multi-GPU systems, and multi-GPU training



DIGITS

Deep Learning GPU Training System

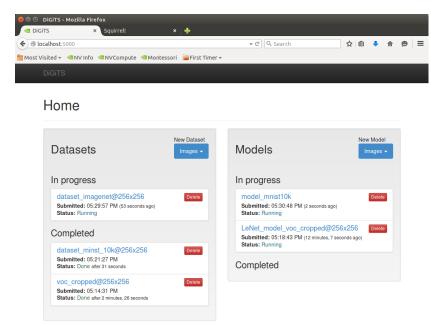
Available at http://developer.nvidia.com/digits

Free to use, Source Code available at Github, latest branch v3.0

https://github.com/NVIDIA/DIGITS

Current release supports classification on images

Future versions: More problem types and data formats (video, speech)



DIGITS

Key Features

Visualize DNN topology and how training data activates your network

Manage training of many DNNs in parallel on multi-GPU systems

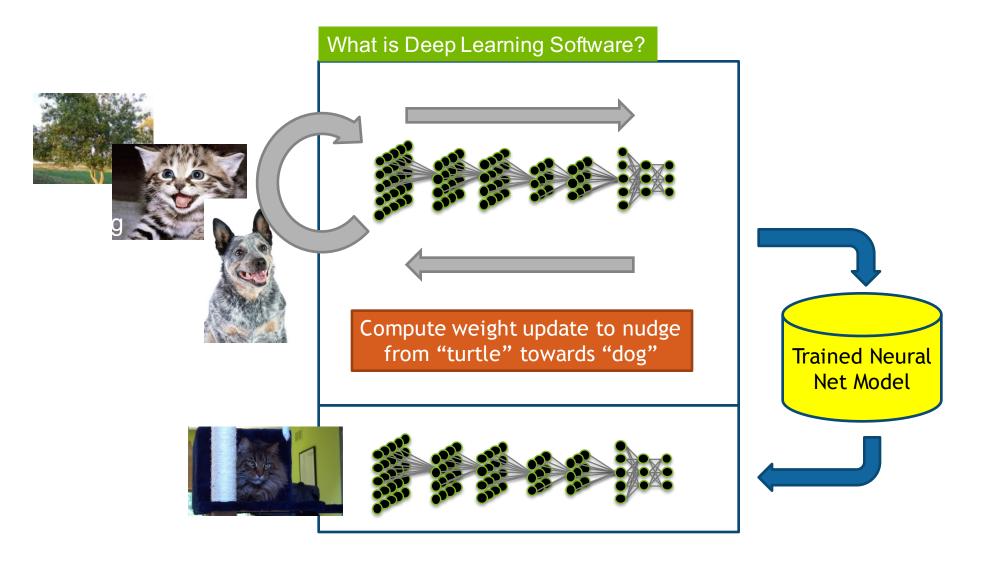
Simple setup and launch

Import a wide variety of image formats and sources

Monitor network training in real-time

Open source, so DIGITS can be customized and extended as needed





Deep Learning

Steps with DIGITS

Creating a Dataset

Define the Network or use existing

Choose a given Framework

Selecting a preconfigured ("standard") network - LeNet, AlexNet, GoogleNet

Previous network

Custom network

Training a Model

Classification





DIGITS Installation

Two Ways

Installation with Pre-Built Packages

Deb packages are provided for easy installation on Ubuntu 14.04.

Packages are provided for major releases, but not minor ones (i.e. v3.0 and v4.0, but not v3.1).

Download the web installer - https://developer.nvidia.com/digits

DIGITS Installation

Two Ways

Installation with Source

Get the latest branch from GitHub

Latest features but carries risk of untested features

Build NVIDIA Caffe branch - https://github.com/NVIDIA/caffe

Download DIGITS from github -https://github.com/NVIDIA/DIGITS

DIGITS Installation

HW and SW

Hardware/software recommendations

GPU(s) with compute capabilities >= 3.0 for cuDNN support

Ubuntu 14.04, CentOS

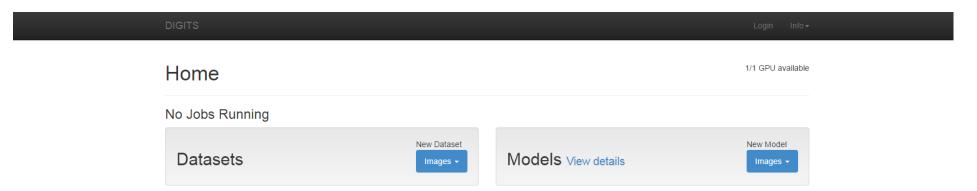
Main Console

Start DIGITS by

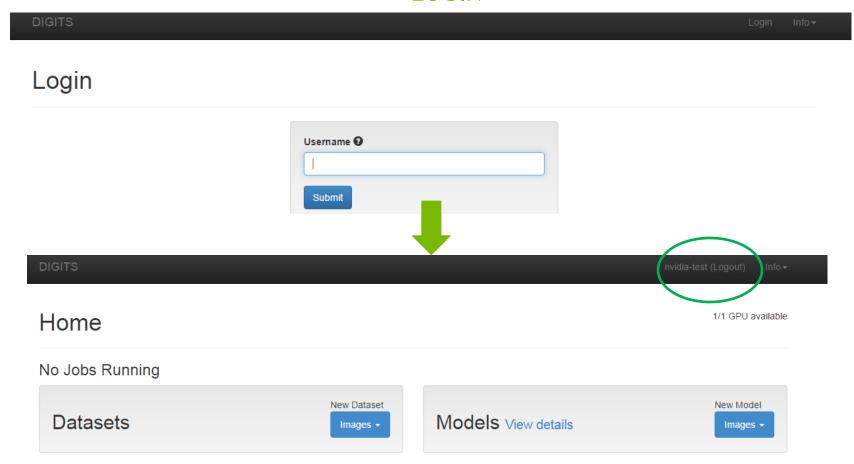
http://localhost/ (if installed from Deb packages),

http://localhost:5000/ (if using digits-devserver) or

http://localhost:34448/ (if using digits-server)



LOGIN



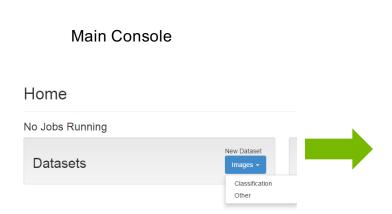


Creating your Dataset

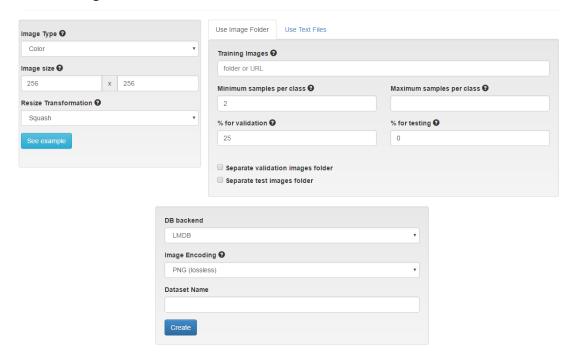
Main Console

Home Datasets In progress • None Completed • None New Dataset Models In progress • None Completed • None

Creating your Dataset



New Image Classification Dataset



In the Datasets section on the left side of the page, click on the blue Images button and select Classification which will take you to the "New Image Classification Dataset" page

Download MNIST dataset

Use the following command to download the MNIST dataset (for Deb package installations, the script is at/usr/share/digits/tools/download_data/main.py):

\$ tools/download_data/main.py mnist ~/mnist

Downloading url=http://yann.lecun.com/exdb/mnist/train-images-idx3-ubyte.gz...

Downloading url=http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz...

Downloading url=http://yann.lecun.com/exdb/mnist/t10k-images-idx3-ubyte.gz...

Downloading url=http://yann.lecun.com/exdb/mnist/t10k-labels-idx1-ubyte.gz...

Uncompressing file=train-images-idx3-ubyte.gz...

Uncompressing file=train-labels-idx1-ubyte.gz...

Uncompressing file=t10k-images-idx3-ubyte.gz...

Uncompressing file=t10k-labels-idx1-ubyte.gz...

Reading labels from /home/username/mnist/train-labels.bin...

Reading images from /home/username/mnist/test-labels.bin...

Reading images from /home/username/mnist/test-images.bin...

Dataset directory is created successfully at '/home/username/mnist'

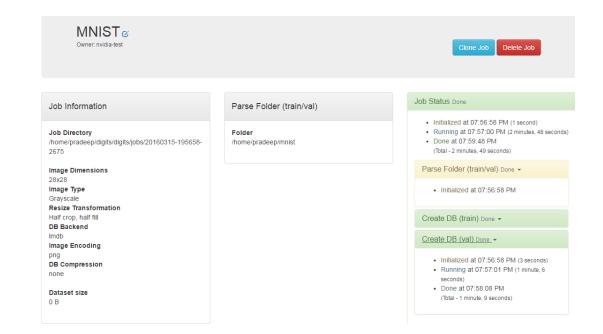
Done after 16.722807169 seconds.

Creating your Dataset

While the model creation job is running, you should see the expected completion time on the right side

When Model creation is done, you can also see completion time and duration

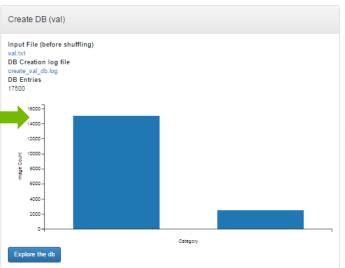
Users may download a copy of .txt files for reference



Database results

- Validation data tests the performance of the network
 - This data is only used for testing the generalization ability of the network
 - Not used to teach/train network
 - Prevents use of and identifies when network is overfit.
 - In current example 17500 images used for validation.





- Training data is used to train our neural network.
- Teaches the network to classify object categories
- Training Data Set,
 Current example uses
 52500 Images for training

Database Results

Database of Images can be explored via Exploring "Explore DB" tab and you can see test images used for training/validation



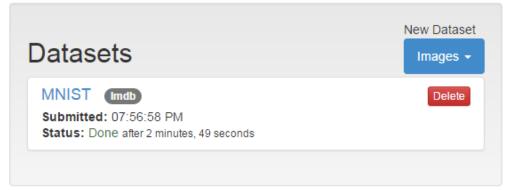
DIGITS DEMO

Creating your Dataset

Home

No Jobs Running

A new database is Created as visible on Home Page of DIGITS





TRAINING

Choose Framework

With DIGITS 3.0, two frameworks are integrated into DIGITS

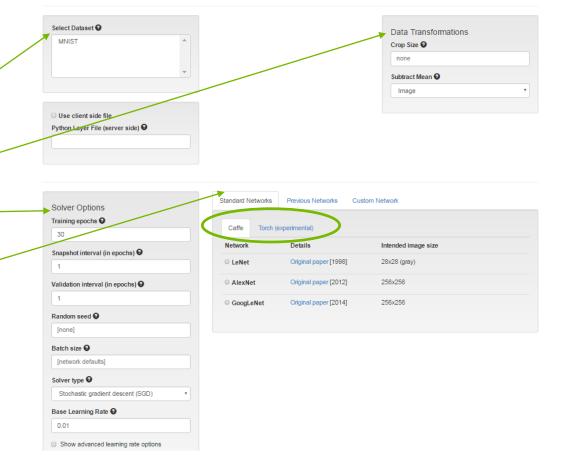
- Caffe
- Torch

Train A network

Training a network interface, please note

- Database selection
- Data Transformations
- Solver Options
- Different Network configurations
 - Caffe
 - Torch(experimental)

New Image Classification Model



Train a Network

Select the Database

Provide a Model Name

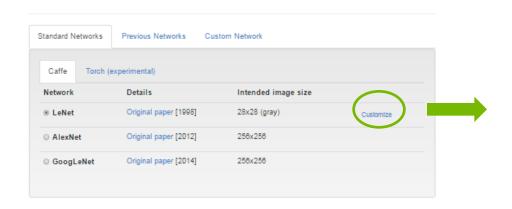
Do any changes in Solver options

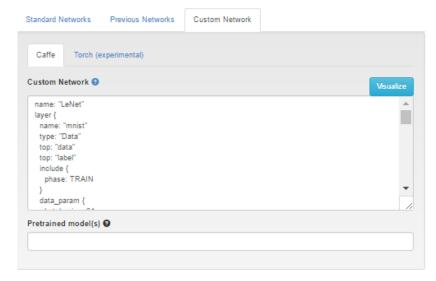
Start with a default Network like LeNet (Framework can be anyone Caffe/Torch)

Click on Create Button

New Image Classification Model Data Transformations Done Tue Mar 15, 07:59:48 PM Crop Size 9 Image Size 28x28 Image Type Subtract Mean DB backend Create DB (train) 52500 images Use client side file Create DB (val) Python Layer File (server side) @ Standard Networks Previous Networks Custom Network Solver Options Training epochs Caffe Network Intended image size Snapshot interval (in epochs) 9 LeNet Original paper [1998] 28x28 (gray) AlexNet Original paper [2012] 258×258 Validation interval (in epochs) 9 Original paper [2014] 258×258 GoogLeNet Random seed @ [none] Batch size 😉 [network defaults] Stochastic gradient descent (SGD) Base Learning Rate @ Model Name MNIST-Model

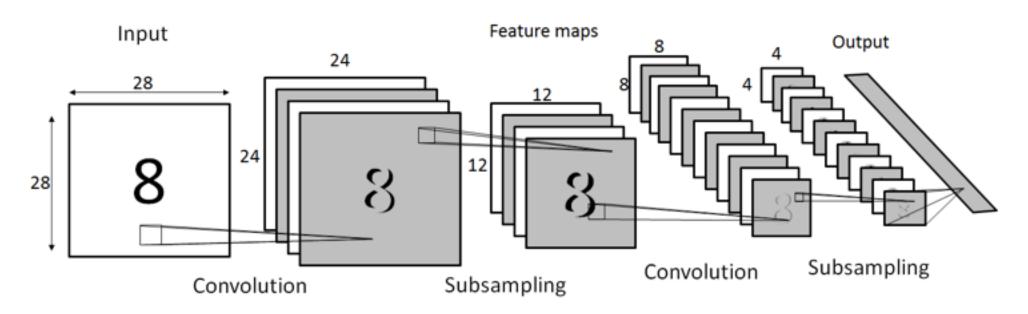
Train a Network-Advance Options



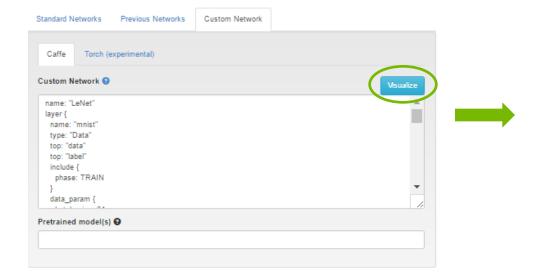


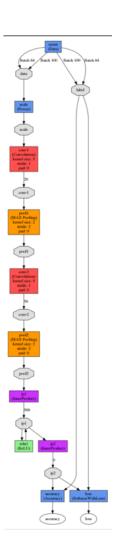
LENET

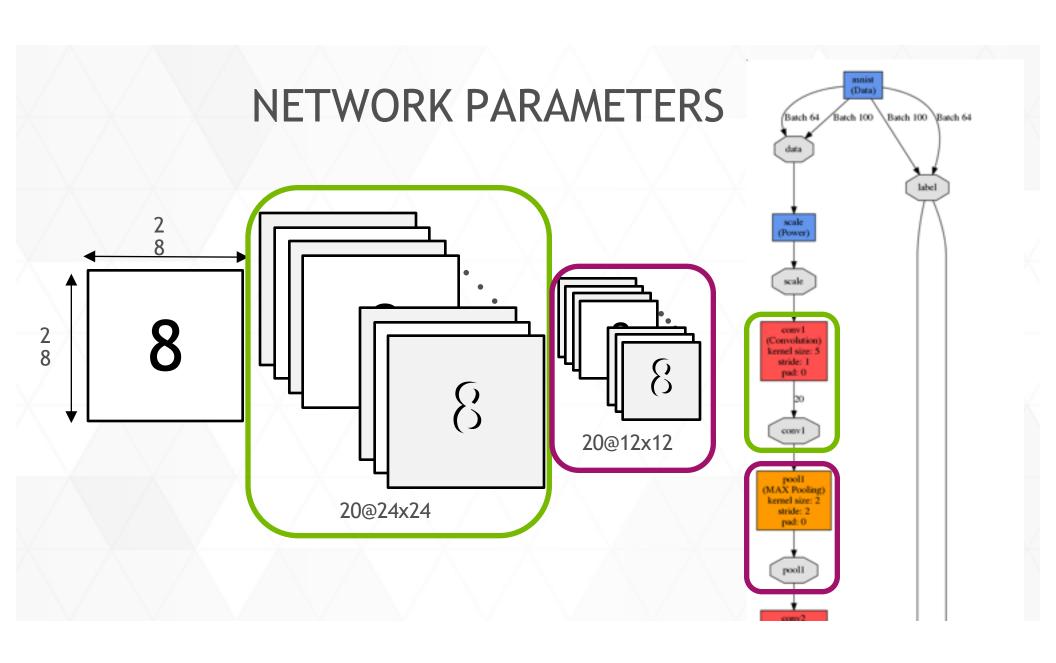
Network Configuration



Train a Network-Advance Options

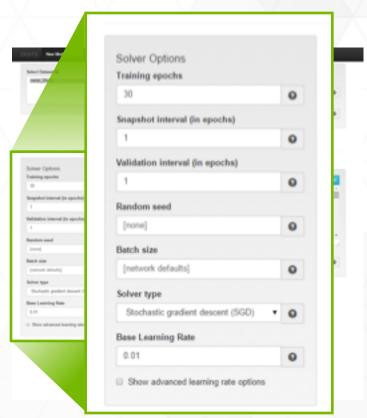






Train a Network-Advance Options

- Training epochs processing of all data
- Snapshot interval saving trained network
- Validation interval DNN test with the validation data
- Batch size number of images processed together
- Solver type SGD, ADAGRAD, NAG
- Learning rate and policy



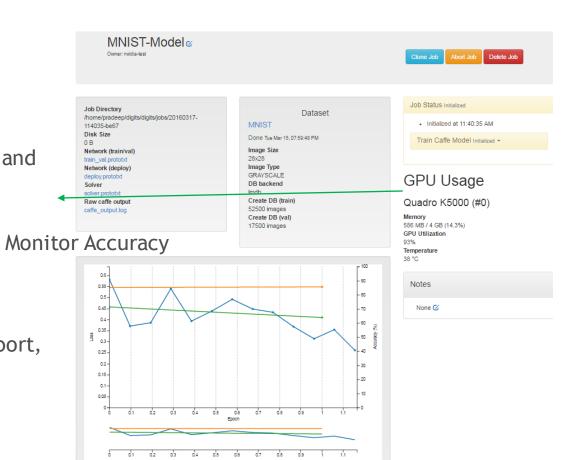
Single and multi-GPU training is easy



Training Results

Monitor GPU, Memory usage and Temperature

If performance is poor, abort, modify, and retrain.

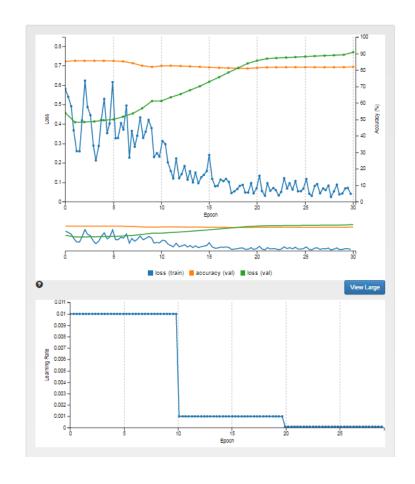


loss (train) accuracy (val) loss (val)

Training Results

First Graph is about the Loss and Accuracy graphs

Second Graph is about the Learning rate, as the training progresses, learning rate goes down as model is getting more mature.



OVERFITTING AND UNDERFITTING

How can I use DIGITS to tell me this is happening?

Overfitting

| Substitute | Sub

Validation data helps you identify when/if this occurs!

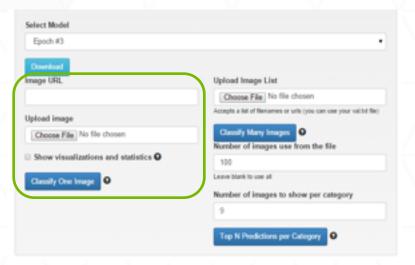
Underfitting



Modifying your Network & Classification

NVIDIA DIGITS

Single Image Classification





8







* Lab tasks

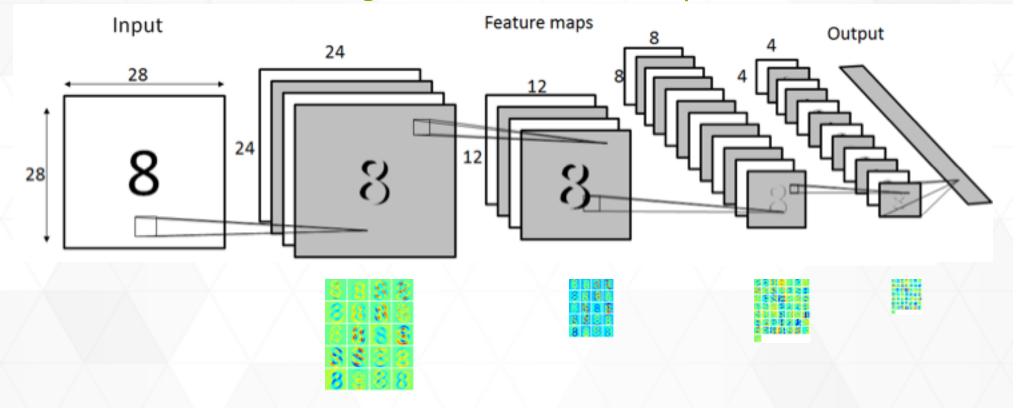
SINGLE IMAGE CLASSIFICATION RESULTS



- Network response at each layer will display
- Visualize responses from different inputs

NETWORK CONFIGURATION

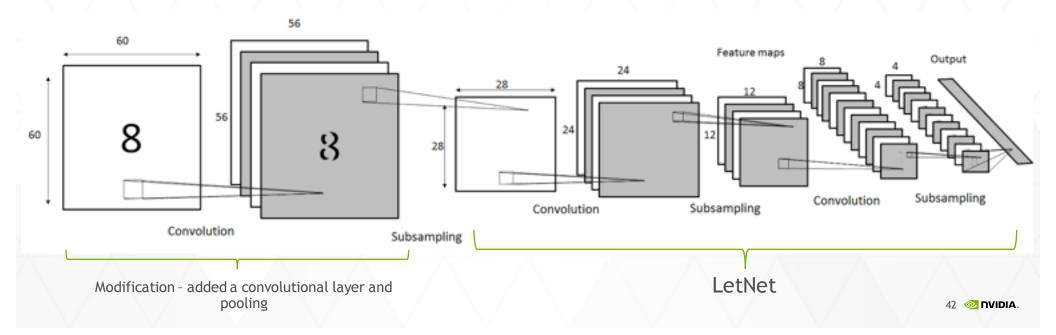
Visualizing LetNet Network Responses

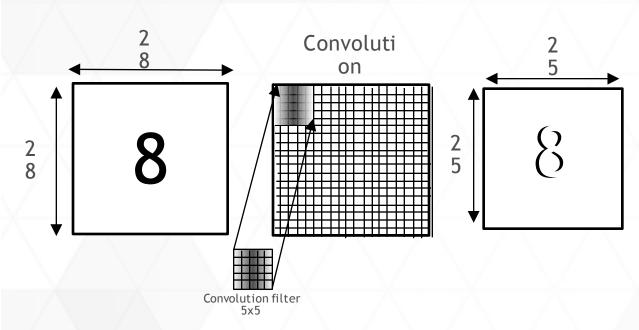


NETWORK CONFIGURATIONS

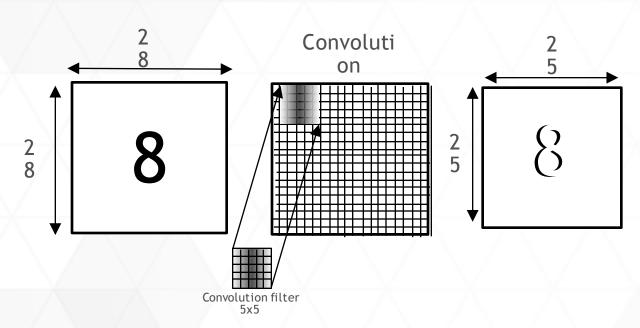
Modifying a Network

- Modifying a network can improve performance
- There are many parameters add or remove a layer, pooling, activation function, zero padding, increasing outputs

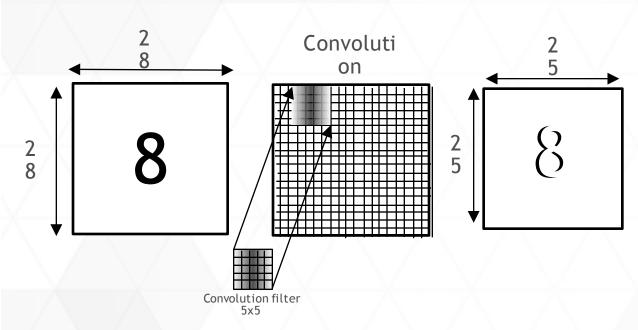




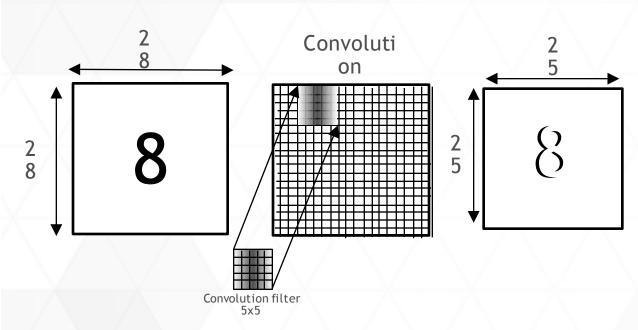
```
layer {
 name: "conv1"
 type: "Convolution"
 bottom: "scale"
 top: "conv1"
 param {
  Ir mult: 1
 param {
  Ir mult: 2
 convolution param {
  num output: 20
  kernel size: 5
  stride: 1
  weight filler {
   type: "xavier"
  bias filler {
    type: "constant"
```



```
layer {
 name: "conv1"
 type: "Convolution"
 bottom: "scale"
 top: "conv1"
 param {
  Ir mult: 1
 param {
  Ir mult: 2
 convolution param {
  num output: 20
  kernel size: 5
  stride: 1
  weight filler {
   type: "xavier"
  bias filler {
    type: "constant"
```



```
layer {
 name: "conv1"
 type: "Convolution"
 bottom: "scale"
 top: "conv1"
 param {
  Ir mult: 1
 param {
  Ir mult: 2
 convolution param {
  num output: 20
  kernel size: 5
  stride: 1
  weight filler {
   type: "xavier"
  bias filler {
    type: "constant"
```



```
layer {
 name: "conv1"
 type: "Convolution"
 bottom: "scale"
 top: "conv1"
 param {
  Ir mult: 1
 param {
  Ir mult: 2
 convolution param {
  num output: 20
  kernel size: 5
  stride: 1
  weight filler {
   type: "xavier"
  bias filler {
    type: "constant"
```

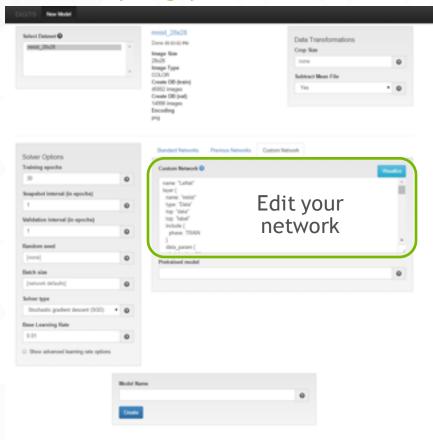
NETWORK PARAMETERS

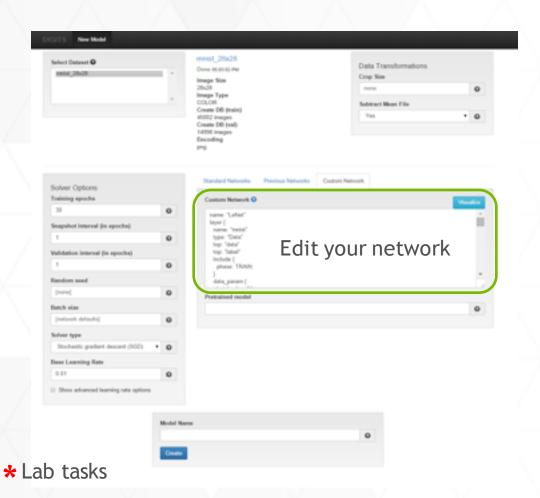
```
Convolution
                                                Pooling/Subsampling
                                                                                       Activation
layer {
                        weight_filler {
                                                     layer {
name: "conv0"
                                                                                        layer {
                           type: "xavier"
                                                      name: "pool0"
type: "Convolution"
                                                                                         name: "relu0"
                                                      type: "Pooling"
bottom: "data"
                                                                                        type: "ReLU"
                        bias_filler {
                                                      bottom: "conv0"
top: "conv0"
                                                                                        bottom: "pool0"
                           type: "constant"
                                                      top: "pool0"
param {
                                                                                         top: "pool0"
                           value: 0.9
                                                      pooling param {
 Ir_mult: 1.0
                                                       pool: MAX
                                                       kernel_size: 2
param {
                                                       stride: 2
 Ir_mult: 2.0
convolution param {
 num output: 20
 kernel size: 5
 stride: 1
```

weight_filler {
 type: "xavier"

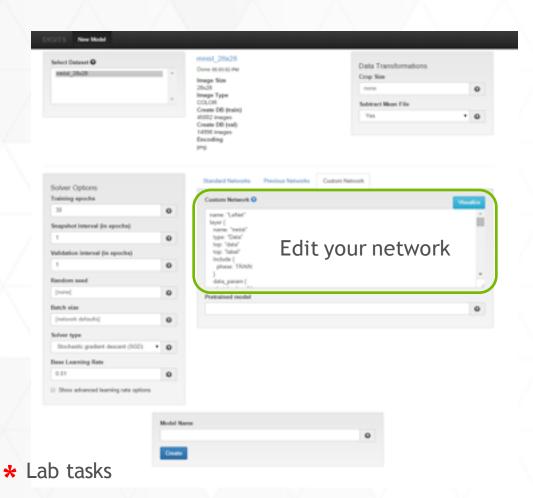
NVIDIA DIGITS

Modifying your Network



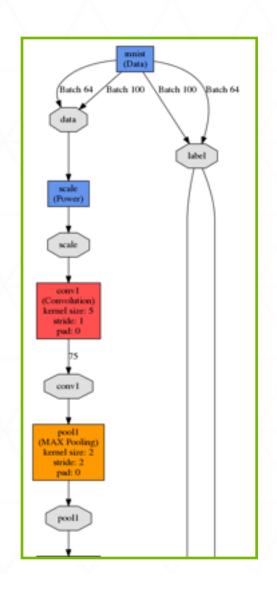


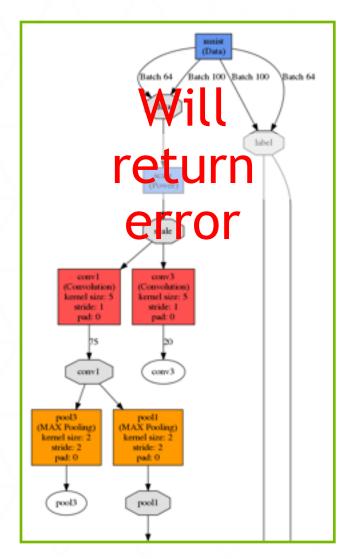
```
layer {
 name: "conv1"
 type: "Convolution"
 bottom: "scale"
                   Input and output to
 top: "conv1"
                           the layer
 param {
  Ir mult: 1
 param {
  Ir mult: 2
 convolution param {
  num output: 20 🛠
  kernel size: 5
  stride: 1
  weight filler {
   type: "xavier"
  bias filler {
   type: "constant"
```



```
# layer {
    name: "relu1"
    type: "ReLU"
    bottom: "conv1"
    top: "conv1"
}
Input and output to
the layer
```

Visualize Configuration Changes





ANOTHER WAY TO IMPROVE PERFORMANCE

Data Augmentation

- Sometimes training data is not a great representation of the field data
 - MNIST data is grayscale, black text with white background
 - 0123456789
- Will these images to be classified correctly when the network is trained with this digit data?









ANOTHER WAY TO IMPROVE PERFORMANCE

Data Augmentation

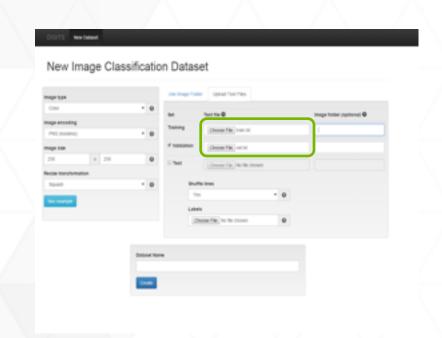
 Depending on the deployment scenario, simple modifications can be made to the training data to improve performance



- There are many ways to augment data
 - Rotations, noise, color distortions, stretching, etc. *
- Many ways to modify images ImageMagick, Pillow, OpenCV

* Lab tasks

USING AN AUGMENTED DATA SET



train.txt

/home/user/train/0/0_1.jpg 0

/home/user/train/0/0_1_invert.jpg 0

/home/user/train/5/5_1.jpg 5

/home/user/train/5/5_1_invert.jpg 5

val.txt

/home/user/mnist/val/7_1.jpg 7

/home/user/mnist/val/7_1_invert.jpg 7



ANOTHER WAY TO IMPROVE PERFORMANCE

Data Augmentation

• Example augmentation - inverted copies of the input data









• Would a network trained with this data augmented, accurately classify these images?

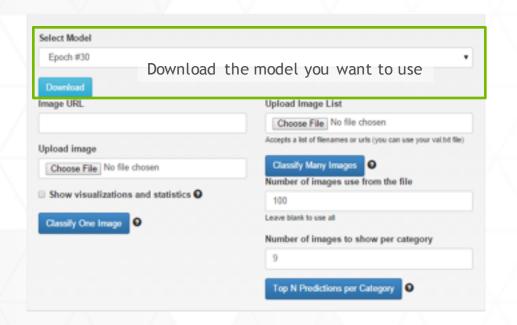








DEPLOYING YOUR NETWORK



Deploy in the cloud



• Deploy on a mobile device



DEPLOYMENT WITH TEGRA

Rapid Classification Anywhere

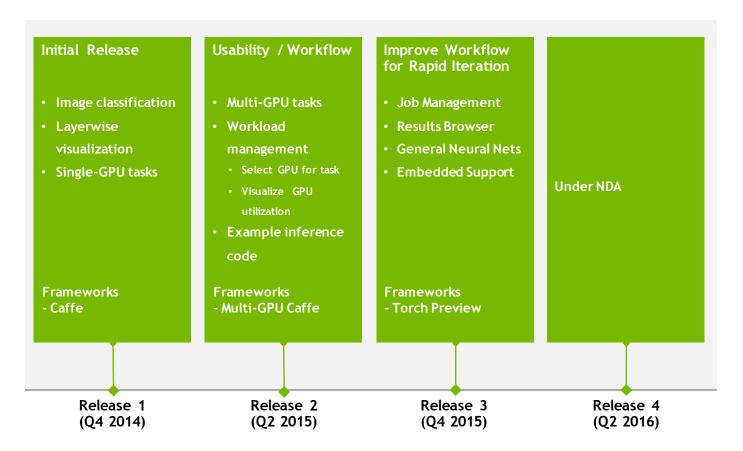
- > Flexible
- ➤ Low Power
- ➤ Easy to use
- ➤ GPU accelerated



Build Caffe on your portable platform Download your trained network from DIGITS

DIGITS Features at a Glance

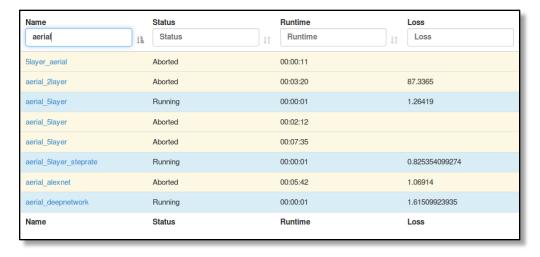
NVIDIA DIGITS ROADMAP



WHAT'S NEW IN DIGITS 3?

DIGITS 3 Improves Training Productivity with Enhanced Workflows

- Train neural network models with Torch support (preview)
- Save time by quickly iterating to identify the best model
- Easily manage multiple jobs to optimize use of system resources
- Active open source project with valuable community contributions



New Results Browser!

developer.nvidia.com/digits



NVIDIA DIGITS

Resources

- Where to get DIGITS
 - Easy to use web installer https://developer.nvidia.com/digits
 - github https://github.com/NVIDIA/DIGITS
 - Remember to install NVIDIA's Caffe branch https://github.com/NVIDIA/caffe
- User support
 - DIGITS Users Google group https://groups.google.com/forum/#!forum/digits-users
- For more information on getting started with DIGITS
 - Parallel forall http://devblogs.nvidia.com/parallelforall/easy-multi-gpu-deep-learning-digits-2/
 - Getting started guide https://github.com/NVIDIA/DIGITS/blob/master/docs/GettingStarted.md

HANDS-ON LAB

- 1. Create an account at <u>nvidia.qwiklab.com</u>
- 2. Go to "Getting Started with DIGITS" lab at bit.ly/dlnvlab2
- 3. Start the lab and enjoy!
- Only requires a supported browser, no NVIDIA GPU necessary!
- Lab is free until end of this Deep Learning Lab series

DEEP LEARNING SERIES

- Review the other seminars in series
- Seminar #3 Getting Started with the Caffe Framework
- Seminar #4 Getting Started with the Theano Framework
- Seminar #5 Getting Started with the Torch Framework
- More information available at <u>developer.nvidia.com/deep-learning-</u> courses