

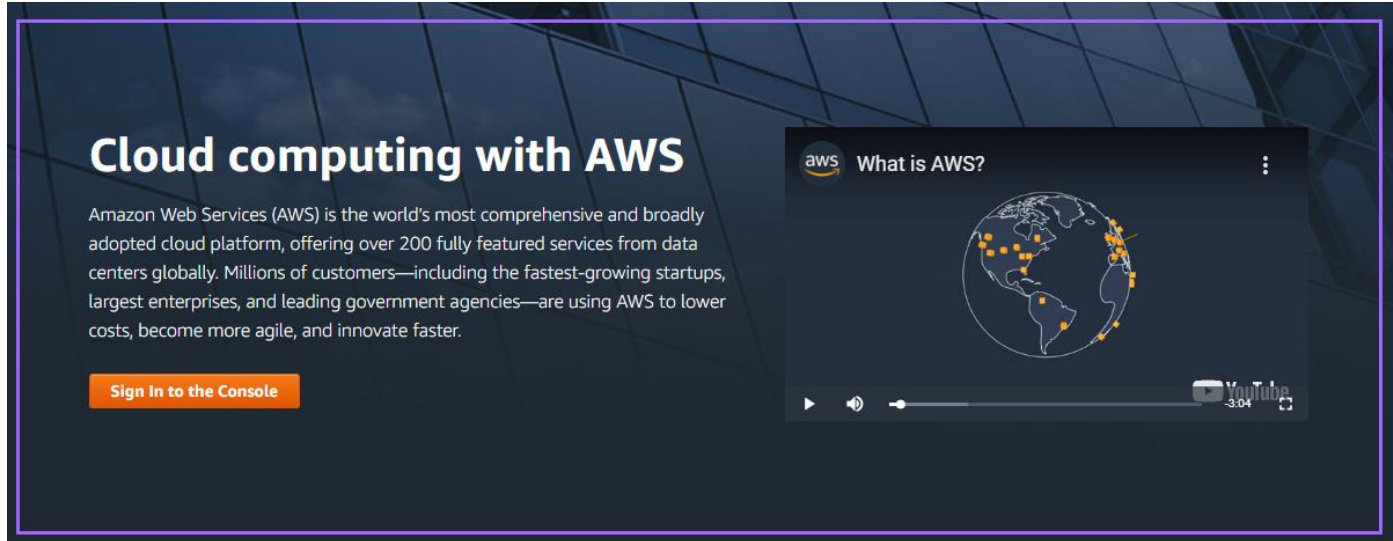
Deploying Accelerated Computing Instances in AWS

YULMAN PEREZ

SUPRI-A
Feb 11, 2021

What is AWS?

- It is a private and public cloud platform.
- It is very useful to build, deploy and manage applications.



How does it work?

Step by Step into AWS

- 1. Create an AWS account
- 2. Sign in to the Console <https://portal.azure.com>.
- 3. Search for EC2 Instances
- 4. Launch Instance, search for the GPU that you are planning to deploy (NVIDIA) or your preferred GPU.
- 5. Define all the connection details.

Step by Step into AWS

The screenshot shows the AWS IAM console interface for selecting an Amazon Machine Image (AMI). The top navigation bar includes the AWS logo, 'Services', a search bar, and user information. The main content area is titled 'Step 1: Choose an Amazon Machine Image (AMI)' and features a progress bar with steps: 1. Choose AMI, 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Add Tags, 6. Configure Security Group, and 7. Review.

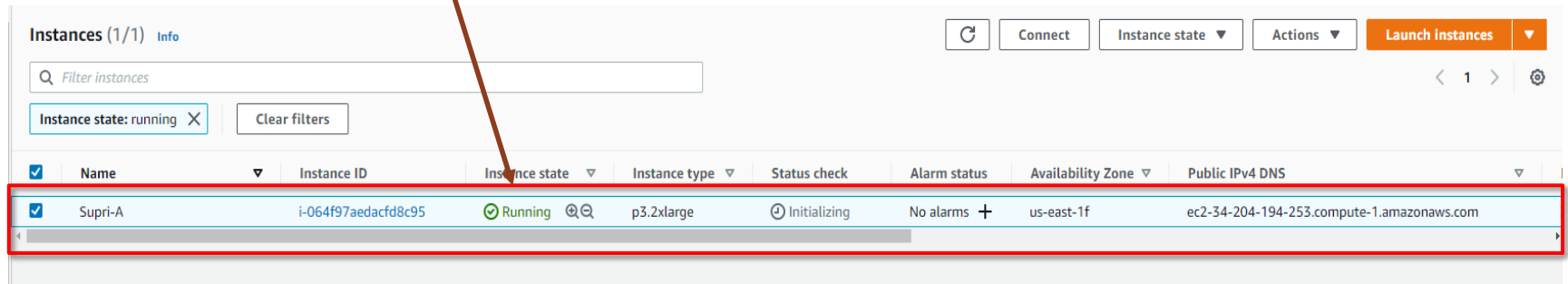
On the left, there are filters for 'Software Free Trial', 'Software Pricing Plans', 'Support', and 'Region'. The 'Region' filter is currently set to 'Current Region (458)'. The main list of AMIs includes:

- AWS Deep Learning AMI (Amazon Linux)**: Linux/Unix, Amazon Linux 2018.03 | 64-bit (x86) Amazon Machine Image (AMI) | Updated: 12/8/20. This AMI is pre-built and optimized for deep learning on EC2 with NVIDIA CUDA, cuDNN, and Intel MKL-DNN. It has popular frameworks like TensorFlow, MXNet, PyTorch, and tools like TensorBoard, TensorFlow Serving, Multi Model Server and Elastic Inference.
- AWS Deep Learning AMI (Ubuntu 16.04)**: Linux/Unix, Ubuntu 16.04 | 64-bit (x86) Amazon Machine Image (AMI) | Updated: 12/8/20. This AMI is pre-built and optimized for deep learning on EC2 with NVIDIA CUDA, cuDNN, and Intel MKL-DNN. It has popular frameworks like TensorFlow, MXNet, PyTorch, and tools like TensorBoard, TensorFlow Serving, Multi Model Server and Elastic Inference.
- Deep Learning Reference Stack**: TensorFlow: MKL-DNN, AVX512 Previous versions | By Intel. The Deep Learning Reference Stack, is an integrated, highly-performant open source stack optimized for Intel Xeon Scalable platforms. This open source community release is part of an effort to ensure AI developers have easy access to all features and functionality of Intel platforms.
- NVIDIA Deep Learning AMI**: Linux/Unix, Ubuntu 16.04 | 64-bit (x86) Amazon Machine Image (AMI) | Updated: 11/23/20. The NVIDIA Deep Learning AMI is an optimized environment for running the GPU-optimized deep learning and HPC containers from the NVIDIA NGC Catalog. The deep learning containers on the NGC container registry require this AMI for GPU acceleration on AWS P4D, P3 and G4 GPU instances.

The 'NVIDIA Deep Learning AMI' is highlighted with a red box, indicating it is the selected option.

Step by Step into AWS

Make sure your instance is running.



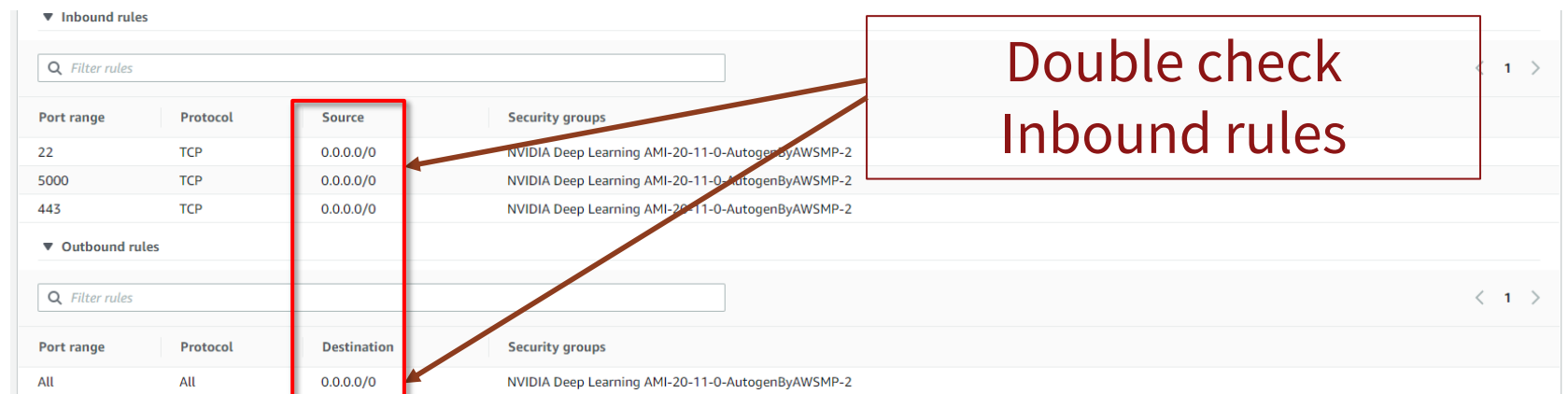
Instances (1/1) Info

Filter instances

Instance state: running X Clear filters

<input checked="" type="checkbox"/>	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
<input checked="" type="checkbox"/>	Supri-A	i-064f97aedacfd8c95	Running	p3.2xlarge	Initializing	No alarms +	us-east-1f	ec2-34-204-194-253.compute-1.amazonaws.com

Double check
Inbound rules



▼ Inbound rules

Filter rules

Port range	Protocol	Source	Security groups
22	TCP	0.0.0.0/0	NVIDIA Deep Learning AMI-20-11-0-AutogenByAWSMP-2
5000	TCP	0.0.0.0/0	NVIDIA Deep Learning AMI-20-11-0-AutogenByAWSMP-2
443	TCP	0.0.0.0/0	NVIDIA Deep Learning AMI-20-11-0-AutogenByAWSMP-2

▼ Outbound rules

Filter rules

Port range	Protocol	Destination	Security groups
All	All	0.0.0.0/0	NVIDIA Deep Learning AMI-20-11-0-AutogenByAWSMP-2

After this step the instance should be ready to use!

Steps to use AI with Matlab?

- 1. Matlab License.
- 2. Install PuTTY (Free SSH and Telnet for Windows).
- 3. Deploy the docker.
- 4. Pull the container using SSH.
- 5. Run the container for Matlab.
- 6. Tunnel SSH and it is ready to use.

Step by Step into AWS

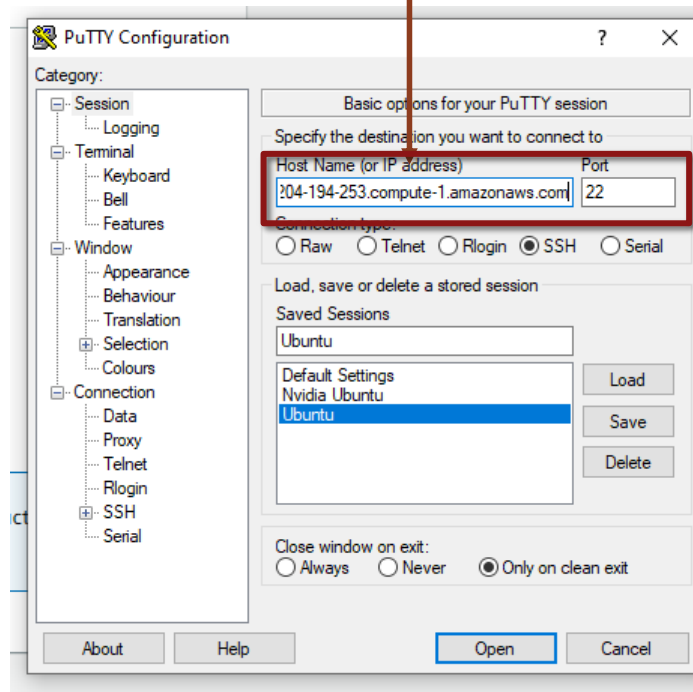
Instance summary for i-064f97aedacfd8c95 (Supri-A) [Info](#)

Updated less than a minute ago

[Refresh](#) [Connect](#) [Instance state ▼](#)

Instance ID i-064f97aedacfd8c95 (Supri-A)	Public IPv4 address 34.204.194.253 open address	Private IPv4 addresses 172.31.71.92
Instance state Running	Public IPv4 DNS ec2-34-204-194-253.compute-1.amazonaws.com open address	Private IPv4 DNS ip-172-31-71-92.ec2.internal
Instance type p3.2xlarge	Elastic IP addresses -	VPC ID vpc-f4ebf78e
AWS Compute Optimizer finding Opt-in to AWS Compute Optimizer for recommendations. Learn more	IAM Role -	Subnet ID subnet-b3d154bd

Start PuTTY



Step by Step into AWS

Run the Instance!

```
* Documentation:  https://help.ubuntu.com
* Management:    https://landscape.canonical.com
* Support:        https://ubuntu.com/advantage

System information as of Wed Feb 10 19:20:09 UTC 2021

System load:  0.08          Processes:           163
Usage of /:   67.0% of 30.96GB Users logged in:       0
Memory usage: 0%           IP address for ens3: 172.31.71.92
Swap usage:   0%

* Introducing self-healing high availability clusters in MicroK8s.
  Simple, hardened, Kubernetes for production, from RaspberryPi to DC.

  https://microk8s.io/high-availability

* Canonical Livepatch is available for installation.
  - Reduce system reboots and improve kernel security. Activate at:
    https://ubuntu.com/livepatch

0 packages can be updated.
0 of these updates are security updates.

Welcome to the NVIDIA GPU Cloud image. This image provides an optimized
environment for running the deep learning and HPC containers from the
NVIDIA GPU Cloud Container Registry. Many NGC containers are freely
available. However, some NGC containers require that you log in with
a valid NGC API key in order to access them. This is indicated by a
pull access denied for xyz ..." or "Get xyz: unauthorized: ..." error
message from the daemon.

Documentation on using this image and accessing the NVIDIA GPU Cloud
Container Registry can be found at
  http://docs.nvidia.com/ngc/index.html

Last login: Thu Jan 28 18:00:04 2021 from 171.66.10.27
ubuntu@ip-172-31-71-92:~$
```

You should be
connected
into the
NVIDIA GPU

Step by Step into AWS

Run the Container for Matlab !

Welcome to the MATLAB Deep Learning Container on NVIDIA GPU Cloud

This container includes commercial software products of The MathWorks, Inc. ("MathWorks Programs") and related materials. MathWorks Programs are licensed under the MathWorks Software License Agreement, available in the MATLAB installation in this container. Related materials in this container are licensed under separate licenses which can be found in their respective folders.

To run MATLAB desktop in this container, make sure you have exposed port 5901 and 6080 in the container.

For Docker 19.03 or later, use a docker run command of the form

```
docker run --gpus all -it --rm -p 5901:5901 -p 6080:6080 --shm-size=512M nvcr.io/partners/matlab:r2020b
```

For Docker 19.02 or earlier, use a docker run command of the form

```
nvidia-docker run -it --rm -p 5901:5901 -p 6080:6080 --shm-size=512M nvcr.io/partners/matlab:r2020b
```

To get started using the desktop in this container you can either

1. Point a browser to port 6080 of the docker host machine running this container

```
http://hostname:6080
```

2. Use a VNC client to connect to display 1 of the docker host machine

```
hostname:1
```

The default password to access the container desktop is

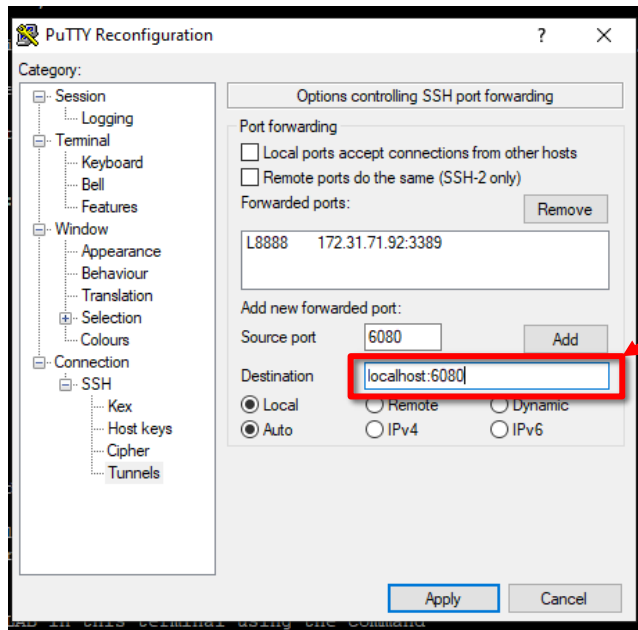
```
matlab
```

Launch MATLAB using the MATLAB icon on the desktop.

To securely access the desktop from a browser or using VNC, you can use

Step by Step into AWS

Tunnel the connection via VNC.

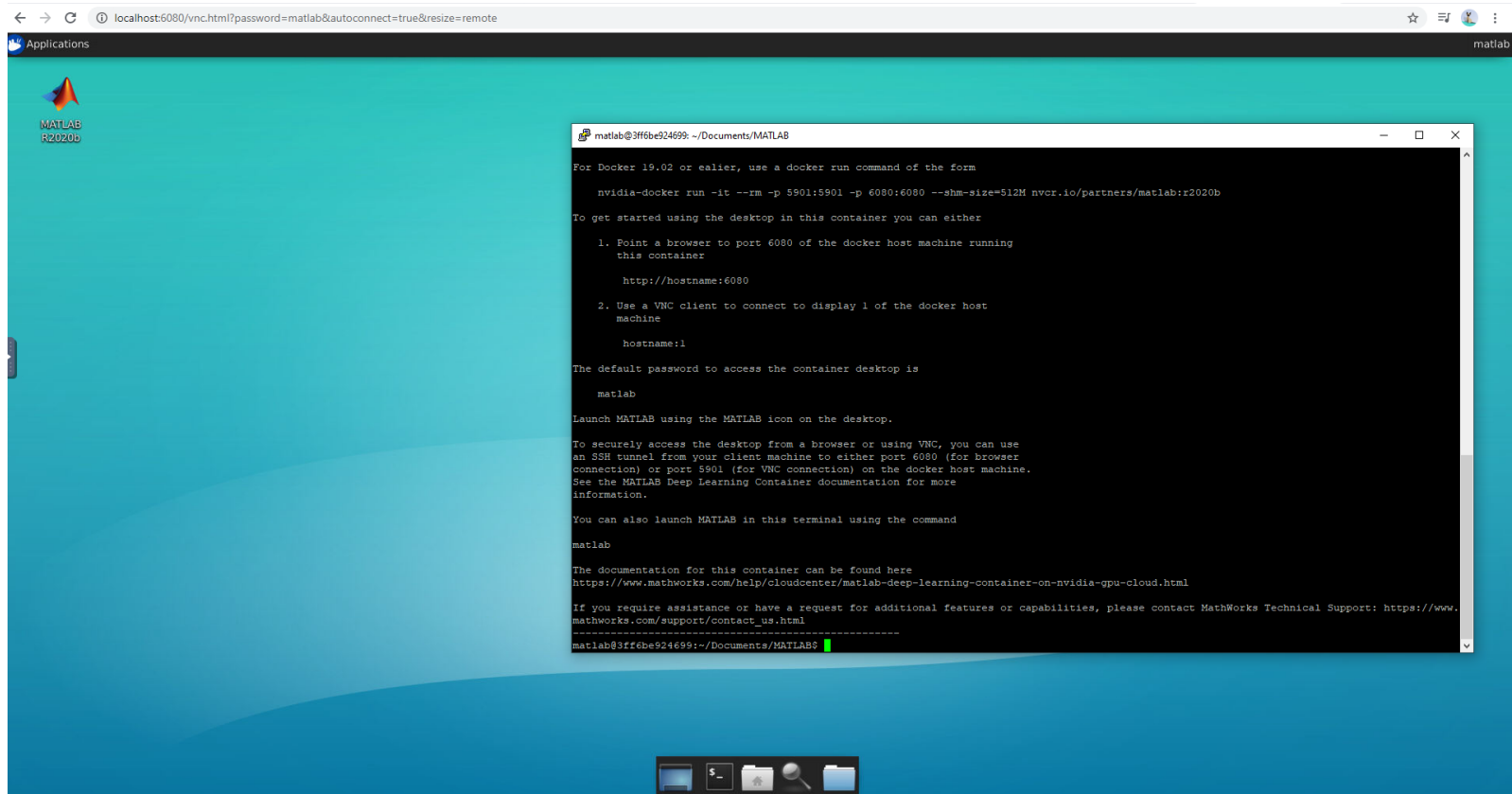


If the port does not work, try a different port (e.g., 6081, 6079)

You can connect with your web browser or VNC.

Step by Step into AWS

Open your browser and type: `http://localhost:6080/`



Matlab is ready!

11

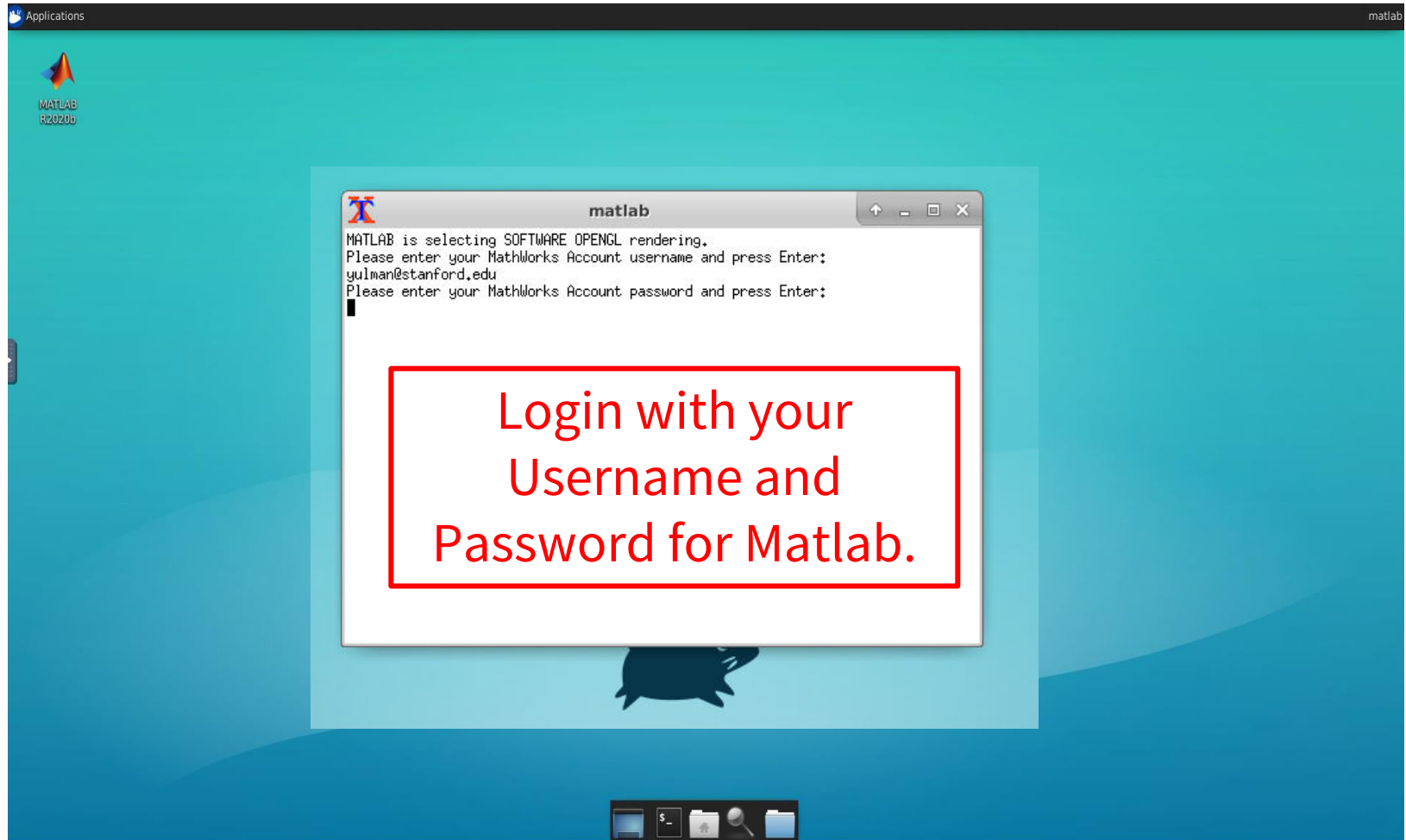
Step by Step into AWS

Applications

MATLAB
R2020b

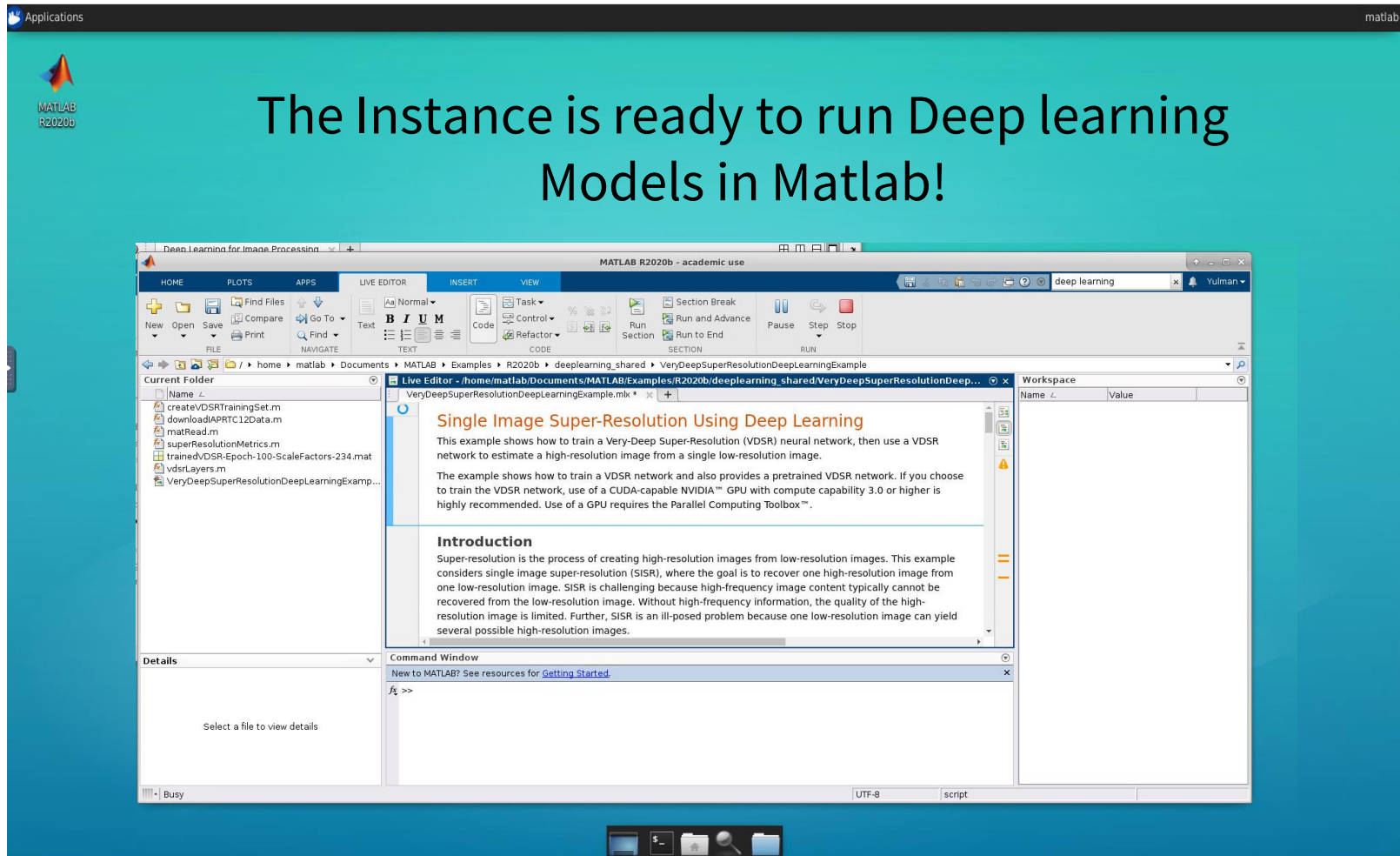
```
matlab
MATLAB is selecting SOFTWARE_OPENGL rendering.
Please enter your MathWorks Account username and press Enter:
yulman@stanford.edu
Please enter your MathWorks Account password and press Enter:
█
```

Step by Step into AWS



13

Step by Step into AWS



The screenshot displays the MATLAB R2020b desktop environment. The top menu bar includes 'HOME', 'PLOTS', 'APPS', 'LIVE EDITOR', 'INSERT', and 'VIEW'. The 'LIVE EDITOR' tab is active, showing a document titled 'Single Image Super-Resolution Using Deep Learning'. The document content includes a title, a description of the VDSR neural network, and an 'Introduction' section. The 'Current Folder' pane on the left lists files such as 'createVDSRTrainingSet.m', 'downloadAPRTC12Data.m', 'matRead.m', 'superResolutionMetrics.m', 'trainedVDSR-Epoch-100-ScaleFactors-234.mat', 'vdsrLayers.m', and 'VeryDeepSuperResolutionDeepLearningExamp...'. The 'Command Window' at the bottom shows the prompt 'f1 >>'. The 'Workspace' pane on the right is empty. The status bar at the bottom indicates 'Busy', 'UTF-8', and 'script'.

Applications

MATLAB R2020b

The Instance is ready to run Deep learning Models in Matlab!

Deep Learning for Image Processing

MATLAB R2020b - academic use

HOME PLOTS APPS LIVE EDITOR INSERT VIEW

File Edit View Live Editor Insert View

Current Folder

- createVDSRTrainingSet.m
- downloadAPRTC12Data.m
- matRead.m
- superResolutionMetrics.m
- trainedVDSR-Epoch-100-ScaleFactors-234.mat
- vdsrLayers.m
- VeryDeepSuperResolutionDeepLearningExamp...

Live Editor - /home/matlab/Documents/MATLAB/Examples/R2020b/deeplearning_shared/VeryDeepSuperResolutionDeep...

Single Image Super-Resolution Using Deep Learning

This example shows how to train a Very-Deep Super-Resolution (VDSR) neural network, then use a VDSR network to estimate a high-resolution image from a single low-resolution image.

The example shows how to train a VDSR network and also provides a pretrained VDSR network. If you choose to train the VDSR network, use of a CUDA-capable NVIDIA™ GPU with compute capability 3.0 or higher is highly recommended. Use of a GPU requires the Parallel Computing Toolbox™.

Introduction

Super-resolution is the process of creating high-resolution images from low-resolution images. This example considers single image super-resolution (SISR), where the goal is to recover one high-resolution image from one low-resolution image. SISR is challenging because high-frequency image content typically cannot be recovered from the low-resolution image. Without high-frequency information, the quality of the high-resolution image is limited. Further, SISR is an ill-posed problem because one low-resolution image can yield several possible high-resolution images.

Details

Command Window

New to MATLAB? See resources for [Getting Started](#).

f1 >>

Workspace

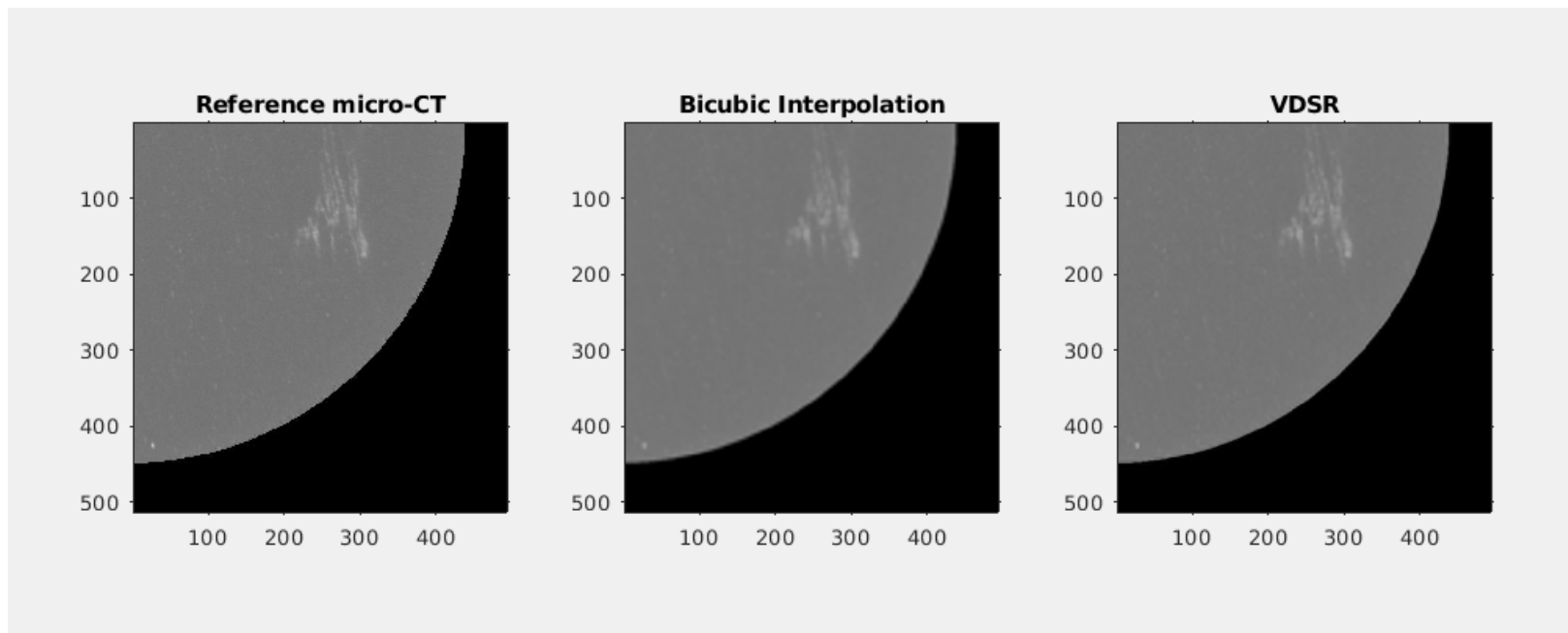
Name Value

Select a file to view details

Busy UTF-8 script

Example in Matlab (Shale Sample)

Numerical evaluations by using different metrics such as :
the Peak Signal-to-Noise Ration (PSNR)
Structure similarity index (SSIM)



Results for Scale factor 4

Average PSNR for Bicubic = 26.659808
Average PSNR for VDSR = 27.418227
Average SSIM for Bicubic = 0.858339
Average SSIM for VDSR = 0.873142

References for AWS

- **AWS Step by Step:**
- <https://docs.aws.amazon.com/dlami/latest/devguide/gpu.html>
- **Matlab AWS Deep Learning:**
- https://www.mathworks.com/help/cloudcenter/ug/matlab-deep-learning-container-on-aws.html#mw_05af73fd-10fe-42c1-bfee-c3539c22507e
- **Matlab 2020b:**
- <https://github.com/mathworks-ref-arch/matlab-on-aws>

Thank you!

Questions?