Hugging Face
BEAR Software Updates 10

Jenny Wong | 29 June 2023

Hugging Face is a <u>community and data science platform</u> that provides tools to enable users to build, train and deploy machine learning models based on <u>open source</u> code and technologies.

🤗

The AI community building the future.

Build, train and deploy state of the art models powered by the reference open source in machine learning.

○ Star | 105,355

More than 5,000 organizations are using Hugging Face

Ai2 **Allen Institute for AI** Non-Profit · 193 models	**⩗|** **Meta AI** Company · 695 models	**aws** **Amazon Web Services** Company · 2 models
intel **Intel** Company · 117 models	**SpeechBrain** Non-Profit · 70 models	**Microsoft** Company · 255 models

G **Google** Company · 587 models
Grammarly Company · 5 models

huggingface.co

Datasets | ImageNet-1k
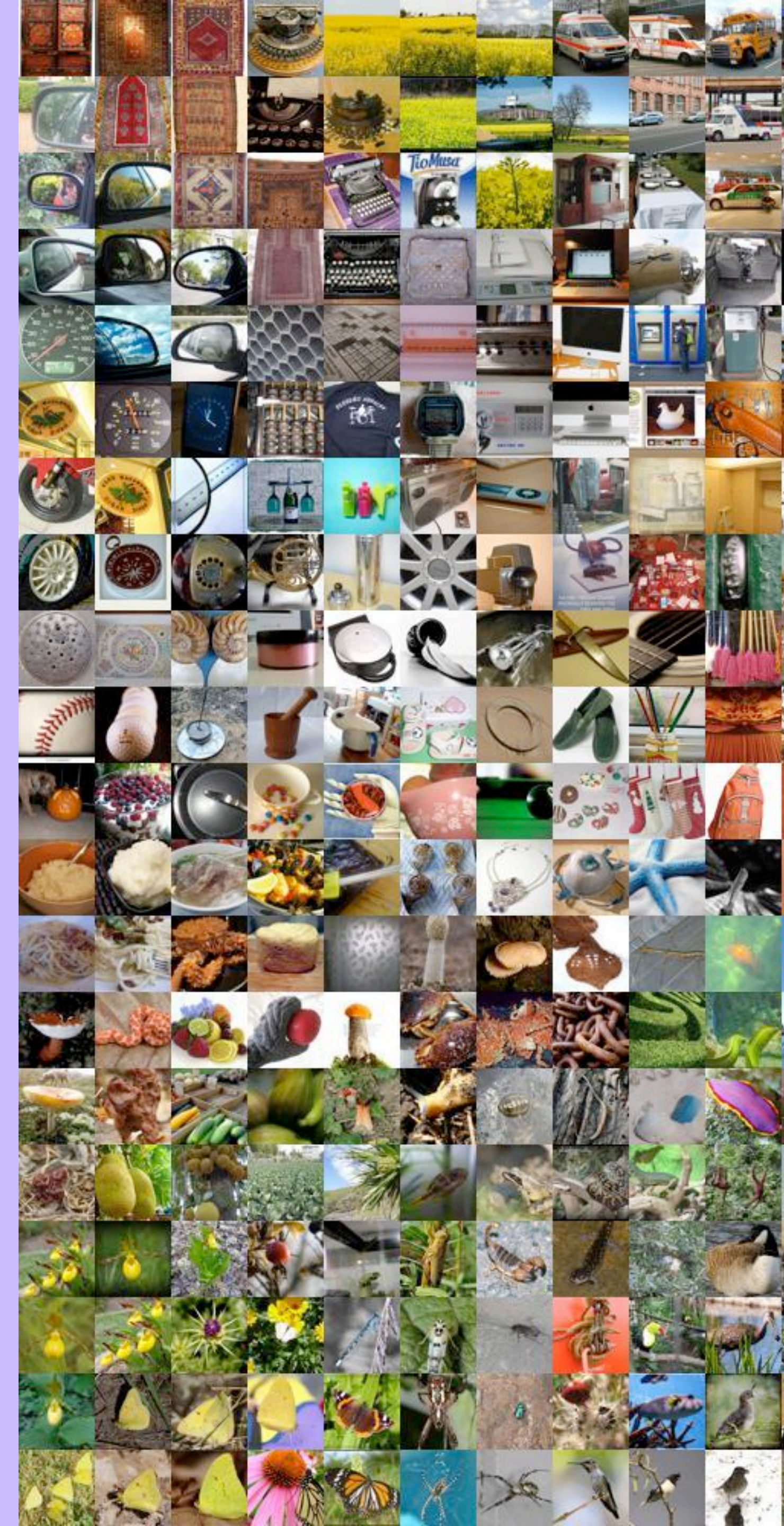
BEAR Software Updates 10

Import modules

```python
import os
from datasets import load_dataset
```

```python
imagenet_data_dir = '/bask/projects/e/edmondac-bear-chall/wongj/challenge_1/challenge_1_participant_info/input_data'
```
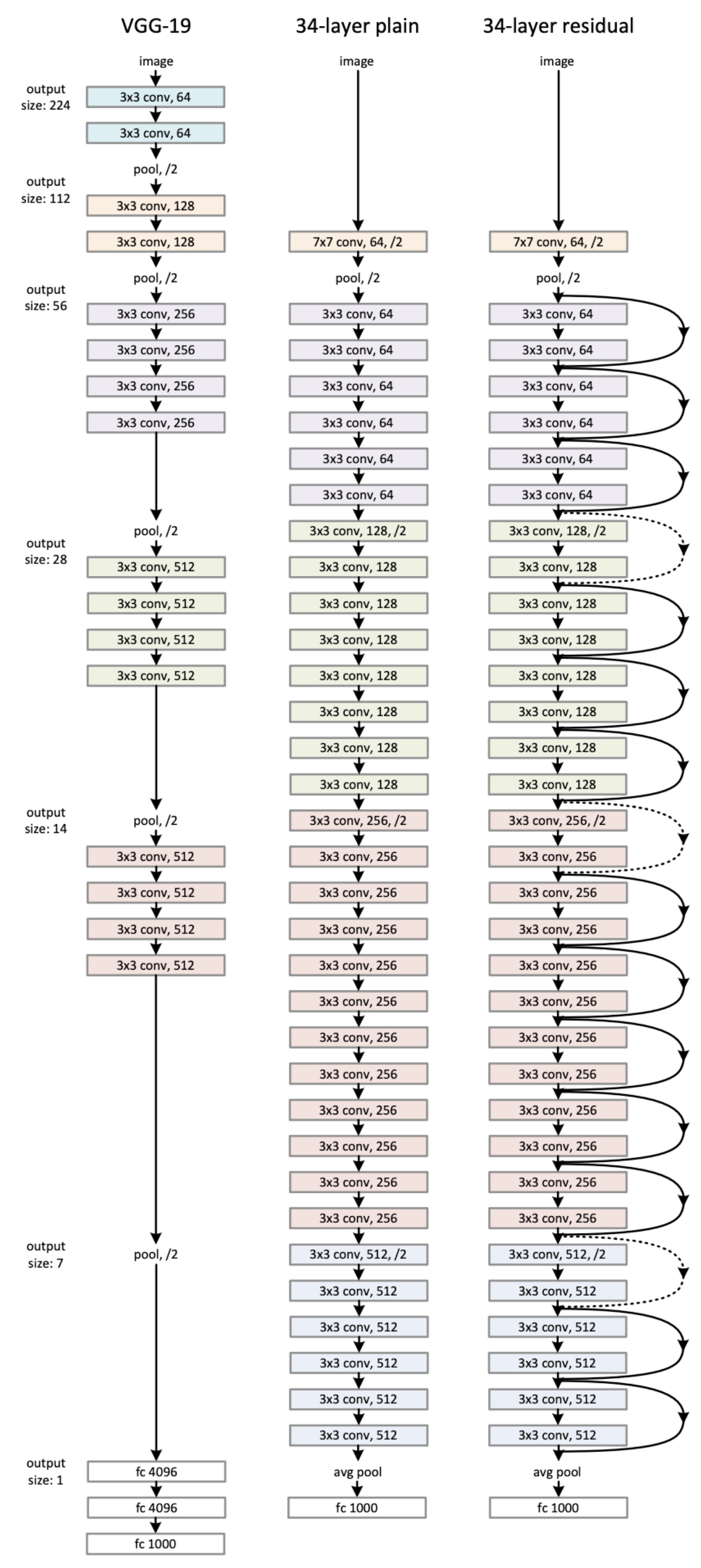
Load HF dataset

```python
dataset = load_dataset(
    'imagenet-1k',
    cache_dir = imagenet_data_dir,
    split = "train",
)
```

Filter for Space Shuttles

```python
space_shuttles = dataset.filter(
    lambda x: x == 812,
    input_columns=['label'],
    num_proc=16
)
```

Models | ResNet-50

Image classification: ResNet-50 model

Import modules

```python
from transformers import AutoImageProcessor, ResNetForImageClassification
import torch
```

Load test image dataset

```python
data_dir = os.getcwd()
dataset_star_trek = load_dataset(
    "vumichien/spaceship_star_trek",
    cache_dir=data_dir,
)

image = dataset_star_trek["train"]["image"][1]
image.show()
```

Load Model

```python
processor = AutoImageProcessor.from_pretrained("microsoft/resnet-50") # for processing the image into a tensor
model = ResNetForImageClassification.from_pretrained("microsoft/resnet-50") # load the ResNet-50 model

inputs = processor(image, return_tensors="pt")

with torch.no_grad(): # feed input into model without calculating gradients
    logits = model(**inputs).logits
```

Summary

HuggingFace promotes open source and open science in ML

FINDABLE HuggingFace Hub is essentially GitHub for ML

ACCESSIBLE Provides tools to lower the barrier to entry

INTEROPERABLE Works with both TensorFlow and PyTorch

REPRODUCIBLE Spaces where creators don't need to set up servers for demos