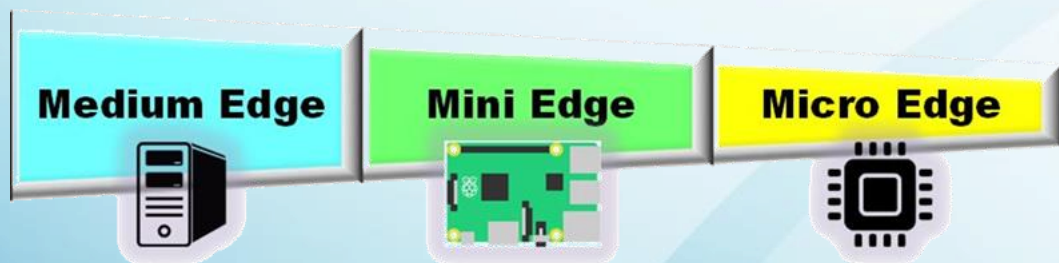
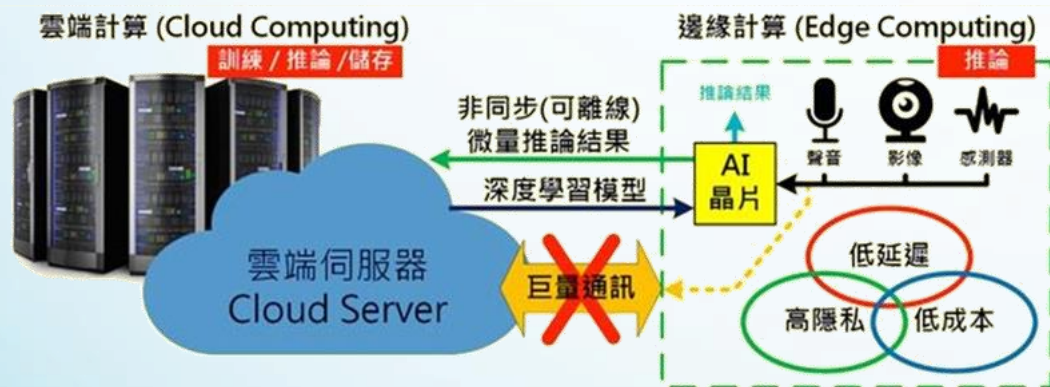


OmniXRI's Edge AI & TinyML 小學堂



沒有最邊



只有更邊

歡迎加入
邊緣人俱樂部



【第8講】

實作案例 – 影像分類



歐尼克斯實境互動工作室 (OmniXRI Studio)
許哲豪 (Jack Hsu)

簡報大綱



- 8.1. 影像分類簡介
- 8.2. 影像分類模型
- 8.3. 影像分類評量
- 8.4. 影像分類實作

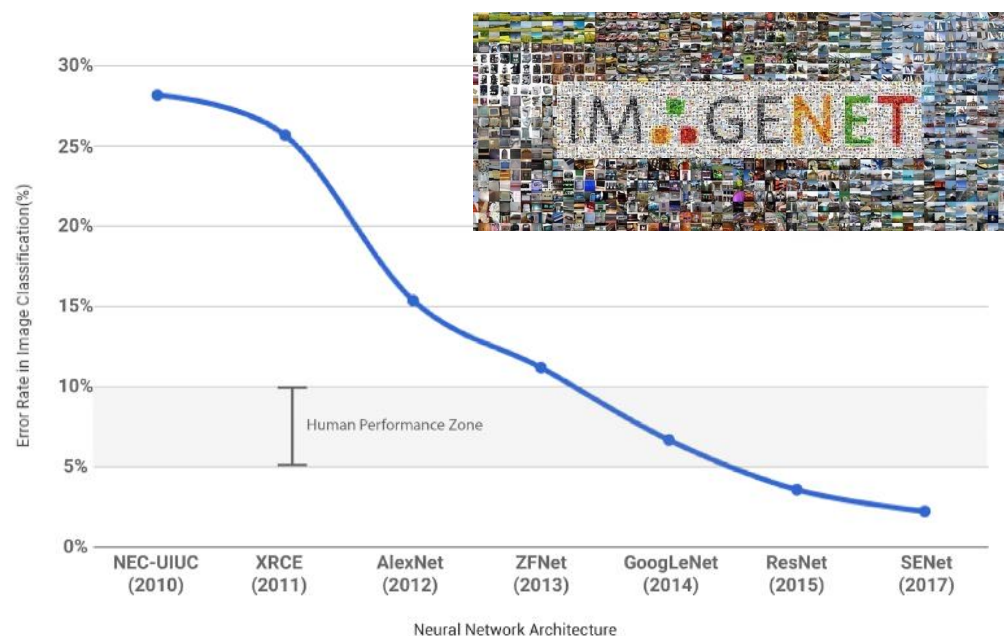
本課程完全免費，請勿移作商業用途！
歡迎留言、訂閱、點讚、轉發，讓更多需要的朋友也能一起學習。

完整課程大綱：<https://omnixri.blogspot.com/2024/02/omnixris-edge-ai-tinymml-0.html>
課程直播清單：<https://www.youtube.com/@omnixri1784/streams>

影像分類簡介

ILSVRC 為 **ImageNet** 大規模視覺挑戰賽，由史丹佛大學李飛飛教授發起。2010年起每年舉行，剛開始辨識錯誤率約在25~30%，2012年AlexNet取得重大突破，以接近15%的錯誤率贏得冠軍，往後開啟卷積神經網路(CNN)及深度學習(Deep Learning)的榮景。直到2017年SeNet以2.3%(Top5)錯誤率，遠低於人類辨識能力，於是停辦該比賽。

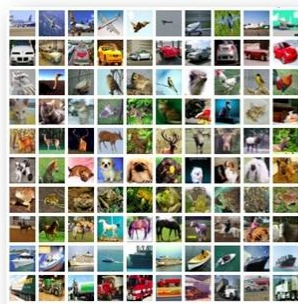
ImageNet擁有1400多萬張影像，具有2萬多個類別。經整理後取**1000**個類別用於ILSVRC比賽。



常見影像資料集

更多內容請參考
第3講資料集建置與標註

- MNIST 手寫數字
 - 28*28灰階影像
 - 共分十類，數字 0 ~ 9
 - 6萬張訓練集，1萬張測試集
- CIFAR-10 彩色影像
 - 32*32 彩色影像
 - 共分十類，飛機、貓、狗...
 - 6萬張，每類6千張影像



Pattern
Analysis
Statistical Modeling and
Computational
Learning
Visual
Object
Classes



Pascal VOC (2005 ~ 2012)

共有1萬7千多張影像，分為20類，標註內容包括影像分類、物件偵測、語義分割等。其標註資料主要採用XML格式。

<http://host.robots.ox.ac.uk/pascal/VOC/>

ImageNet
Large
Scale
Visual
Recognition
Challenge



ImageNet (2010 ~ 2017)

共有超過1400萬張影像，透過Amazon Mechanical Turk外包協助下進行手動標註，包含2萬多個類別，超過100萬個物件偵測邊界框(Bounding Box)被標註。

<http://image-net.org/>

Common
Objects
In
Context



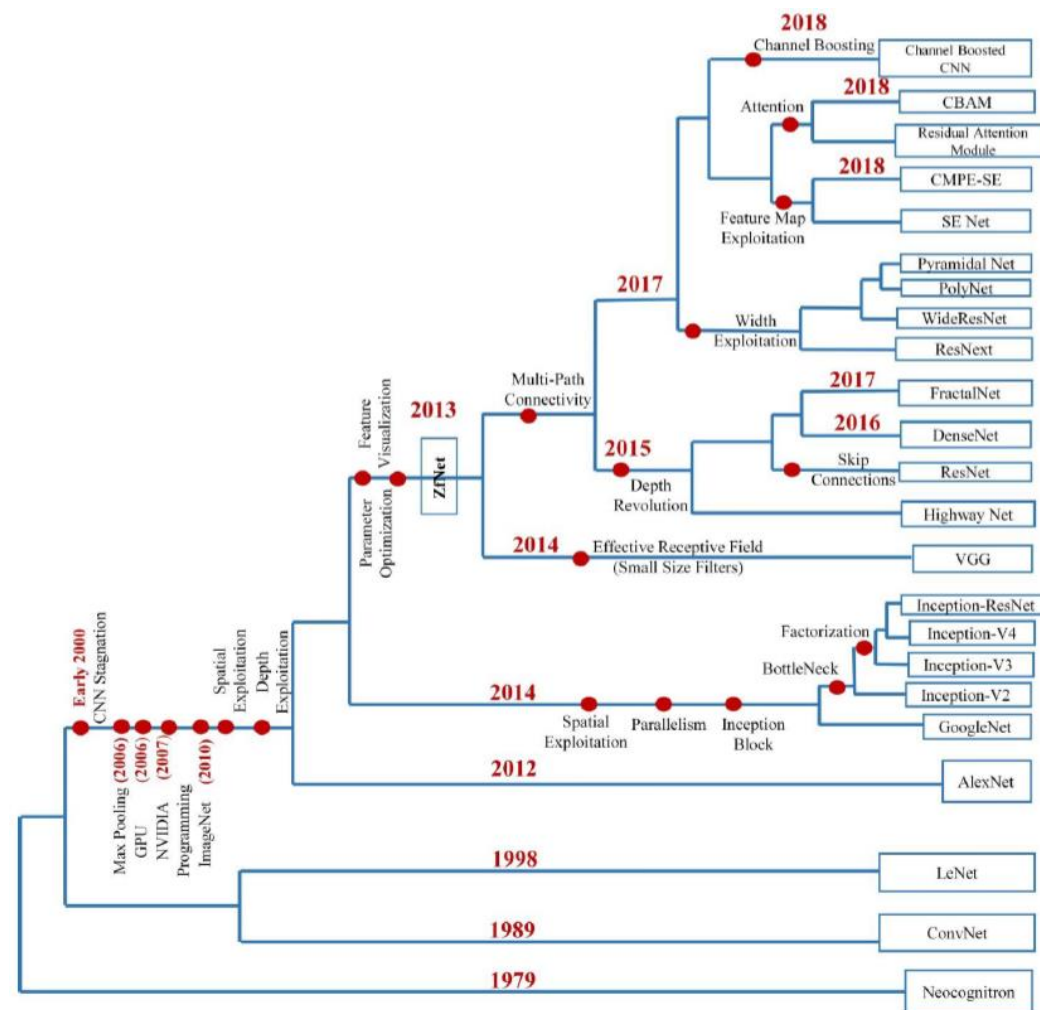
MS COCO (2015 ~)

共有超過32萬張影像，包含91(80)個類別，超過250萬個物件測邊界框被標註。其標註資料主要採用JSON格式。

<https://cocodataset.org/>

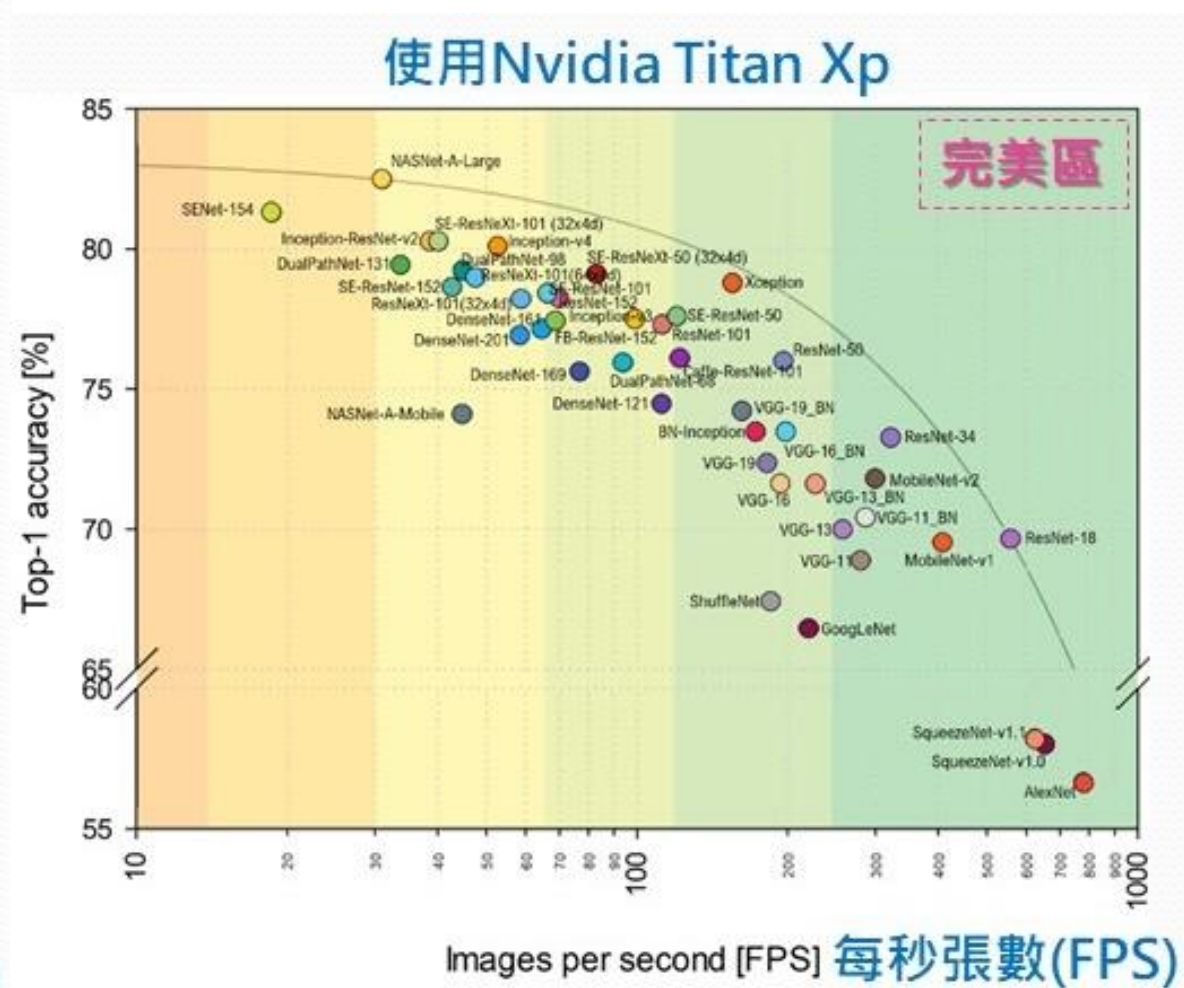
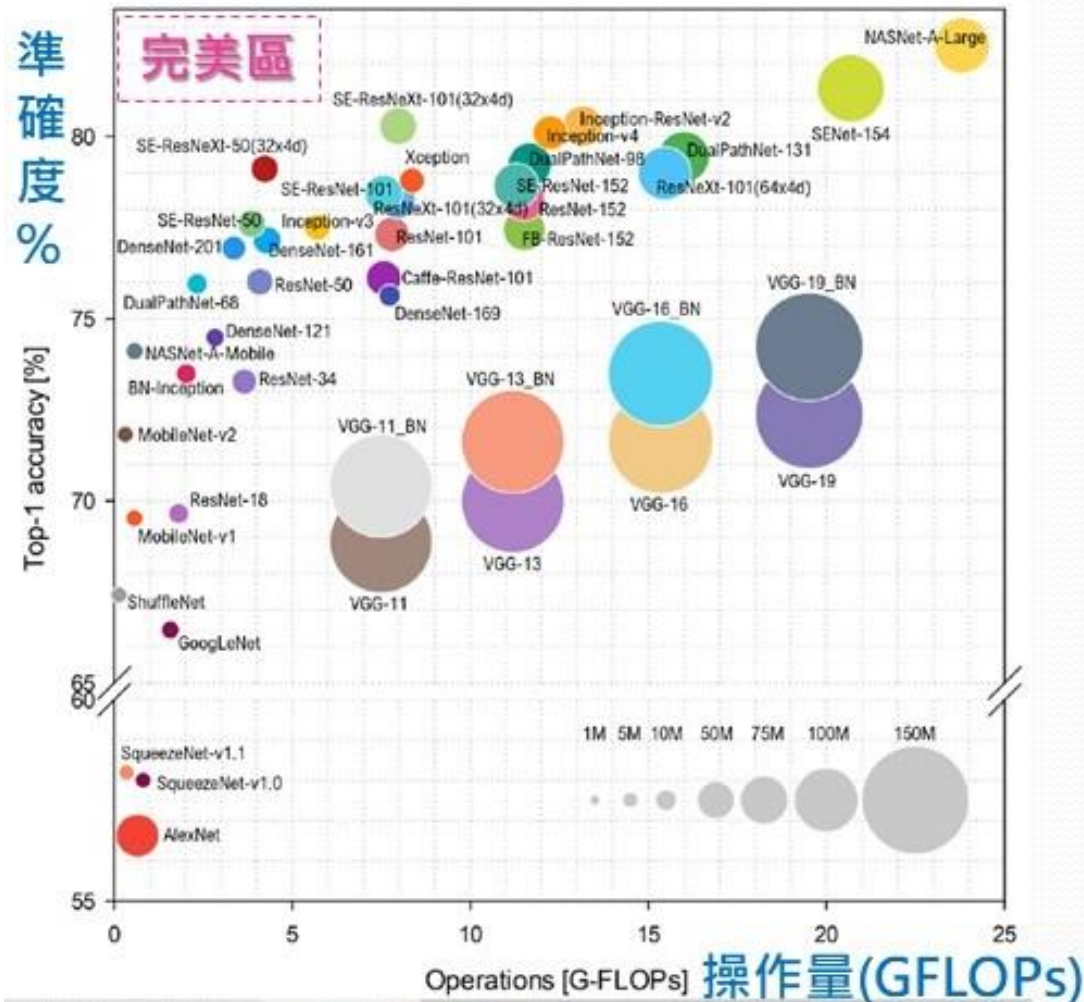
影像分類模型

- AlexNet
- DenseNet
- EfficientNet (v1,v2)
- Inception (v1,v2,v3,v4)
- MobileNet (v1,v2,v3)
- ResNet 18 (26, 34, 50)
- Shufflenet (v1,v2)
- Squeezenet
- VGG 16 (19)



影像來源：<https://arxiv.org/abs/1901.06032>

操作量 vs. 準確度 vs. 推論速度



影像來源：<https://arxiv.org/abs/1810.00736>

影像來源：<https://arxiv.org/abs/1810.00736>

Intel OpenVINO 範例來源

更多OpenVINO內容請參考
第5講開源模型推論工具

▼ LEARN OPENVINO

- Interactive Tutorials (Python)
- Sample Applications (Python & C++)
- Large Language Model Inference Guide

▼ Legacy Features

- OpenVINO Development Tools package
- Model Optimizer / Conversion API
- Open Model ZOO

2021.4版後提供
Notebooks
(本機/雲端Colab)

2023.0版後從主程式獨立出來
Open Model Zoo
(淡出OpenVINO)

▼ Open Model ZOO

Overview of OpenVINO™
Toolkit Intel's Pre-Trained
Models

Overview of OpenVINO™
Toolkit Public Pre-Trained
Models

Model Downloader and other
automation tools

Deep Learning accuracy
validation framework

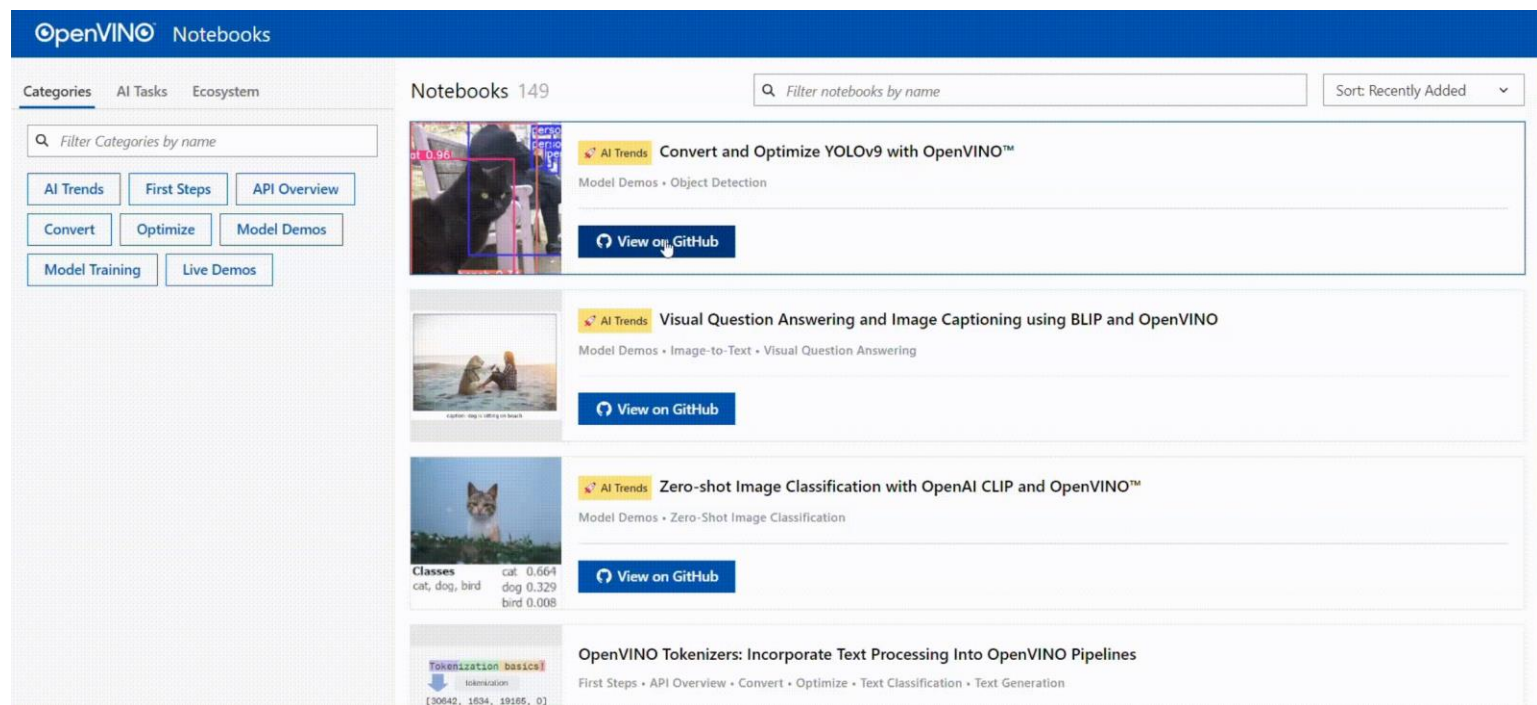
Dataset Preparation Guide

Open Model Zoo Demos

OpenVINO Model Server
Adapter

OpenVINO Notebooks 新版網頁簡介

2024.0 Notebooks



https://openvinotoolkit.github.io/openvino_notebooks/

2022.1 ~ 2023.3 Notebooks

First steps (001~099)
Covert & Optimize (100~199)
Model Demos (200~299)
Model Training (300~399)
Live Demos (400~499)

2024.0 Notebooks (取消編號)

可依分類、AI工作項目及生態系統方式直接查詢，或輸入關鍵字查詢，大幅改善工作效率。

OpenVINO Notebooks — Image Classification

The screenshot displays the OpenVINO Notebooks web interface. The top navigation bar includes the OpenVINO logo and the word 'Notebooks'. Below this, there are tabs for 'Categories', 'AI Tasks' (highlighted with a red box and a '1' badge), and 'Ecosystem'. A search bar is present with the placeholder text 'Filter AI Tasks by name'. On the left sidebar, under the 'Multimodal' section, there are buttons for 'Text-to-Image', 'Image-to-Text', 'Text-to-Video', 'Video-to-Text', 'Text-to-Audio', 'Audio-to-Text', 'Visual Question Answering', 'Image Captioning', 'Feature Extraction', 'Text-to-Image Retrieval', 'Image-to-Text Retrieval', 'Image-to-3D', 'Image-to-Video', and 'Computer Vision' (highlighted with a red box). Under 'Computer Vision', the 'Image Classification' button is also highlighted with a red box. The main content area shows a list of notebooks. The first notebook is 'Hello NPU' with the subtitle 'API Overview • Image Classification' and a 'View on GitHub' button. The second notebook is 'Big Transfer Image Classification Model Q' with the subtitle 'Optimize • API Overview • Image Classification' and a 'View on GitHub' button. The third notebook is 'OpenVINO™ model conversion API' with the subtitle 'Convert • API Overview • Image Classification • Text C' and a 'View on GitHub' button. On the right side, there are three overlapping notebook preview cards. The top card is 'Hello Image Classification' with the subtitle 'First Steps • Image Classification' and buttons for 'View on GitHub', 'Open in Colab', and 'Launch in Binder'. The middle card is 'Convert a PyTorch Model to OpenVINO™ IR' with the subtitle 'Convert • Image Classification' and buttons for 'View on GitHub' and 'Open in Colab'. The bottom card is 'Video Recognition using SlowFast and OpenVINO™' with the subtitle 'Model Demos • Object Detection • Image Classification' and a 'View on GitHub' button. Below these cards, there are two more preview cards: 'Convert a Tensorflow Lite Model to OpenVINO™' with the subtitle 'Convert • Image Classification' and buttons for 'View on GitHub' and 'Open in Colab', and another 'Hello Image Classification' card with the same subtitle and buttons.

https://openvinotoolkit.github.io/openvino_notebooks/

OpenVINO OMZ Public 模型挑選

Public Pre-Trained Models

36項

預訓練模型

下載和轉換模型到IR檔(xml, bin)

Classification Models

模型名稱	AI框架	OMZ名稱	推論精度	計算量
Model Name	Implementation	OMZ Model		
↕	↕	Name ↕	Accuracy ↕	GFlops
AlexNet	Caffe*	alexnet	56.598%/79.812%	1.5
AntiSpoofNet	PyTorch*	anti-spoof-mn3	3.81%	0.15
CaffeNet	Caffe*	caffenet	56.714%/79.916%	1.5

點擊OMZ名稱，
查看模型完整說明。

An example of using the Model Downloader:

OMZ名稱
omz_downloader --name <model_name>

An example of using the Model Converter:

OMZ名稱
omz_converter --name <model_name>

Open Model Zoo 2023.0版後
雖不再更新維護但可使用。

資料來源：https://docs.openvino.ai/2023.1/omz_models_group_public.html

OpenVINO OMZ Public 模型說明

alexnet

Use Case and High-Level Description

The **alexnet** model is designed to perform image classification. Just like other common classification models, the **alexnet** model has been pre-trained on the ImageNet image database. For details about this model, check out the [paper](#).

The model input is a blob that consists of a single image of **1, 3, 227, 227** in **BGR** order. The BGR mean values need to be subtracted as follows: [104, 117, 123] before passing the image to the model.

Specification

Metric	Value
Type	Classification
GFLOPs	1.5
MParams	60.965
Source framework	Caffe*

Accuracy

Metric	Value
Top 1	56.598%
Top 5	79.812%

Input

Original model

Image, name - **data**, shape - **1, 3, 227, 227**, format is **B, C, H, W**, where:

- **B** - batch size
- **C** - channel
- **H** - height
- **W** - width

Channel order is **BGR**. Mean values -

Converted model

Image, name - **data**, shape - **1, 3, 227, 227**, format is **B, C, H, W**, where:

- **B** - batch size
- **C** - channel
- **H** - height
- **W** - width

Channel order is **BGR**.

Output

Original model

Object classifier according to ImageNet classes, name - **prob**, shape - **1, 1000**, output data format is **B, C**, where:

- **B** - batch size
- **C** - predicted probabilities for each class in [0, 1] range

Converted model

Object classifier according to ImageNet classes, name - **prob**, shape - **1, 1000**, output data format is **B, C**, where:

- **B** - batch size
- **C** - predicted probabilities for each class in [0, 1] range

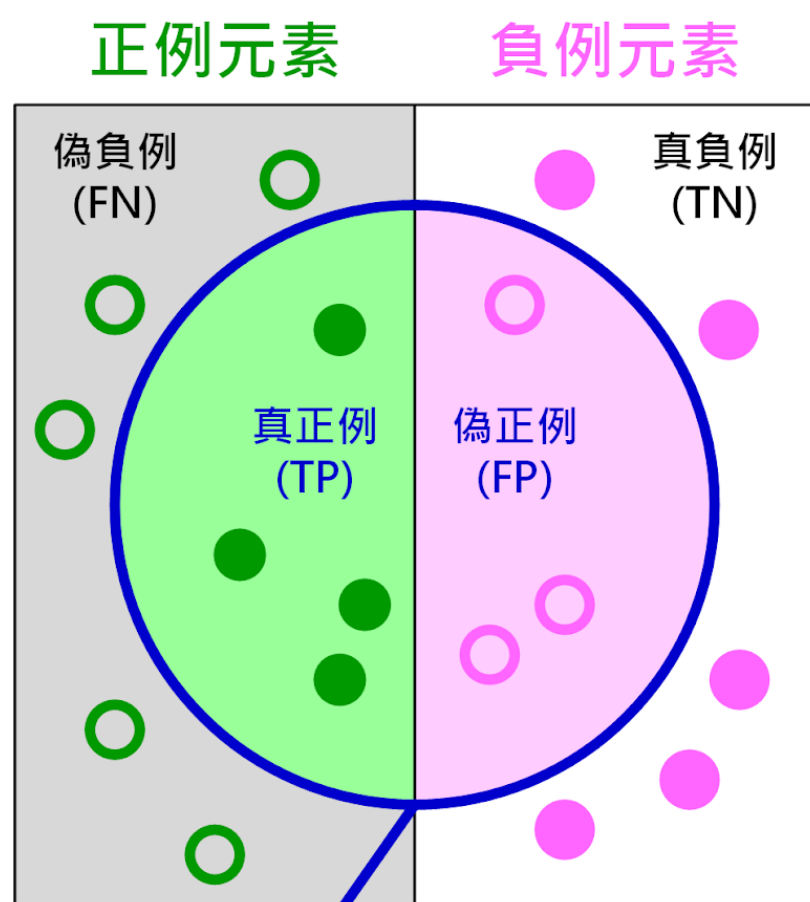
Demo usage

The model can be used in the following demos provided by the Open Model Zoo to show its capabilities:

- [Classification Benchmark C++ Demo](#)
- [Classification Python* Demo](#)

資料來源：https://docs.openvino.ai/2023.1/omz_models_model_alexnet.html

影像分類評量方式 – mAP



被提取元素

TP : True Positives
TN : True Negatives

FP : False Positives
FN : False Negatives

(真正例 + 真負例) / 總樣本數

正確率
(Accuracy)



真正例 / 被提取總數

精確率
(Precision)



真正例 / 正例元素總數

召回率
(Recall)



$$\frac{1}{F_1} = \frac{1}{P} + \frac{1}{R} \rightarrow F_1 = \frac{2PR}{P + R} = \frac{2TP}{2TP + FP + FN}$$

mean Average Precision

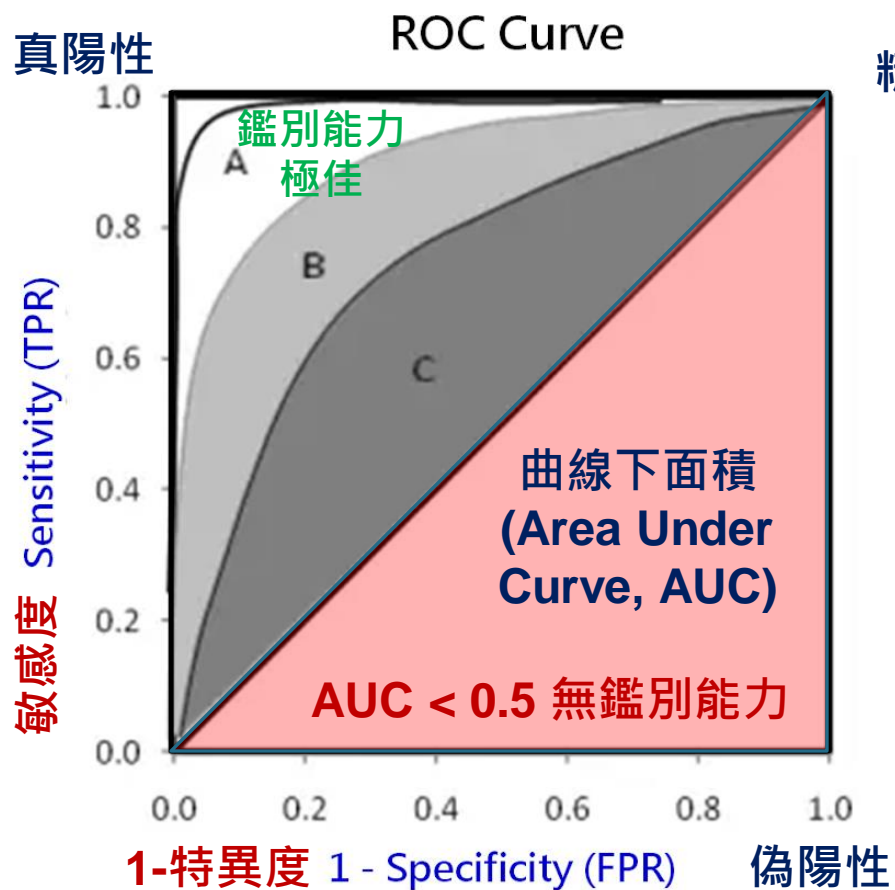
$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k$$

AP_k = the AP of class k
 n = the number of classes

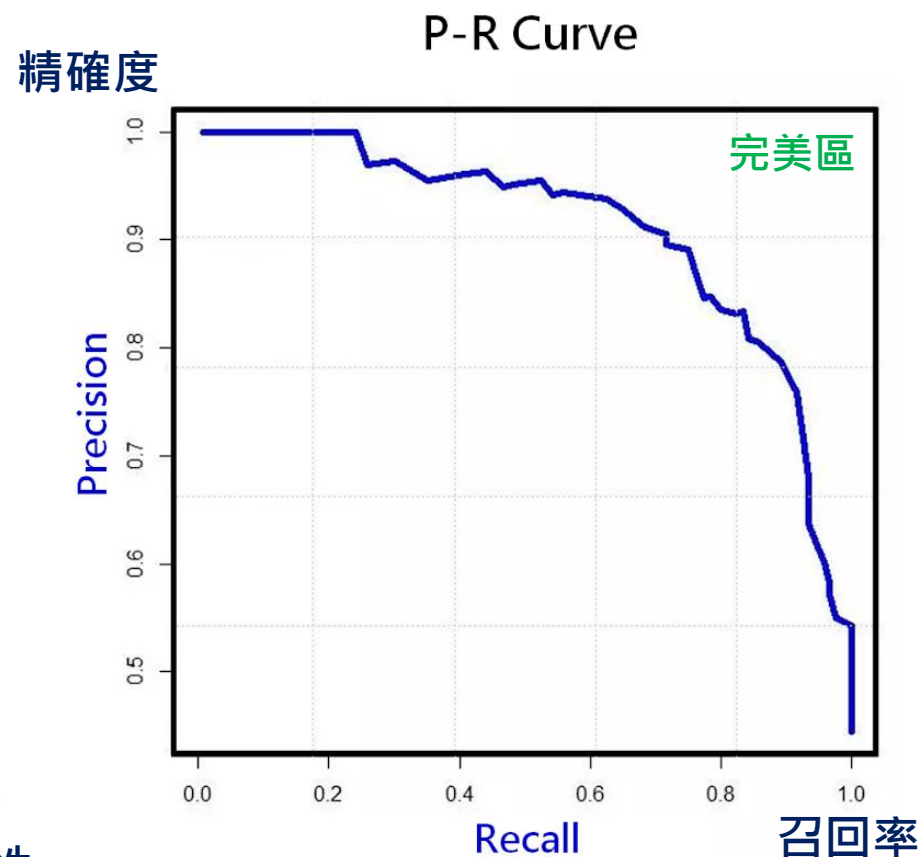
圖片來源：<http://omnixri.blogspot.com/2019/07/aigoai.html>

影像分類評量方式 — ROC / PR曲線

Receiver Operating Characteristic Curve



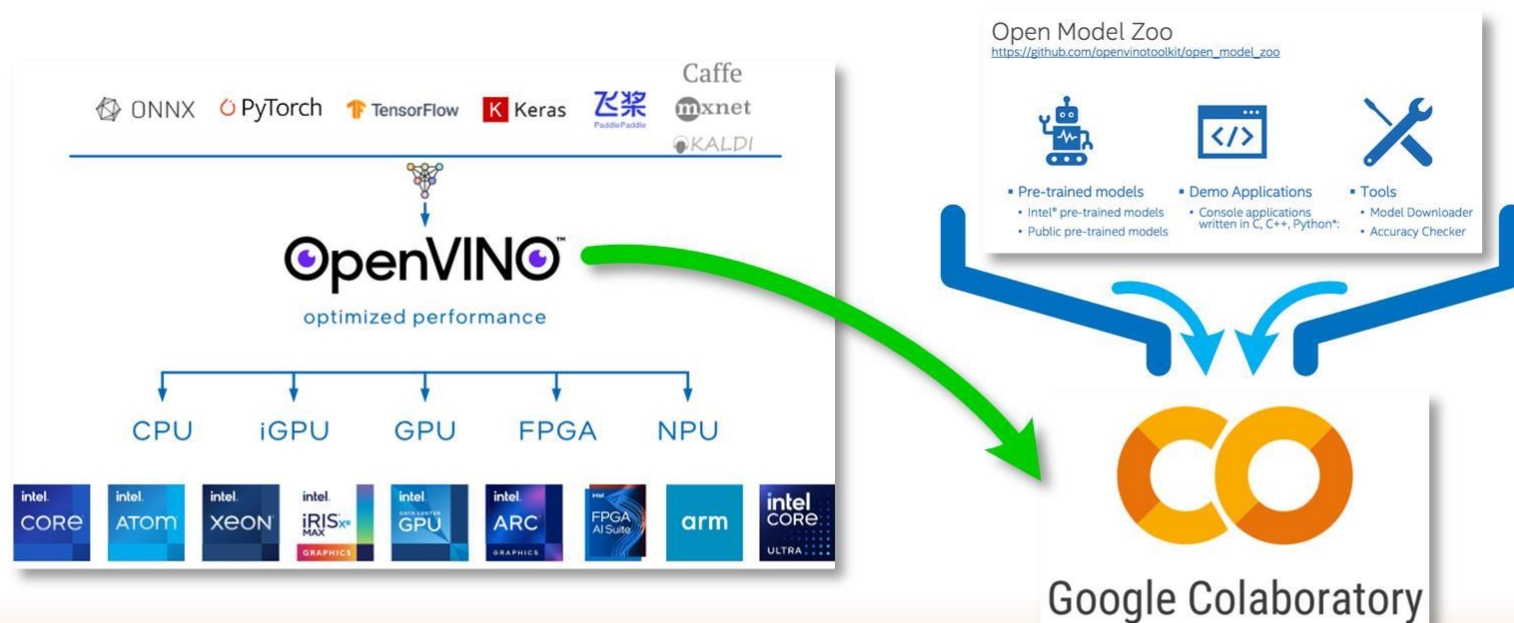
Precision-Recall Curve



適合類別不平衡資料集

圖片來源：<http://omnixri.blogspot.com/2019/07/aigoai.html>

OpenVINO Open Model Zoo Colab範例



如何運行Intel OpenVINO Open Model Zoo
(OMZ) 範例於Google Colab上

2023.1版後移入
「Legacy Features」

- * Open Model Zoo
- * Model Downloader
- * Model Converter
- * Info Dumper
- * Benchmark Tool

<https://omnixri.blogspot.com/2024/02/intel-openvino-open-model-zoomzgoogle.html>

Colab Open Model Zoo 影像分類範例

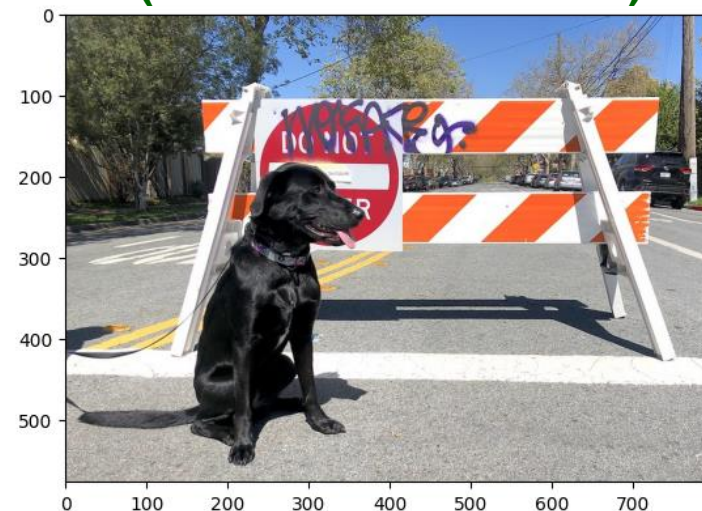
工作流程

1. 安裝openvino-dev
2. 導入必要模組
3. 下載預訓練模型
4. 模型轉換
5. 指定推論裝置
6. 將模型載入裝置並編譯
7. 下載測試資料

整合原Notebooks
001-hello-world
104-model-tools

https://colab.research.google.com/github/OmniXRI/OpenVINO2023LTS_on_Colab_Samples/blob/main/OpenVINO_2023_LTS_Demo_Classification.ipynb

平毛尋回犬
(Flat-coated Retriever)



Do Inference

```
result_infer = compiled_model([input_image])[output_layer]
result_index = np.argmax(result_infer)
```

```
# Convert the inference result to a class name.
imagenet_classes = open("../data/datasets/imagenet/imagenet_2012.txt").read().splitlines()

# The model description states that for this model, class 0 is a background.
# Therefore, a background must be added at the beginning of imagenet_classes.
imagenet_classes = ['background'] + imagenet_classes

imagenet_classes[result_index]
```

'n02099267 flat-coated retriever'

'n02099267 flat-coated retriever'

參考文獻

- 許哲豪，臺灣科技大學資訊工程系「人工智慧與邊緣運算實務」（2021~2023）
<https://omnixri.blogspot.com/p/ntust-edge-ai.html>
- 許哲豪，如何運行Intel OpenVINO Open Model Zoo（OMZ）範例於Google Colab上
<https://omnixri.blogspot.com/2024/02/intel-openvino-open-model-zoomzgoogle.html>

沒有最邊



只有更邊

歡迎加入
邊緣人俱樂部



歐尼克斯實境互動工作室
(OmniXRI Studio)

許哲豪 (Jack Hsu)



Facebook : Jack Omnixri

FB社團 : Edge AI Taiwan邊緣智能交流區

電子信箱 : omnixri@gmail.com

部落格 : <https://omnixri.blogspot.tw>

開 源 : <https://github.com/OmniXRI>

YOUTUBE 直播 : <https://www.youtube.com/@omnixri1784/streams>