THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

KUDOS DATA

Practical Data Science

Predictive Image Analysis Using Deep Transfer Learning

Instructor
Sam.Gu@KudosData.com

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

KUDOS DATA

Agenda

- Why Data Scientists Are Highly Paid ?
- Data Science vs. Machine Learning
- Ethical Challenges in Data Science & Prediction
- Predictive Image Analysis Using Deep Transfer Learning

Why Data Scientists Are Highly Paid ?

KUDOS DATA

Biz Domain Knowledge

Subject Matter Expert
($70k+)

Modeling Knowledge

Traditional Analyst
($60k+)

Programming Technical Skills

Software Engineer
($50k+)

...

Taste (Market)

Recipe (Model)

Sourcing (Software)

Cooking (Scaling)

...

Biz Strategy
Sales & Market
Finance
Competition

Biz Data Scientist
($90k+)

Data Science Engineer
($80k+)

Prod Support
Customer Care
Operation
Miscellaneous

Full Stack Data Scientist
($130k+)

Entrepreneur Data Scientist
($???)

Why Data Scientists Are Highly Paid ?

KUDOS DATA

A highly paid data scientist is a superman capable of:

- modeling using machine learning
- writing quality code (full stack / web master)
- designing data driven software architecture
- industrial domain knowledge
- explanation to layman business sponsor/customer (visualization and storytelling)
- workaround organizational red-tapes (even hostility)

Data Science vs. Machine Learning

KUDOS DATA

Data Science Elements:

- Framing Business Problem
- Data Acquisition & Preparation
- Exploratory Data Analysis
- Data-driven Modeling (machine learning, tools & techniques)
- Result Presentation & Visualization (actionable insights)

Machine Learning Elements :

- Task
- Experience (data)
- Performance

Ethical Challenges in Data Science & Prediction

KUDOS DATA

- New technology makes people's life miserable ?
- Sales Prediction
- Education / Medical Resource Allocation
- Data Science = Data-driven Discrimination ?

Predictive Image Analysis Using Deep Transfer Learning

KUDOS DATA

1. Deep Learning Basics for Image Analysis
2. Real World Image Analysis Needs
3. Idea of Transfer Learning
4. Architecture of Transfer Learning

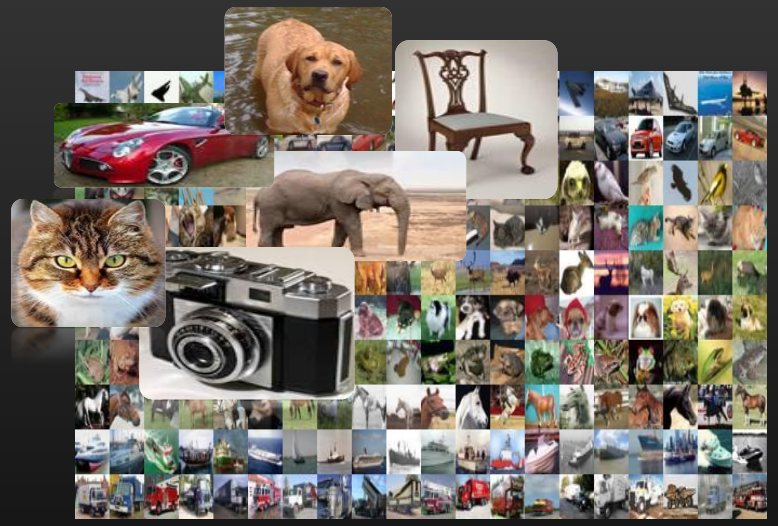
---------- Short break / Workshop environment setup ----------

5. Hands-on Datalab Workshop on Google Cloud Platform (GCP)

1 Deep Learning Basics for Image Analysis (1)

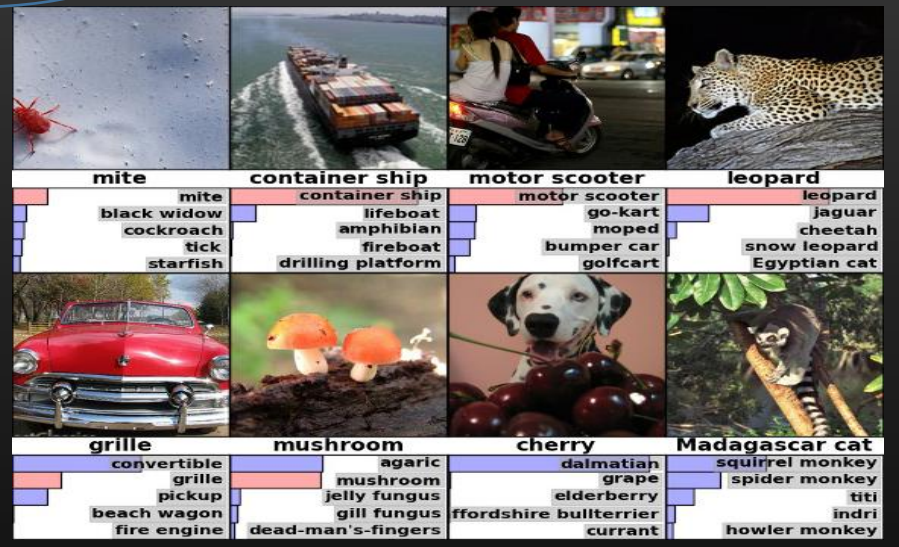
KUDOS DATA

Please! Tell me the correct image category. **And how you did it?!**



Huge images as learning inputs (millions)

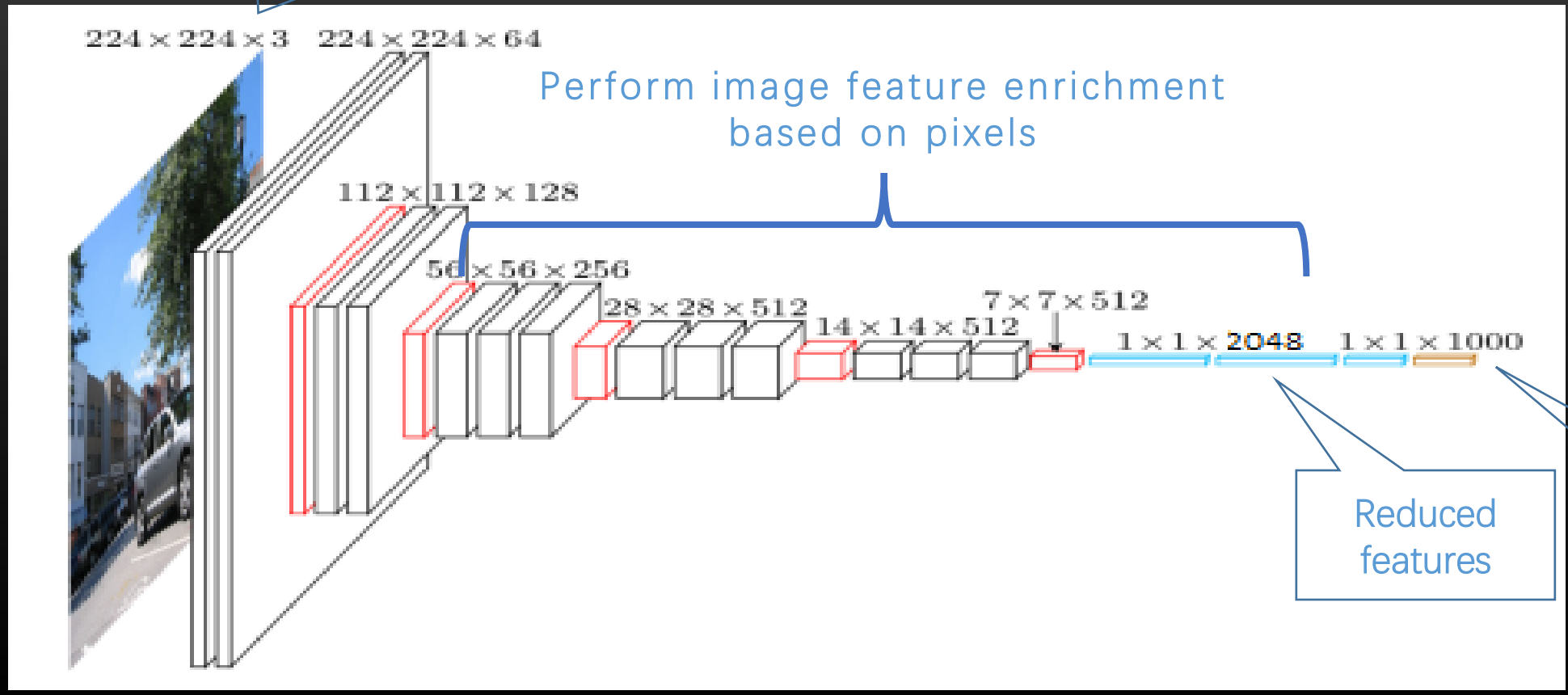
Deep Learning Magic Crystal Ball



Very accurate predicted class (1000 classes)

1 Deep Learning Basics for Image Analysis (2)

150,528 features per image
(224 x 224 pixels x 3 RGB channels)



2 Real World Image Analysis Problem

KUDOS DATA

In my workplace, I have some images of **healthy working valves** and **malfunctioned ones**. I'd like to explore the promising benefit of automatically monitor, classify and alert me if there is any anomaly happening in the industrial fields…

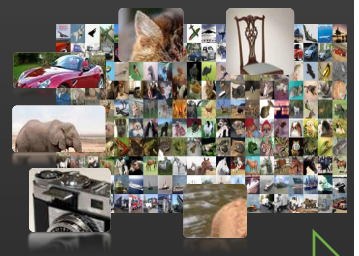



But I

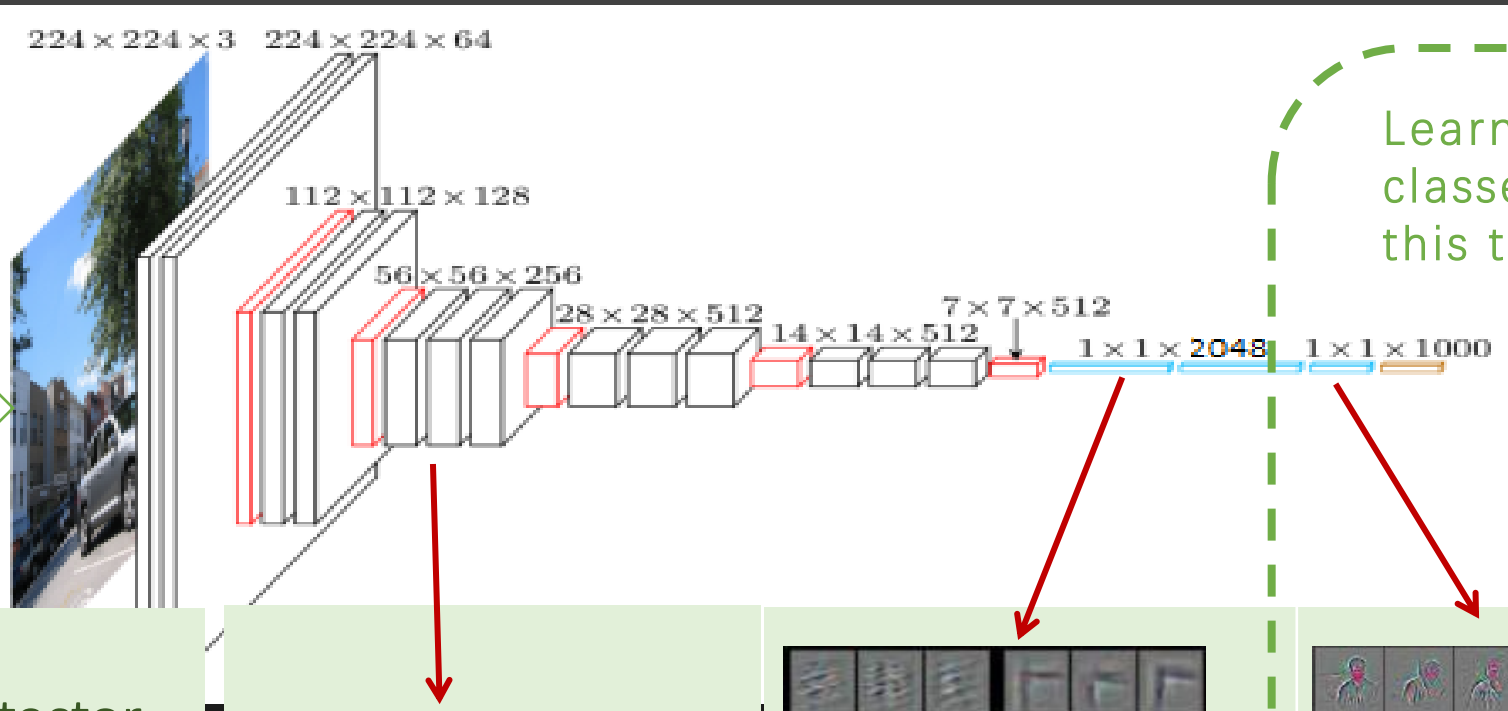
- Lack of powerful computing resources
- Difficulty to design network architecture and tune millions of hyper-parameters of deep neural network
- Lack of enough labeled images (around one hundred images)

Is there a solution?

3 Idea of Transfer Learning



Task of ImageNet

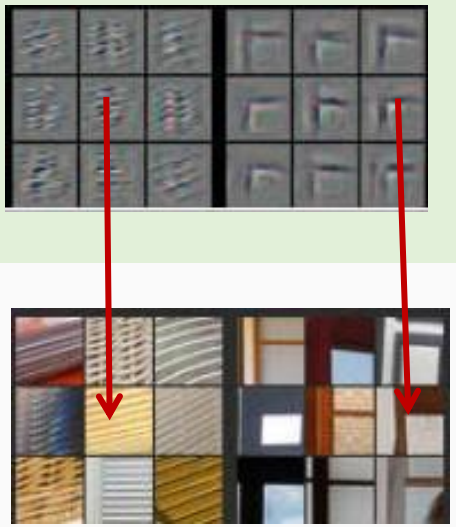


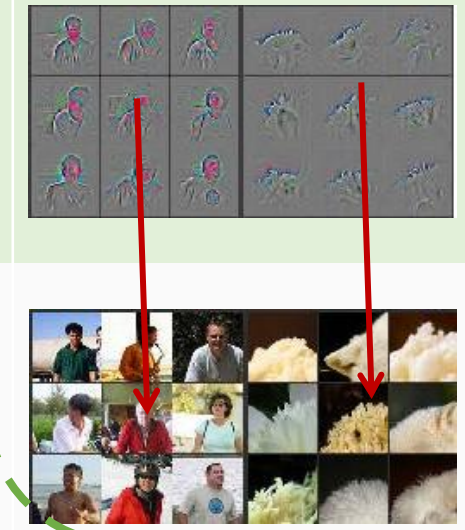
Learn and predict 1000 classes: Very specific to this task

Example of detector learnt in deep neural net

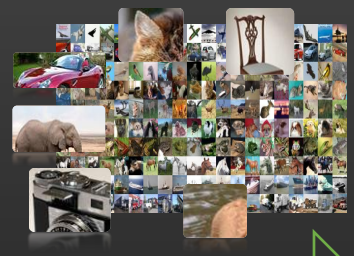
Example of matched pattern from individual image

[Zeiler & Fergus]

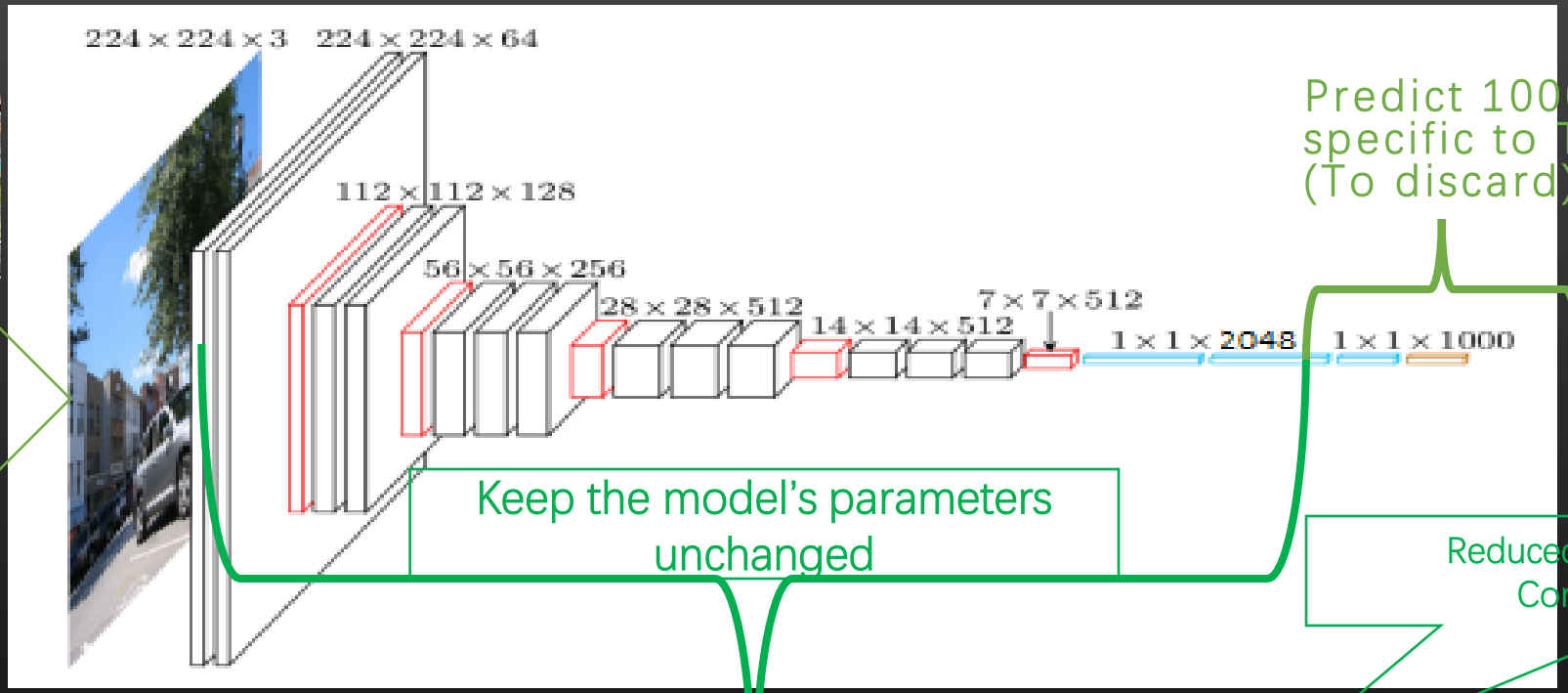




4 Architecture of Transfer Learning



Task1
ImageNet



Predict 1000 classes: Very specific to Task1 (To discard)

Keep the model's parameters unchanged

Reduced features per image: 2048
Concentration Ratio: 1.4%

Task2 Valves

Raw features per image: 150,528

Feature Extraction / Transformation (Reusable for Task2)

New Simple ML Model

Normal Valve

Abnormal Valve

5 Hands-on Datalab Workshop on GCP (1)

KUDOS DATA

Overview: In this lab, you will carry out a transfer learning example based on Google Inception-v3 image recognition neural network.

You will learn:

1. Explore images in customer's industry.
2. Reposition a pre-trained deep neural net for new image recognition task.
3. Perform feature extraction.
4. Obtain deep feature representation of customer's original image.
5. Train a simple machine learning model for new classification task.
6. Evaluate results of this transfer learning model.

Prerequisites:

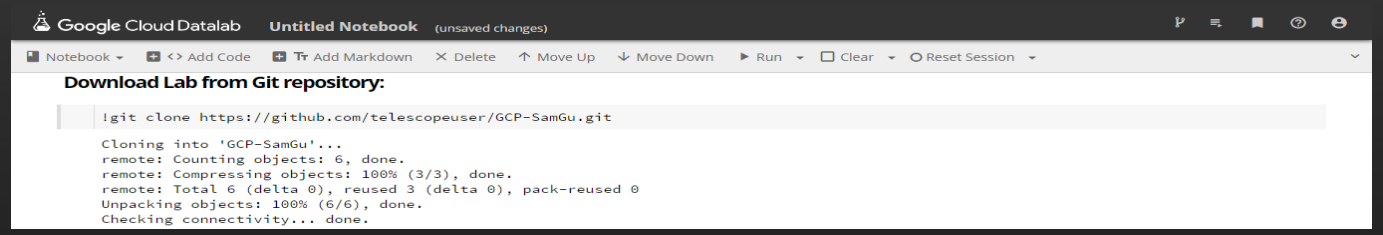
- Google Cloud Platform Account
- Basic working knowledge of GCP, Datalab and Python

5 Hands-on Datalab Workshop on GCP (2)

KUDOS DATA

Steps to access lab workshop:

1. Login Google Cloud Platform to start Datalab.
2. Create a new notebook to download this lab by running command: !git clone https://github.com/telescopeuser/GCP-SamGu.git



3. Go to folder GCP-SamGu/Lab/, then open notebook Lab_Image_Analysis.ipynb to follow.

Reference:

- https://github.com/telescopeuser/GCP-SamGu
- Google Cloud Platform Free Registration
- Google Datalab Quick Start

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

KUDOS DATA

Thank you !
And you can help us by

Star our projects at:
https://github.com/telescopeuser

Write to participate in real world projects to democratize AI:
Sam.Gu@KudosData.com

WeChat