Problem Description and Pipeline

Case study focused around an application called **Photo OCR**

There are hree reasons to see this:

1. **Look at how a complex system can be put together**
2. **The concept of a machine learning pipeline**

   - How to allocate resources when we're trying to decide what to do next.
3. **Some ideas of how to apply machine learning to computer vision problems**

   - Applying machine learning to tangible problems
   - **Artificial data synthesis**

The Photo OCR problem

Photo OCR stands for: **Photo Optical Character Recognition**

With the growth of digital photography and more recently the growth of camera in our cell phones we now have **tons of visual pictures** that we take all over the place, and one of the things that has interested many developers is how to get our computers to **understand the content** of these pictures a little bit better.

The photo OCR problem focuses on how to get computers to read the text to the purest in images that we take.

Given an image like this it might be nice if a computer can read the text in this image so that if we're trying to look for this picture again we type in the words, Lula B's and and have it automatically pull up this picture, so that we're not spending lots of time digging through our photo collection.

The photo OCR problem is getting computers to read text in an image:

- One Possible application for this would include:
  - Make searching easier (e.g. searching for photos based on words in them)
- OCR of documents is a comparatively easy problem
  - From photos it's really hard

Photo OCR pipeline

OCR pipeline

1. **Text detection**

2. **Character segmentation**

3. **Character classification**

Optional some may do spell check after this too (we're not focussing on such systems though). Here's a picture showing the photo OCR pipeline:

We have an **image**, which then fed to the **text detection system** text regions, we then **segment out the characters** (the individual characters in the text) and then finally we **recognize the individual characters**.

- **Pipelines are common in machine learning**
  - Separate modules which may each be a **machine learning component** or **data processing component**
- If we're designing a machine learning system, pipeline design is one of the **most important questions**
  - **Performance** of pipeline and each module often has a **big impact on the overall performance a problem**
  - We would often have different engineers working on each module
    - Offers a natural way to **divide up the workload**

Video Question: When someone refers to a "machine learning pipeline," he or she is referring to:

- A PhotoOCR system.
- A character recognition system.

> A system with many stages / components, several of which may use machine learning.

- An application in plumbing. (Haha.)