Applications

Module 1: Computer Vision

AI Camp - Scholarships (due by EOM)

"AI Camp (www.ai-camp.org) is offering 10 scholarships for their Introduction to AI lecture series. This lecture series is the same series that offered at Stanford University this summer and will introduce AI and careers in tech to you. No coding experience is needed. The opportunity is first-come, first serve, and the deadline for registration is at the end of this month. You can register it here: www.ai-camp.org/free-lesson. AI Camp is founded by Stanford Ph.Ds and advised by a Stanford Professor and CTOs in Silicon Valley."

A bit more about AI Camp:

1. **AI Camp was founded by Stanford Ph.Ds, and our instructors are from Stanford and Carnegie Mellon University**. Our mission is to bring AI education to all middle and high schools across the nation.
2. Our AI curriculum is specially crafted for middle and high school students. **In just one year, we have taught more than 2,000 students for this introductory class.**
3. Learning AI can spark students' interest in learning about math, coding, critical thinking, and science. **AI is the biggest technology tidal wave in our student's lifetime.**

The deadline for the scholarships is at the end of this month.

"

What are some applications of CNNs within the field of computer vision that you have heard of?

Applications (this list is not all-inclusive!)

◎ Facial Recognition + Emotion Detection
◎ Optical Character Recognition (OCR)
◎ Action Classification
◎ Spam Detection
◎ Colorization
◎ Medical Diagnosis
◎ **And many others!**

Facial Recognition + Emotion Detection

◎ Problem of identifying and verifying people in a photo by their face

◎ **4-step pipeline**
- ○ Detection: Locate one or more faces in the image and mark with a bounding box
- ○ Alignment: Normalize the face to be consistent with the database
- ○ Feature Extraction: Extract features from the face that can be used for the recognition task
- ○ Recognition: Perform matching of the face against one or more known faces in a prepared database.

◎ Use cases: service authentication, passport verification, etc.
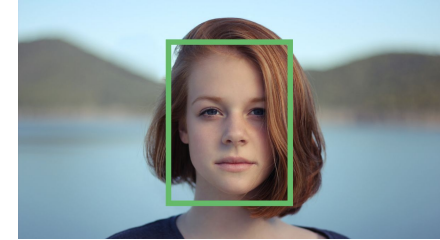
(a)   (b)   (c)

OCR

◎ Conversion of images of typed, handwritten or printed text into machine-encoded text

◎ **Popular precoded frameworks**
   ○ Google Vision
   ○ Amazon Rekognition
   ○ Tesseract

◎ Use cases: data entry, passport information extraction, traffic sign detection (self-driving cars), etc.


```
Store #05666
3515 DEL MAR HTS,RD
SAN DIEGO, CA  92130
(858) 792-7040

Register #4 Transaction #571140
Cashier #56661020  8/20/17  5:45PM

      wellness+ with Plenti
    Plenti Card#: 31XXXXXXXXXX4553
1 G2 RETRACT BOLD BLK 2PK      1.99 T
    SALE 1/1.99, Reg 1/4.69
    Discount  2.70-

1 Items          Subtotal      1.99
                      Tax       .15
                    Total      2.14
                               2.14
*MASTER*
    MASTER card * #XXXXXXXXXXXX5485
    App #AA APPROVAL AUTO
    Ref # 05639E
    Entry Method: Chip
```

OCR Receipt Example

Output
Store #056663515
DEL MAR HTS,RD
SAN DIEGO, CA 92130
(858) 792-7040Register #4 Transaction #571140
Cashier #56661020 8/20/17 5:45PMwellnesst+ with Plenti
Plenti Card#: 31XXXXXXXXXX4553
1 G2 RETRACT BOLD BLK 2PK 1.99 T
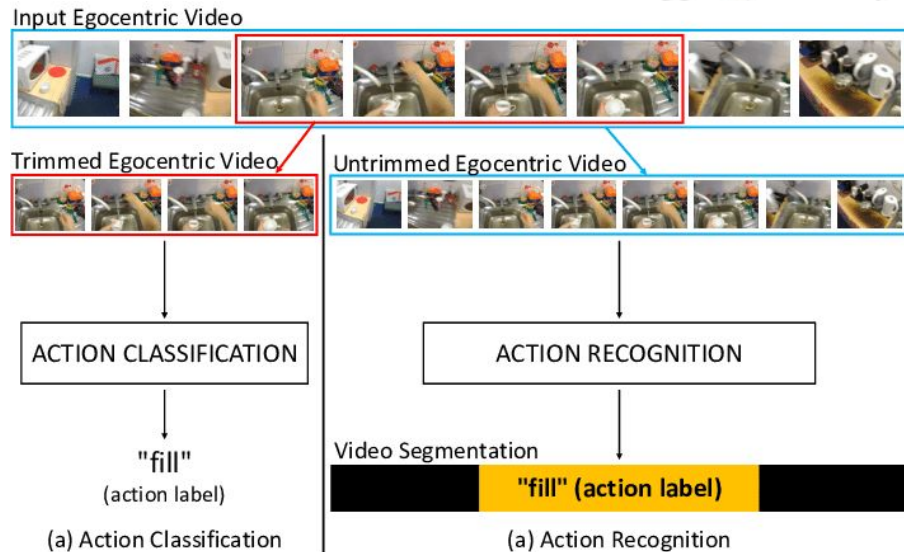SALE 1/1.99, Reg 1/4.69
Discount 2.70-

1 Items Subtotal 1.99
Tax .15

Total 2.14
xMASTER 2.14
MASTER card * #XXXXXXXXXXXX548S
Apo #AA APPROVAL AUTO
Ref # 05639E
Entry Method: Chip
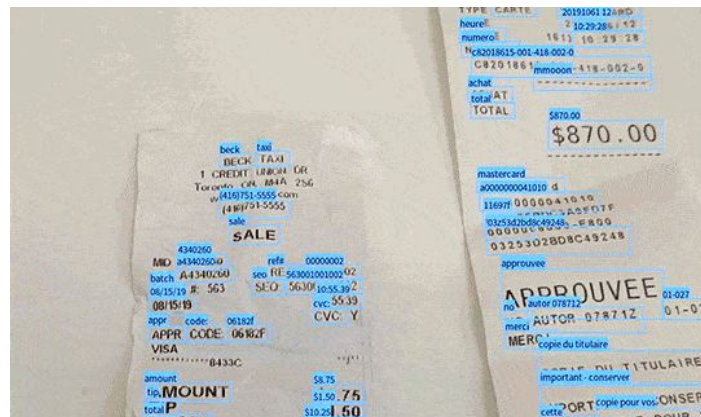
Action Classification

◎ Using video instead of images
◎ Features are extracted w/ temporal convolutional network
- ○ Time as another dimension
- ○ Sequence of frames → fixed size vector
◎ Simple classification afterward
◎ Some examples of well-known models are S3D and I3D
- ○ 3D stands for the temporal aspect of it

Input Egocentric Video

Trimmed Egocentric Video

Untrimmed Egocentric Video

ACTION CLASSIFICATION

ACTION RECOGNITION

"fill"
(action label)

(a) Action Classification

Video Segmentation

"fill" (action label)

(a) Action Recognition

Spam Detection

◎ Problem of determining whether or not an email should be classified as spam

◎ Image spam vs. text spam
  ○ Keyword-based analysis (text extraction + text spam filter) + **image classification**

◎ Often considered a logistic regression task, but spam-not spam classification is a common task for CNNs
 ○ DeepCapture, DeepImageSpam, SMSSpam

Colorization

◎ **Very difficult CV pixel prediction task**
◎ **Unconstrained**
◎ Importance:
  ○ SEO
  ○ Photorealism & historiography
◎ Samples: Algorithmia
  ○ Relevant publication: Colorful Image Colorization
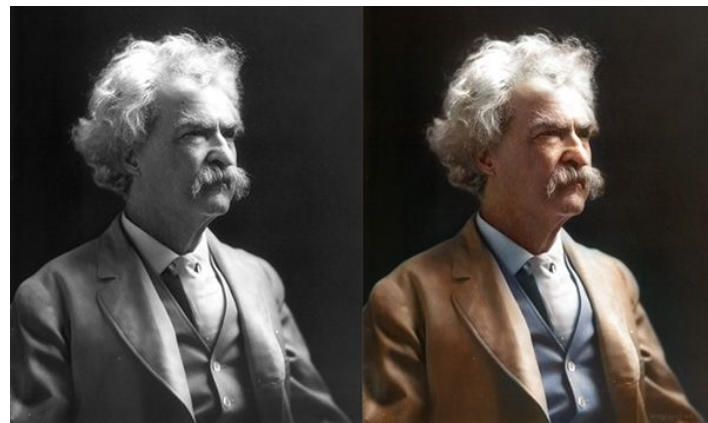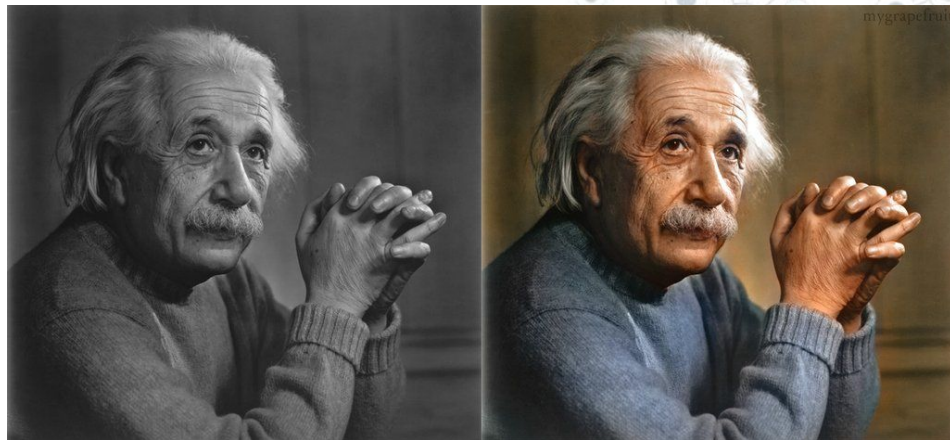
Medical Diagnosis

◎ **Image-based** medical diagnosis
   ○ Problem of determining which disease or condition explains a patient's symptoms (visual signs of illness)
   ○ Most commonly posed as a binary classification problem: given an image and a disease, determine whether or not the image is positive or negative for the disease
◎ **Common examples**: glaucoma diagnosis, brain tumor detection, skin cancer detection (often a multi-class classification problem)
◎ COVID-19 diagnosis!
   ○ DeepCOVID-XR
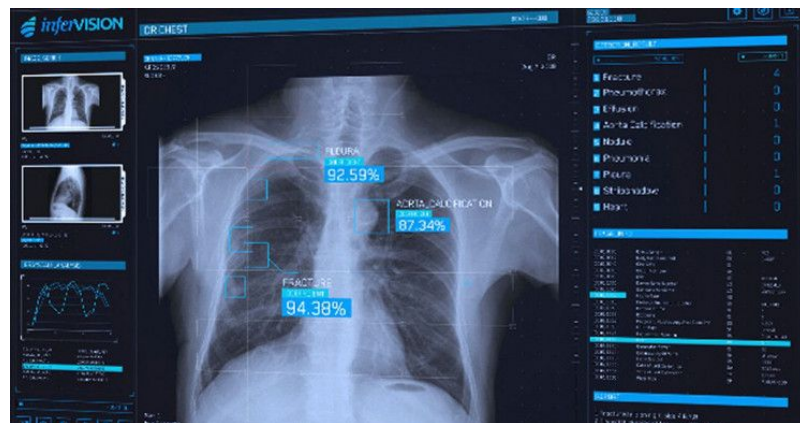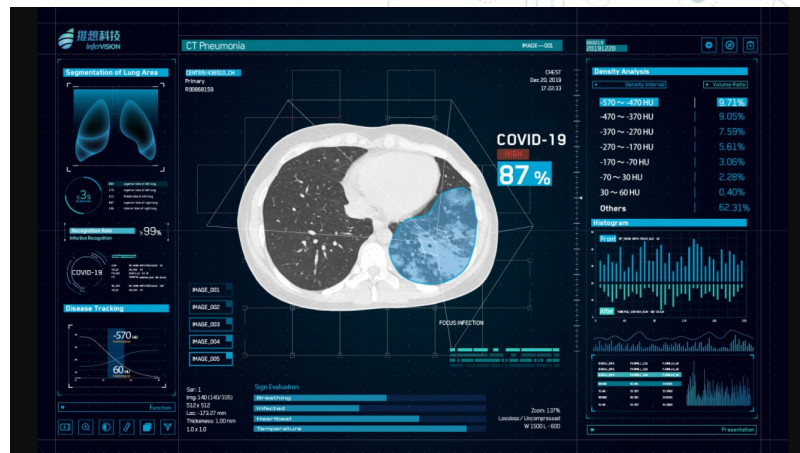
Example!

◎ **Medical imaging: Notebook**
  ○ Classification between chest X-rays and abdominal X-rays
  ○ Raises the idea of domain-specific learning (while domain-agnostic learning is preferrable, pathological features often necessitate task-specific models)
    ◉ Domain-specific multitask AI needs to be able to distinguish between different modalities