CSE 673
COMPUTATIONAL VISION

VENU GOVINDARAJU
DEEN DAYAL MOHAN

University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

Covid-19 Guidelines

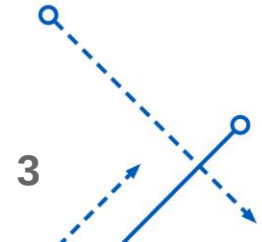
- Effective Aug. 3, the University at Buffalo will require all students, employees and visitors – regardless of their vaccination status – to wear face coverings while inside campus buildings. This includes classrooms, hallways, libraries and other common spaces, as well as UB buses and shuttles.

- Students are expected to wear mask in class during lectures (unless you have a UB approved exception)

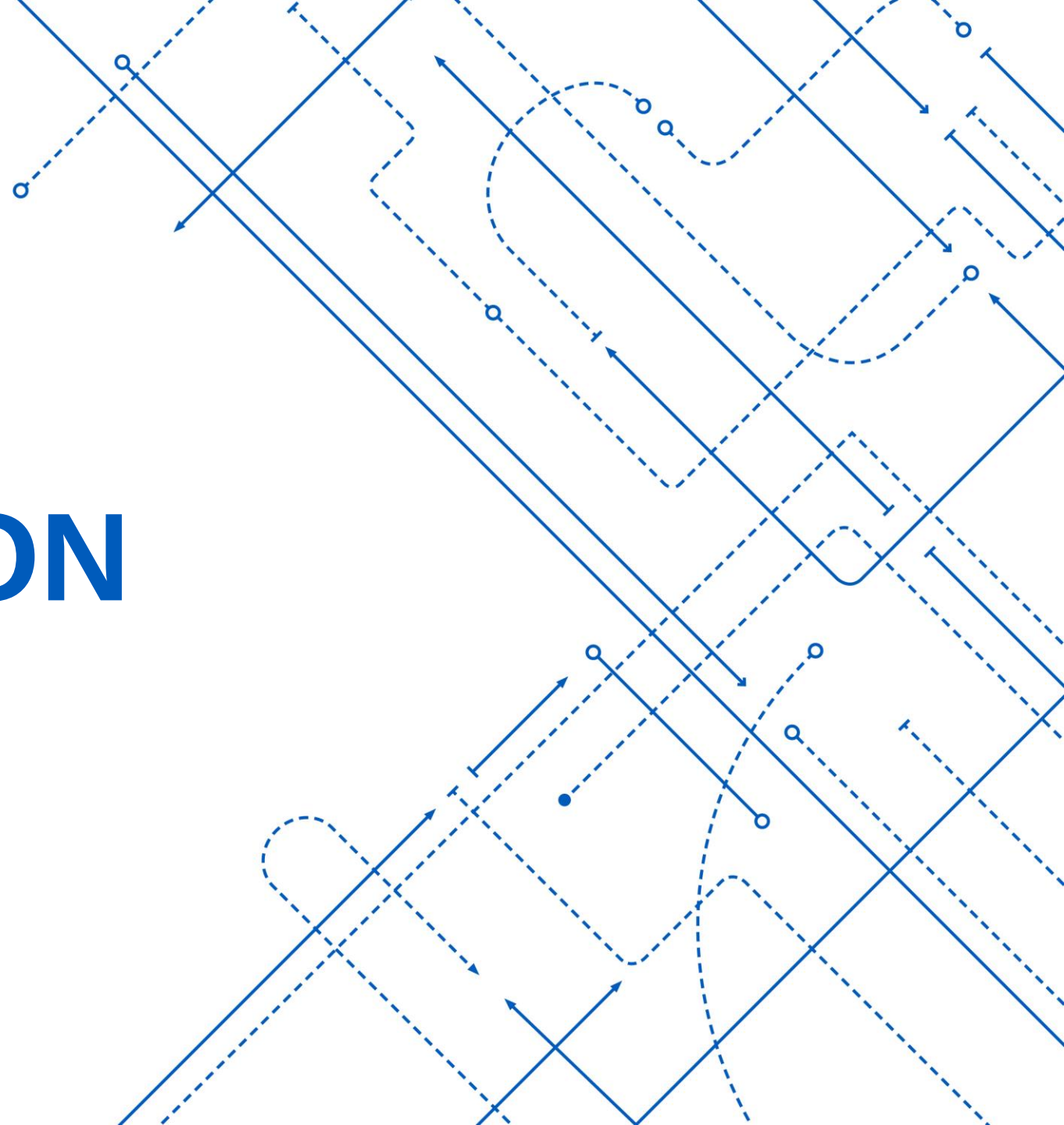
- Public Health Behavior Expectations https://www.buffalo.edu/studentlife/who-we-are/departments/conduct/coronavirus-student-compliance-policy.html

Agenda

- Introduction
- Applications of Computer Vision
- Course Structure and Contents



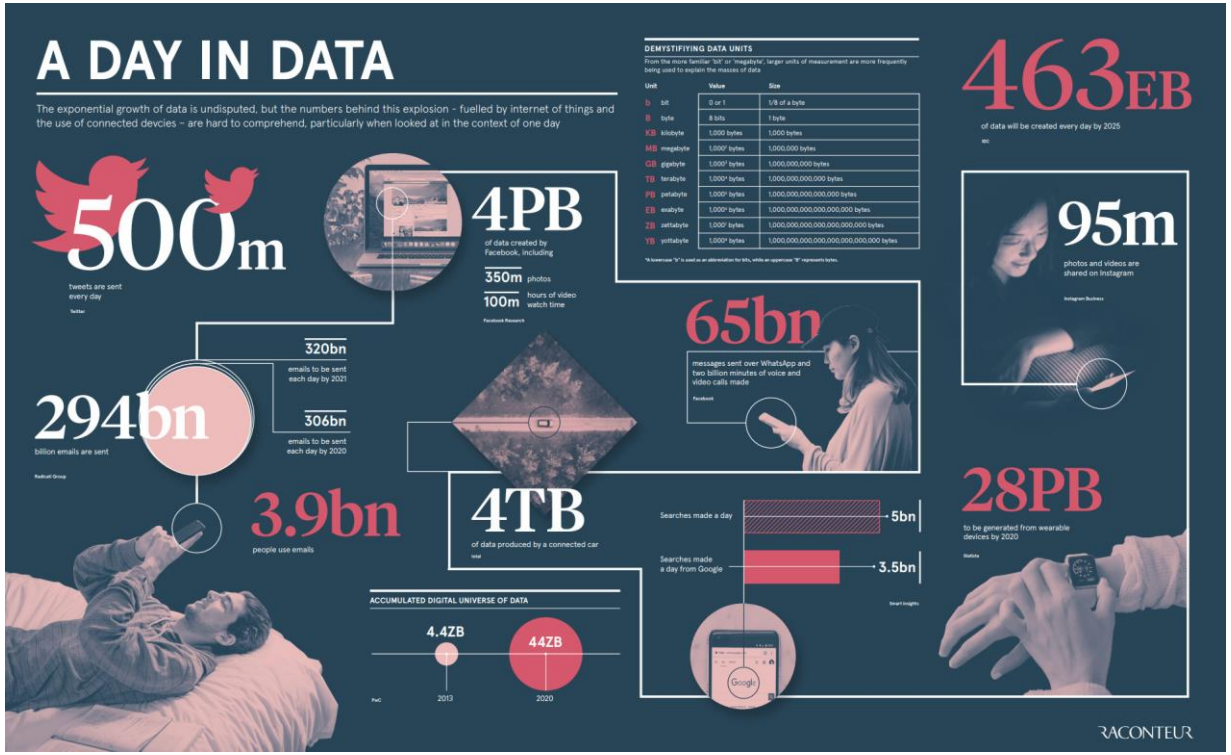
INTRODUCTION



University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

Data is Ubiquitous

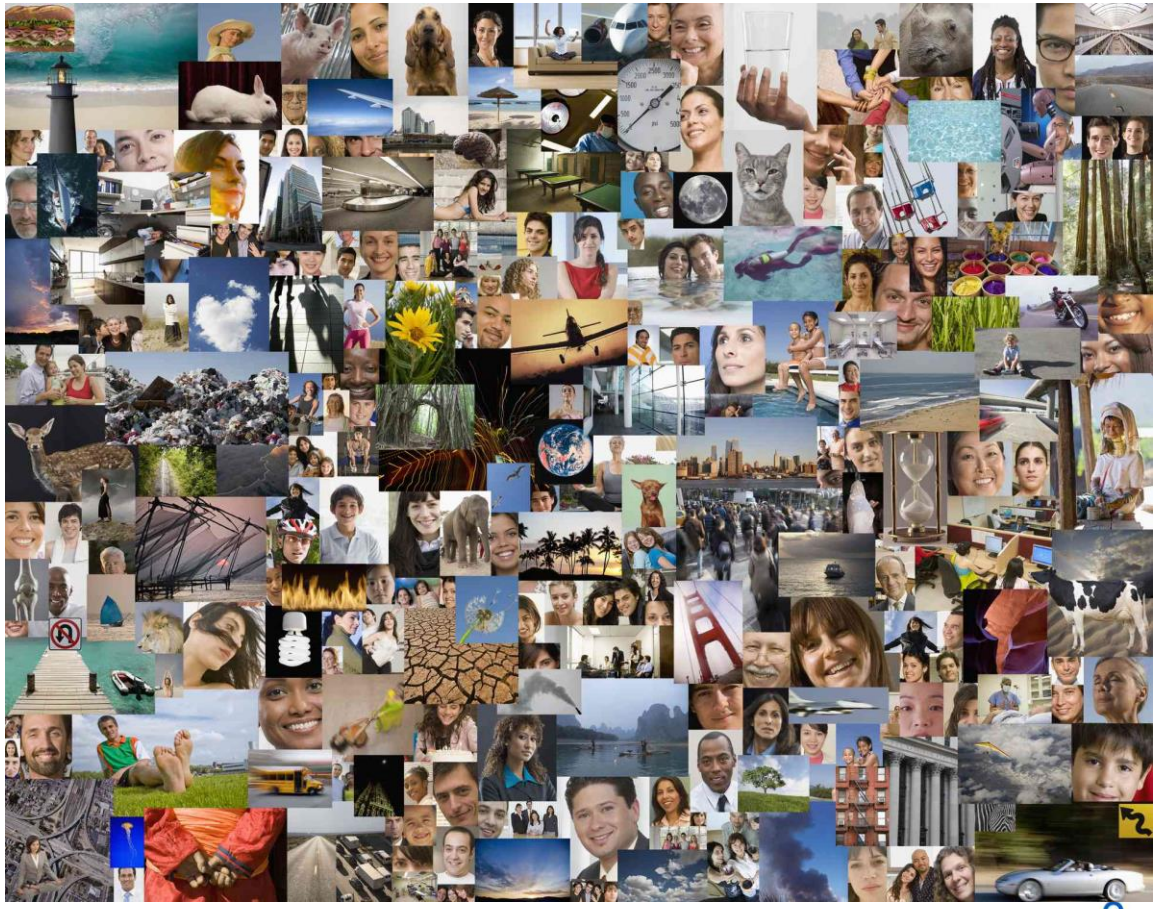
- The entire digital universe is expected to be 44 zettabytes in 2020.
- This will mean there are 40 times more bytes than there are stars in the observable universe.
- By 2025, it's estimated that 463 exabytes of data will be created each day globally – that's the equivalent of 212,765,957 DVDs per day



https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/

What about images ?

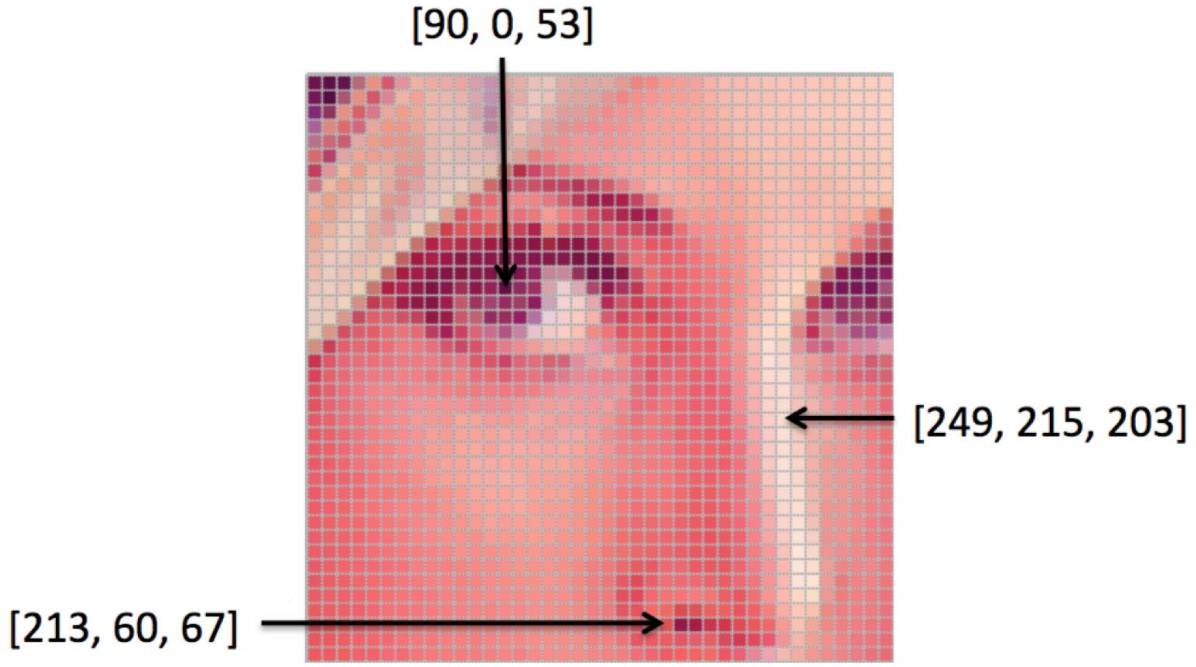
- Instagram users upload over 100 million photos and videos everyday. That is 69,444 million posts every minute
- 300 hours of video are uploaded to YouTube every minute!
- 7.15% of the content on Facebook is video. In 2020, there were 10.5 million videos on the site
- In 2020, the watch time for videos on Twitter rose by 72%



https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/
https://blog.microfocus.com/how-much-data-is-created-on-the-internet-each-day/

Images

- 2D Array of numbers (indicating intensity levels)
- Gray scale image has one channel and the value ranges from 0-255
- A color images has 3 channels Red, Green and Blue.
- The number of rows and columns in the 2D array indicates the resolution of the image.



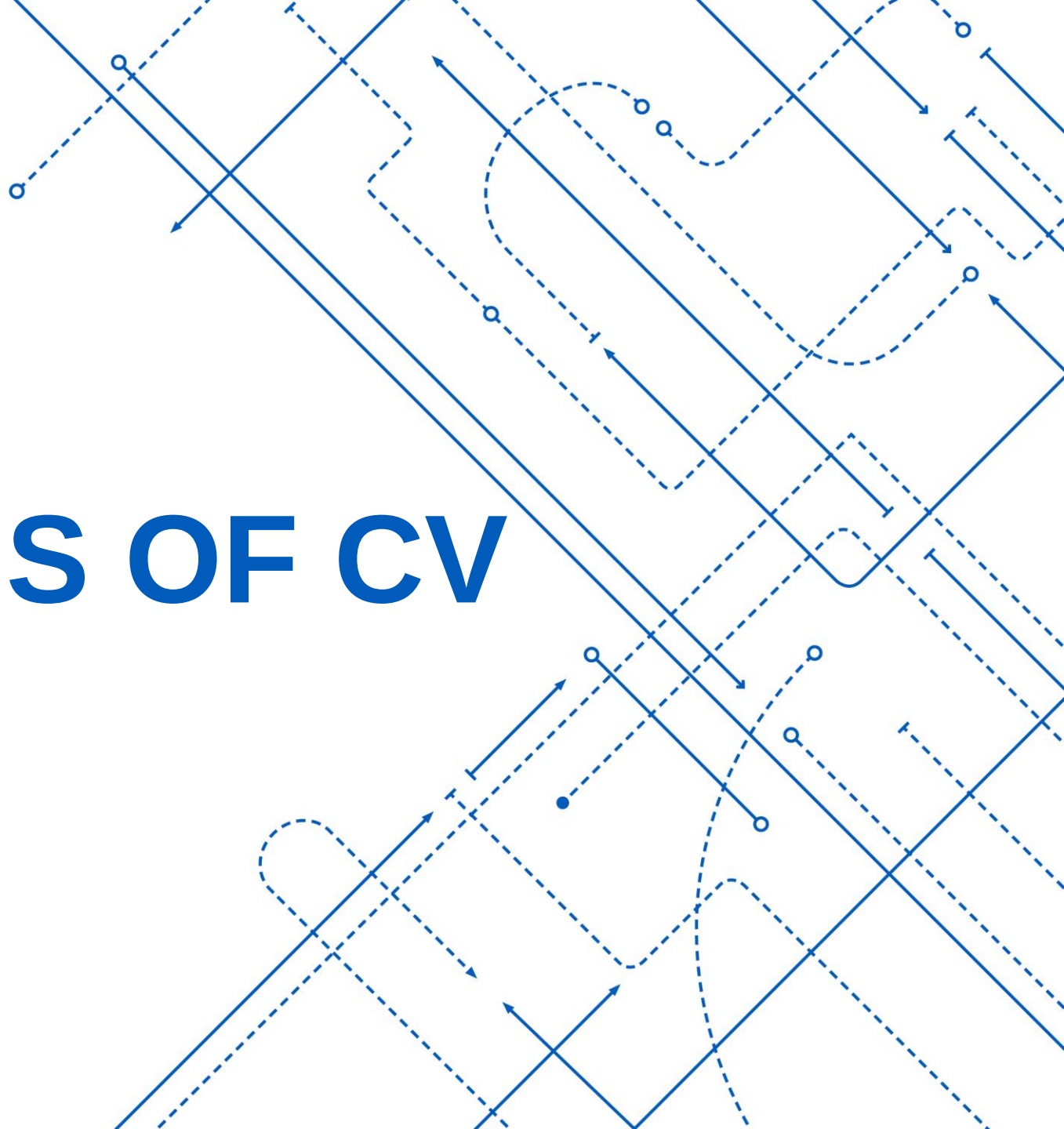
Computer Vision

- Computer Vision, often abbreviated as CV, is defined as a field of study that seeks to develop techniques to help computers “see” and understand the content of digital images such as photographs and videos.

- It is an interdisciplinary field having connections with fields like Mathematics, Pattern Recognition, Physics, Computer Science etc.

- Computer Vision is used in industries ranging from energy and utilities to manufacturing and automotive – and the market is continuing to grow. It is expected to reach USD 48.6 billion by 2022

APPLICATIONS OF CV

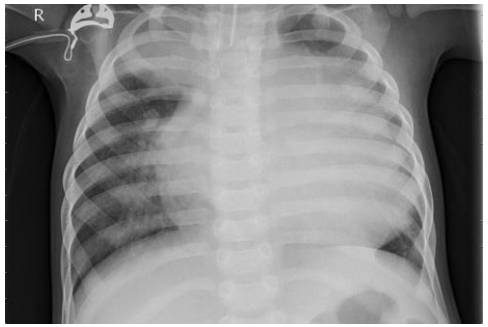


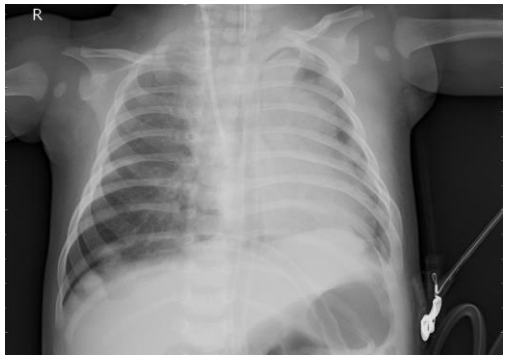
University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

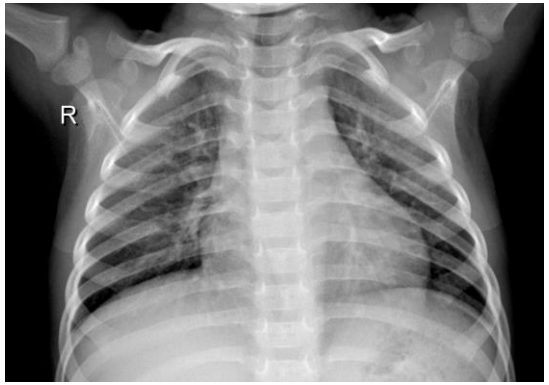
Recognition

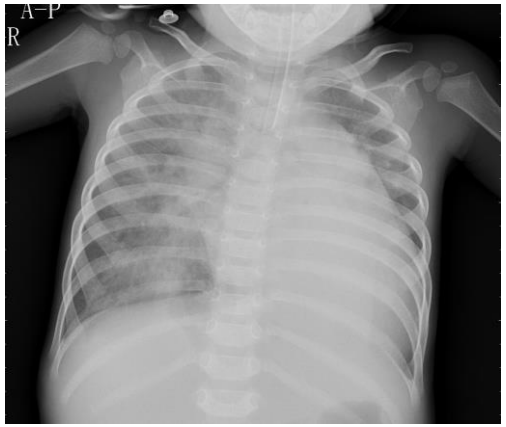
Recognition

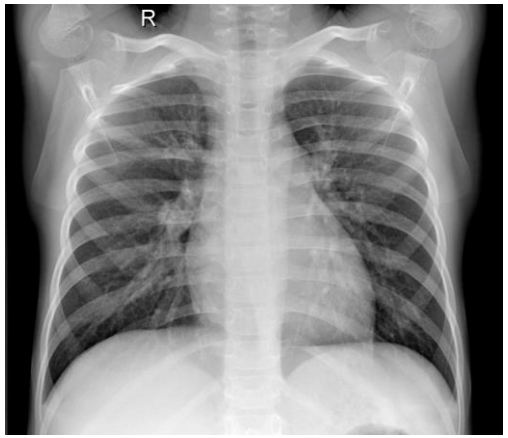
Recognition

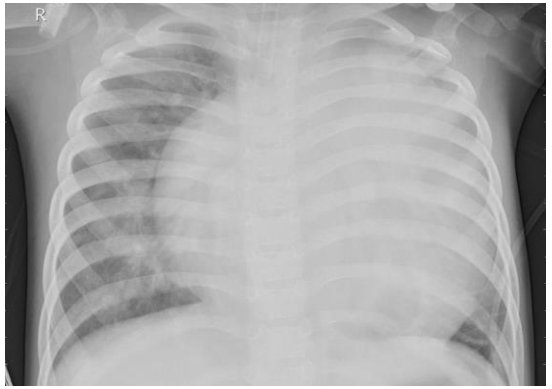












Multi-Class Classification

Cat

Dog

Dog

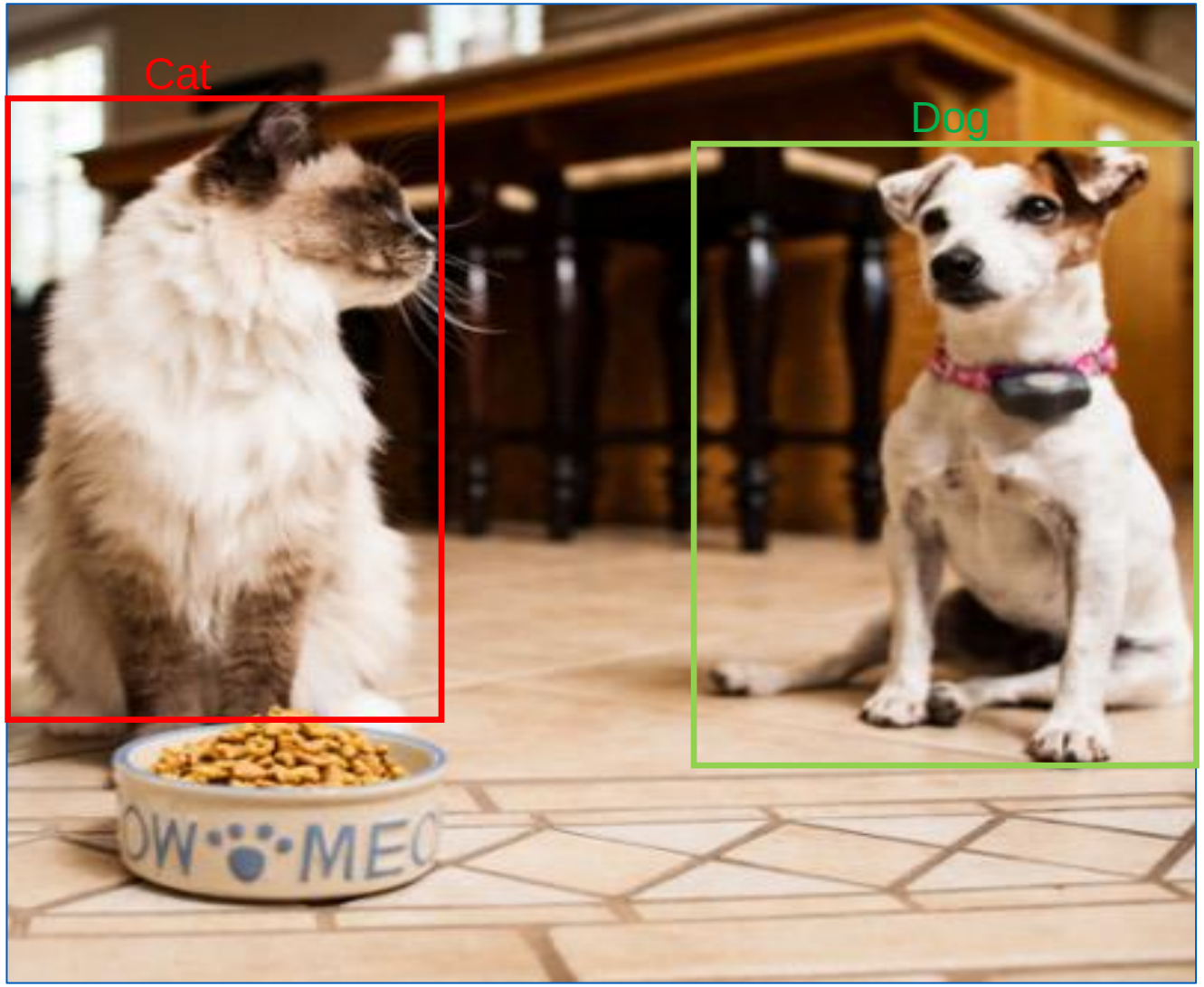
Bird

Cat

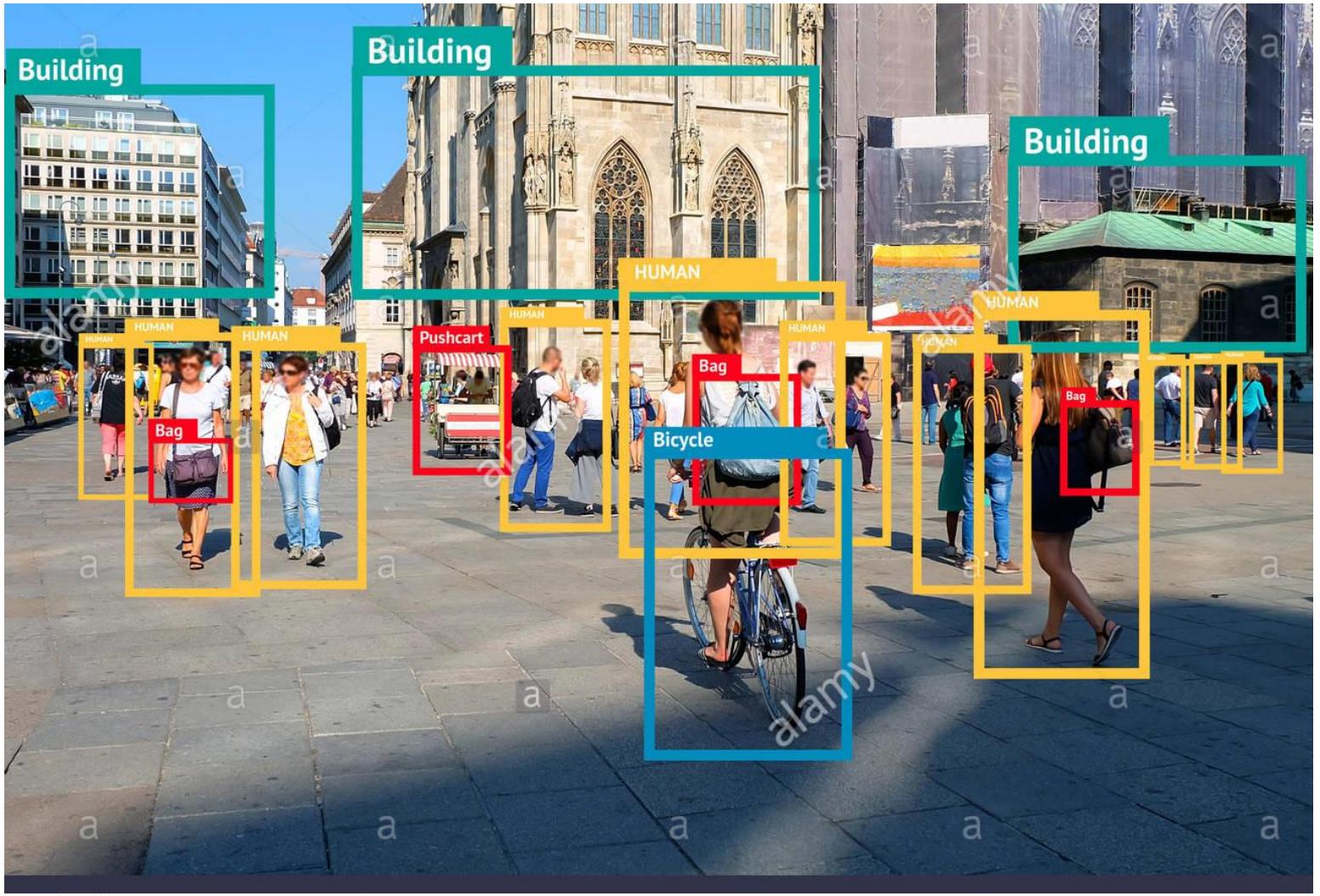
Bird

Detection

- Classification + Location
- Prediction would be Class of the object + 4 end points of the box enclosing the object
- More Challenging problem than recognition



When does it become Challenging



Multi-Camera Multi-trajectory Tracking

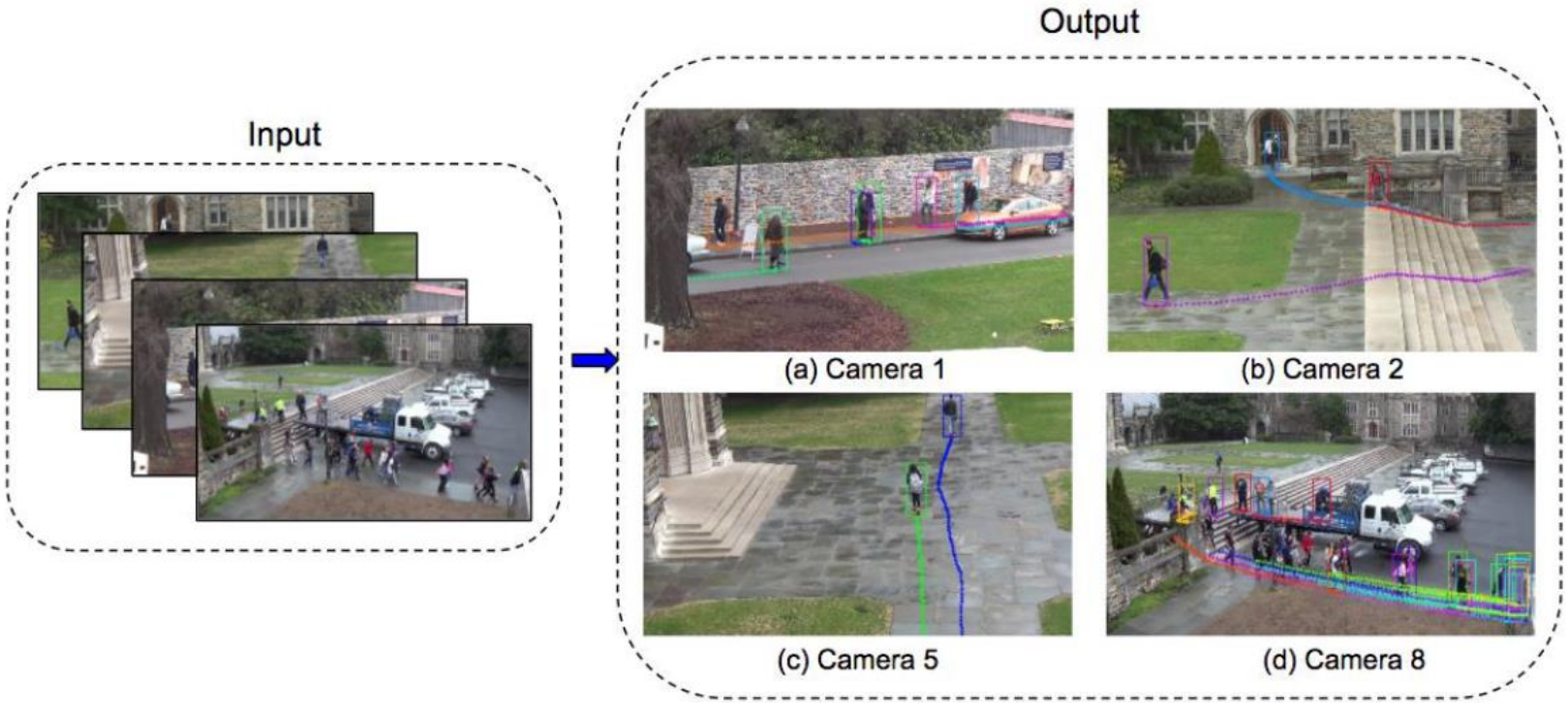
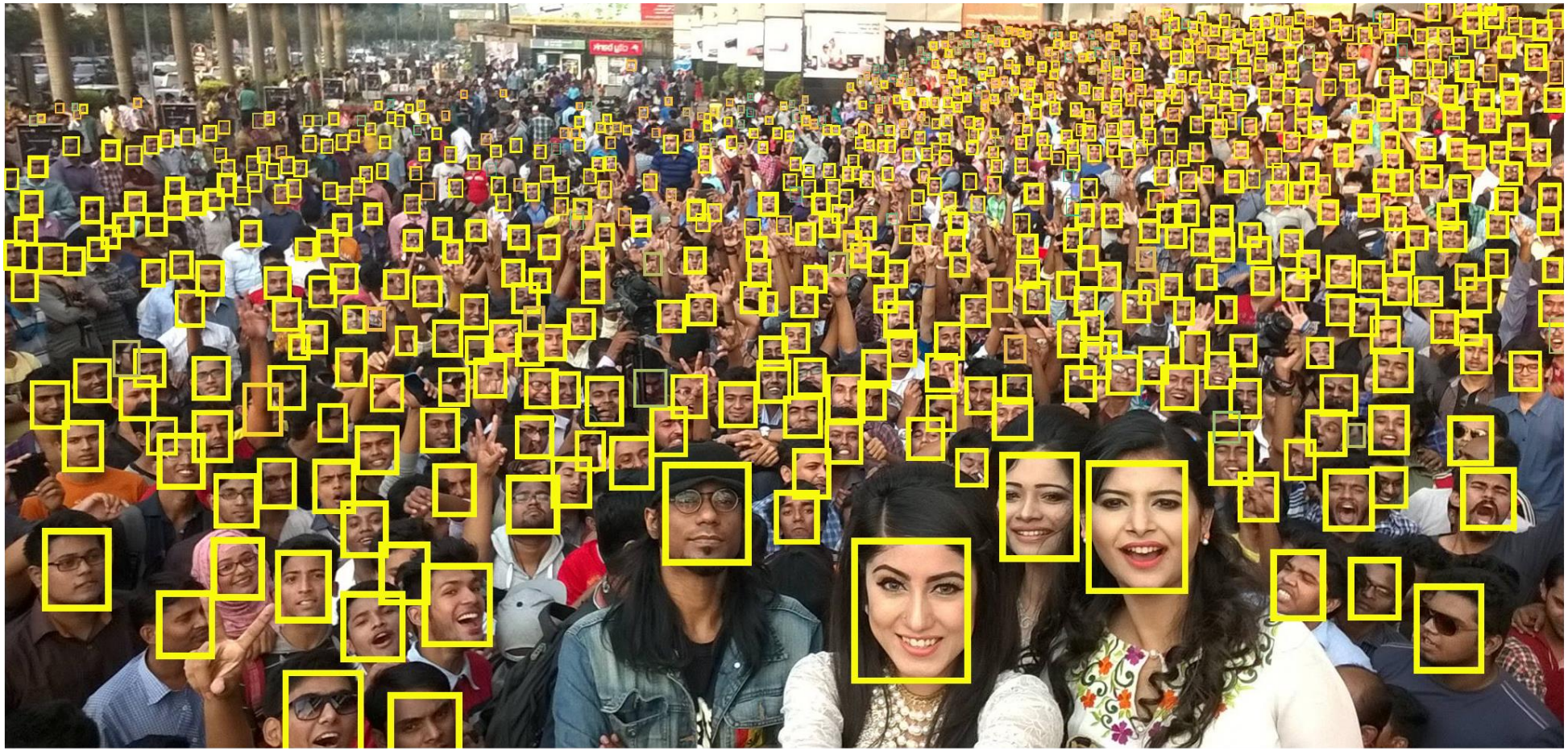


Fig 1: Camera sequences from DukeMTMC dataset with person trajectories highlighted

Multi-Camera Multi-trajectory Tracking

Face Recognition Demo

Face Detection Challenges



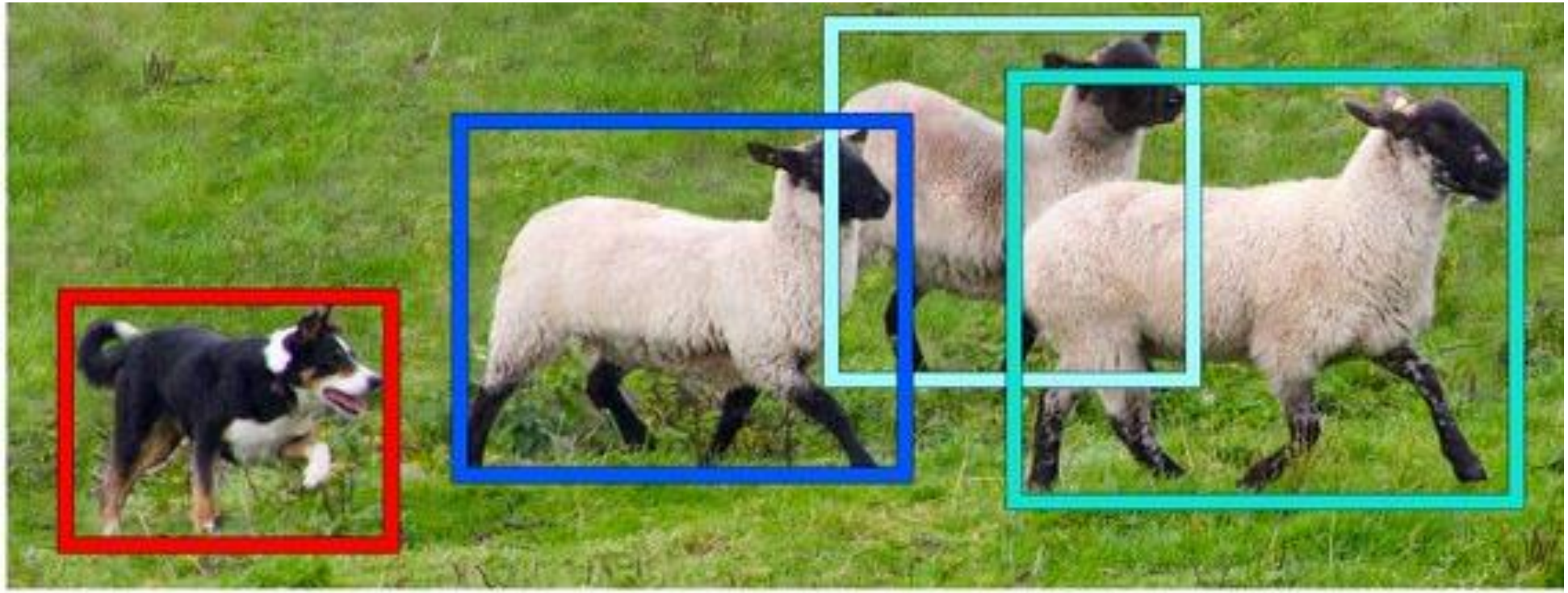
Facial Action Unit Recognition

More Faces

Other Biometrics

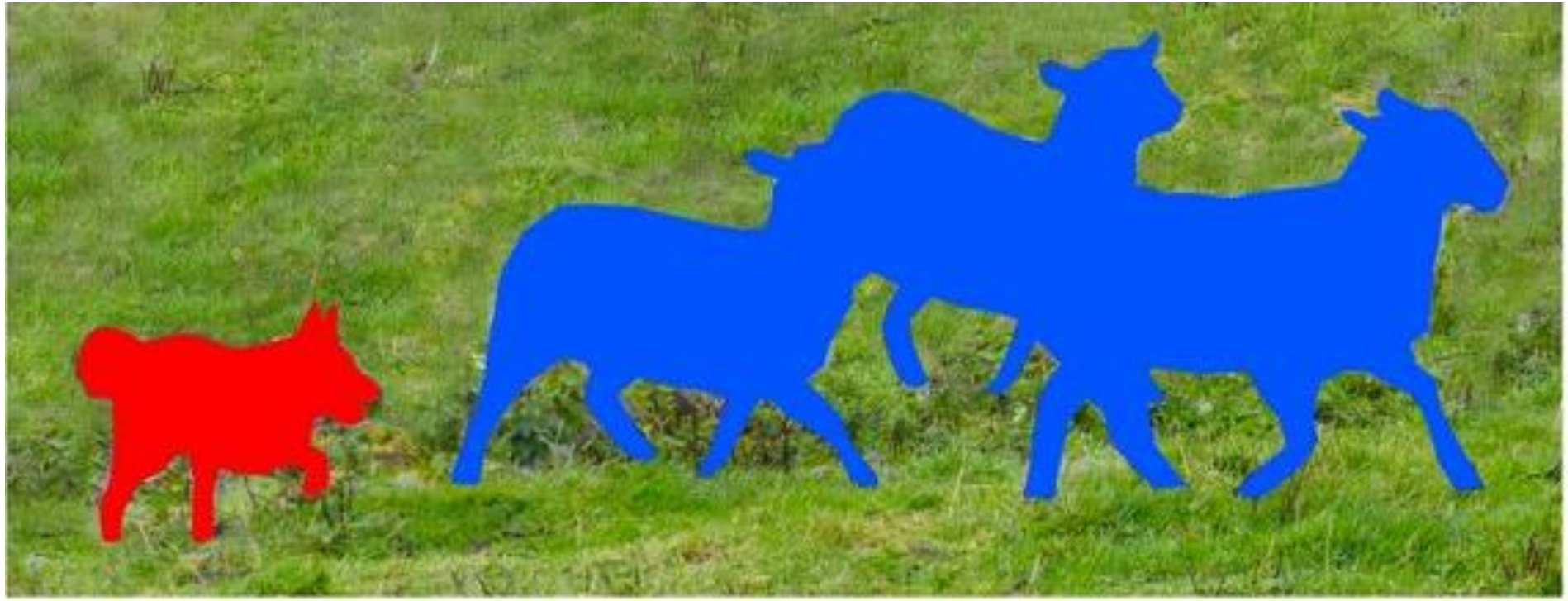
Segmentation

Segmentation



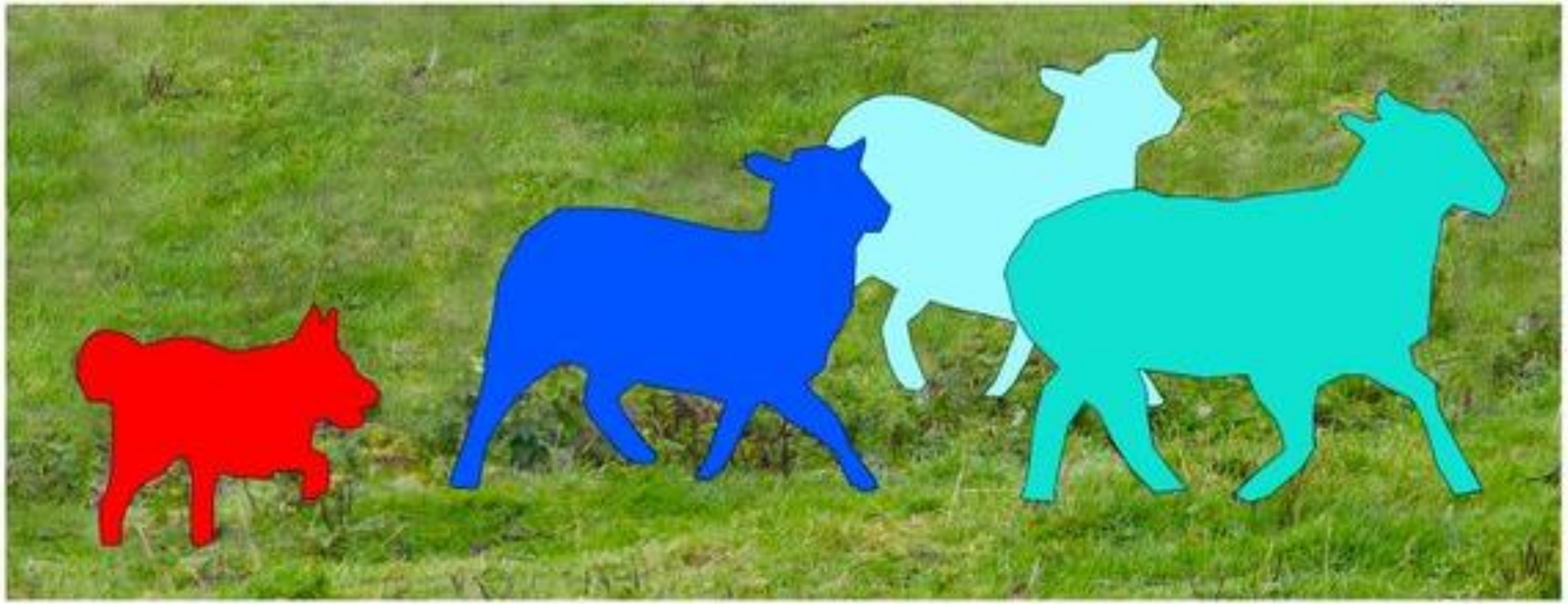
Object Detection

Segmentation



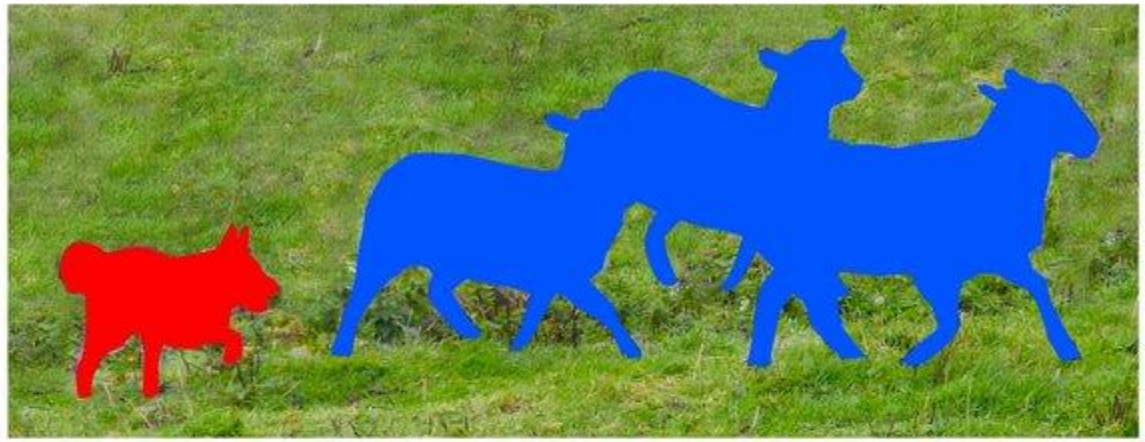
Semantic Segmentation

Segmentation

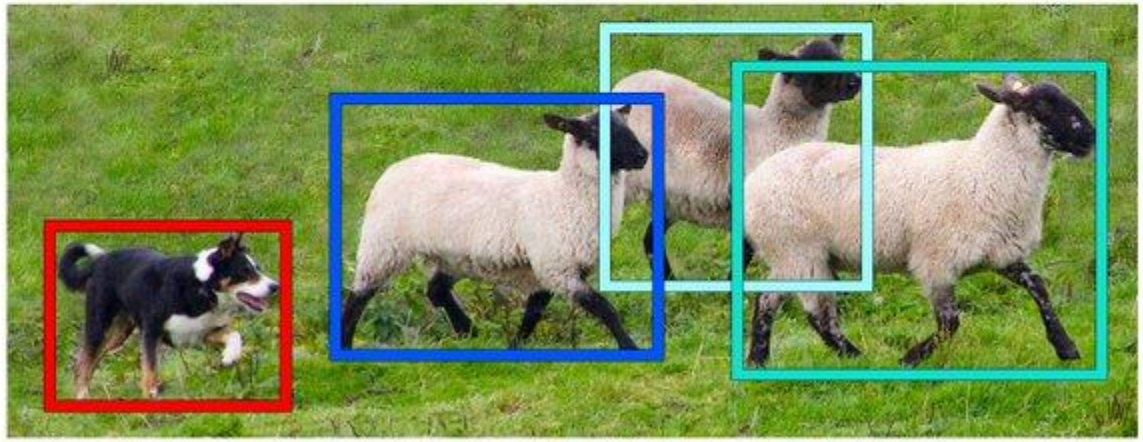


Instance Segmentation

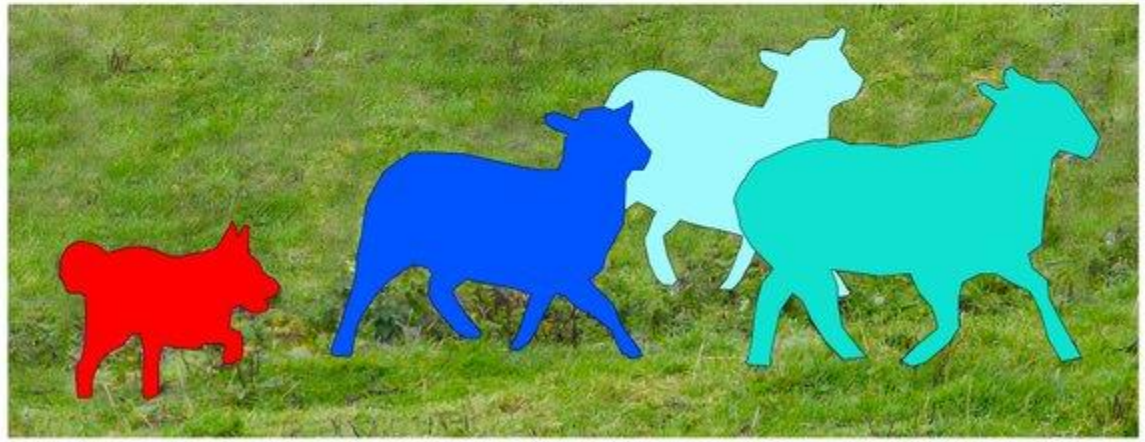
Image Recognition



Semantic Segmentation

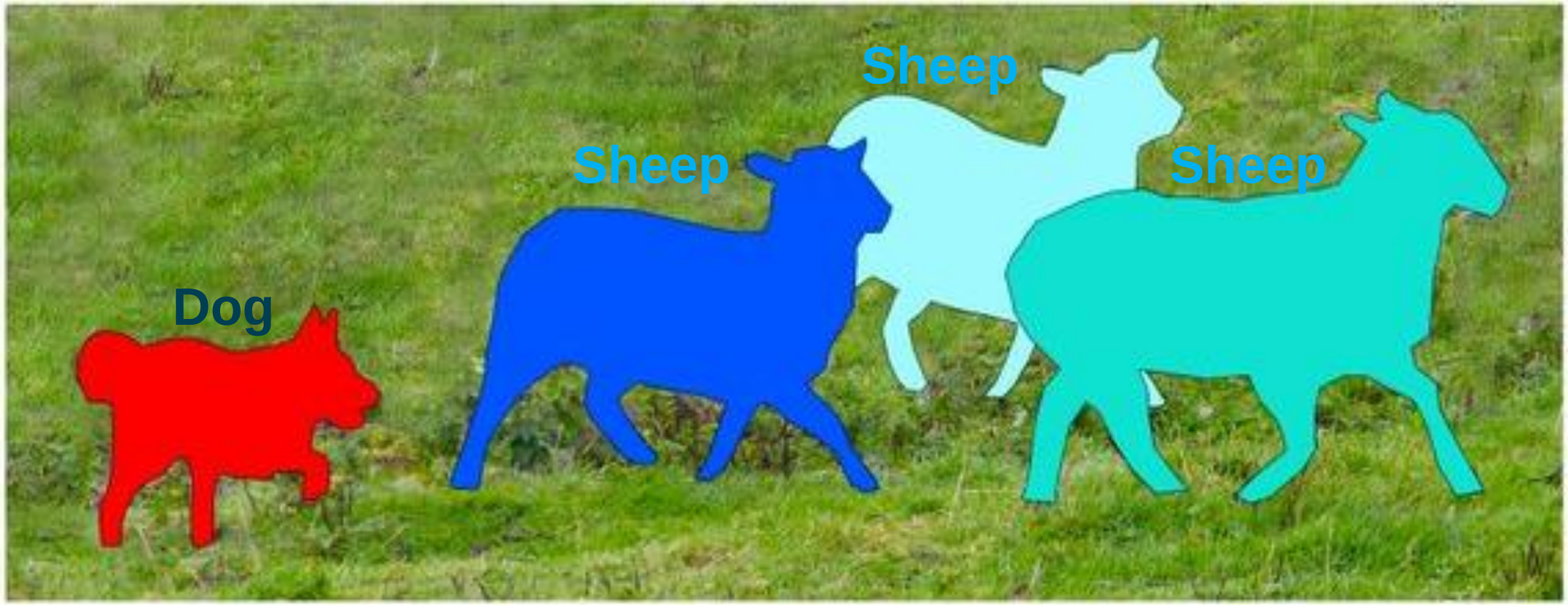


Object Detection



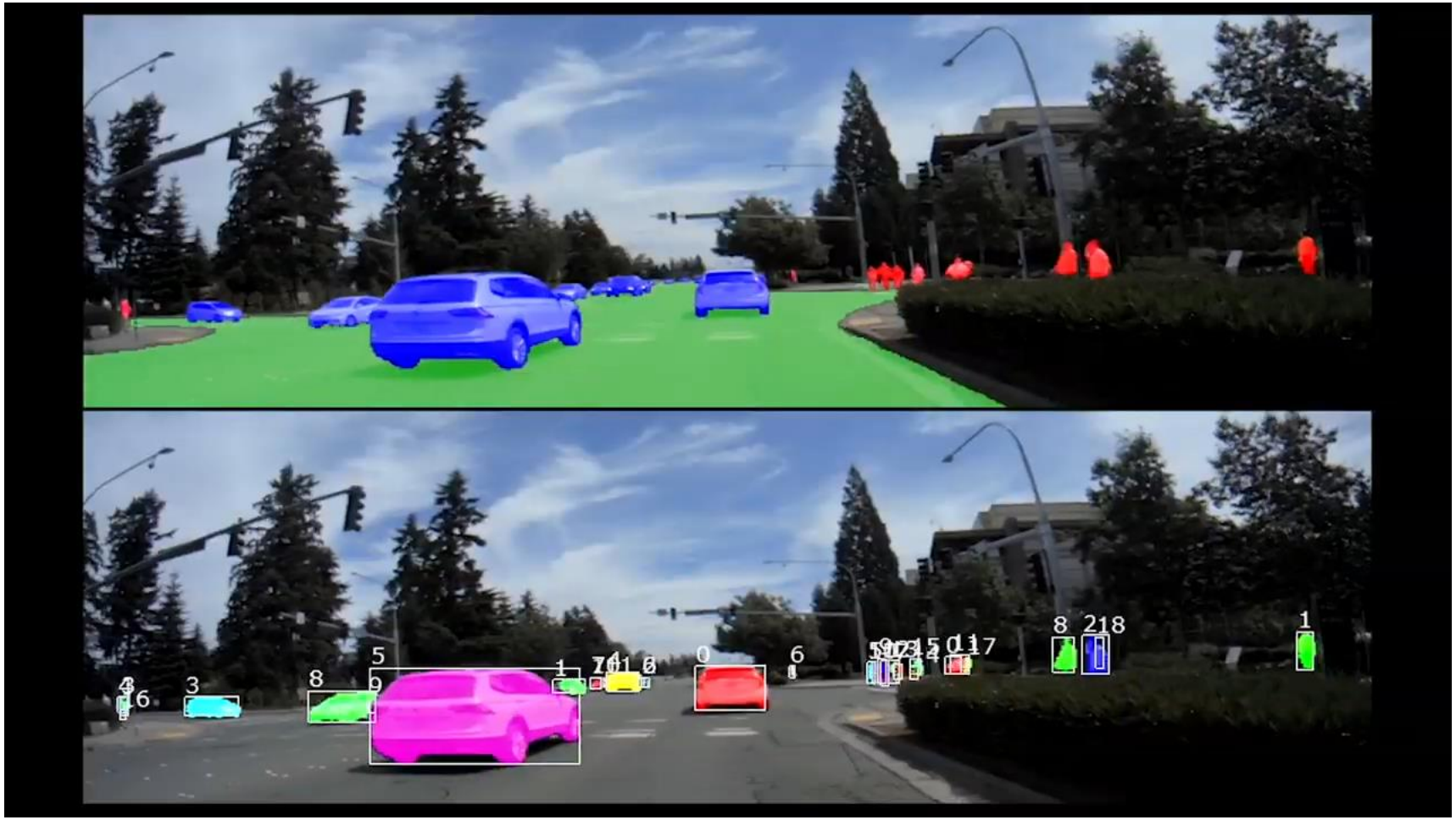
Instance Segmentation

Segmentation



Panoptic Segmentation

Segmentation in Autonomous Driving



Ref : https://www.youtube.com/watch?v=HS1wV9NMLr8

Scene Text Detection

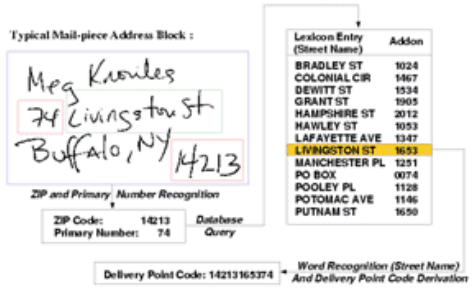
Handwritten Text

A Machine Learning Success Story

Pioneering work on Postal Automation at UB

- Handwriting recognition for postal automation
- Saving hundreds of millions of dollars in labor costs for the US Postal Service
- Over 95% of US letter mail sorted without manual intervention
- Technology licensed to Australia Post and UK's Royal Mail

Mail Transport Hardware
(Processing/Sorting)



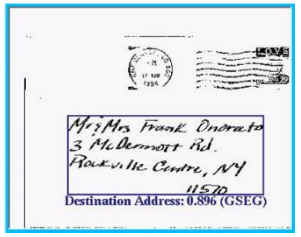
Automated Handwritten
Address Interpretation

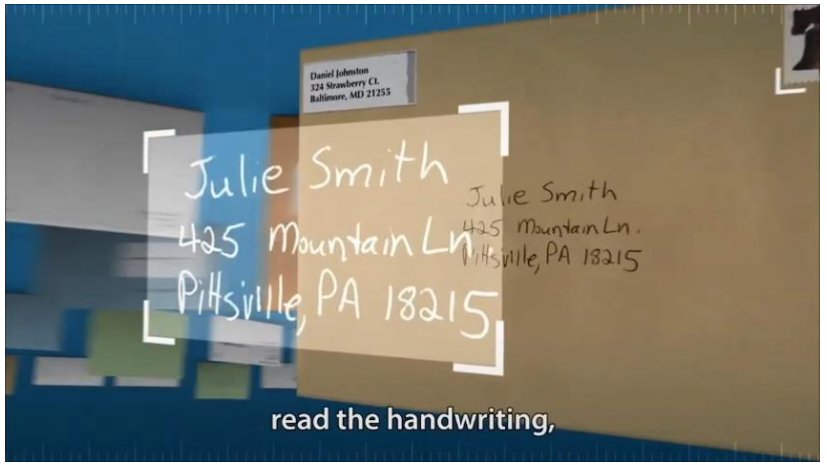
Remote Computer
Reader (RCR) Software

"*A lexicon driven approach to handwritten word recognition for real-time applications*",
IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 4, pp. 366-379, 1997

SIEMENS · AUSTRALIA POST · UNITED STATES POSTAL SERVICE · Royal Mail · LOCKHEED MARTIN

Handwritten Text





Handwriting recognition for postal automation

(click image for video)

Source: Systems at Work: a USPS TV Production
https://www.youtube.com/watch?v=WX16-52bHvg

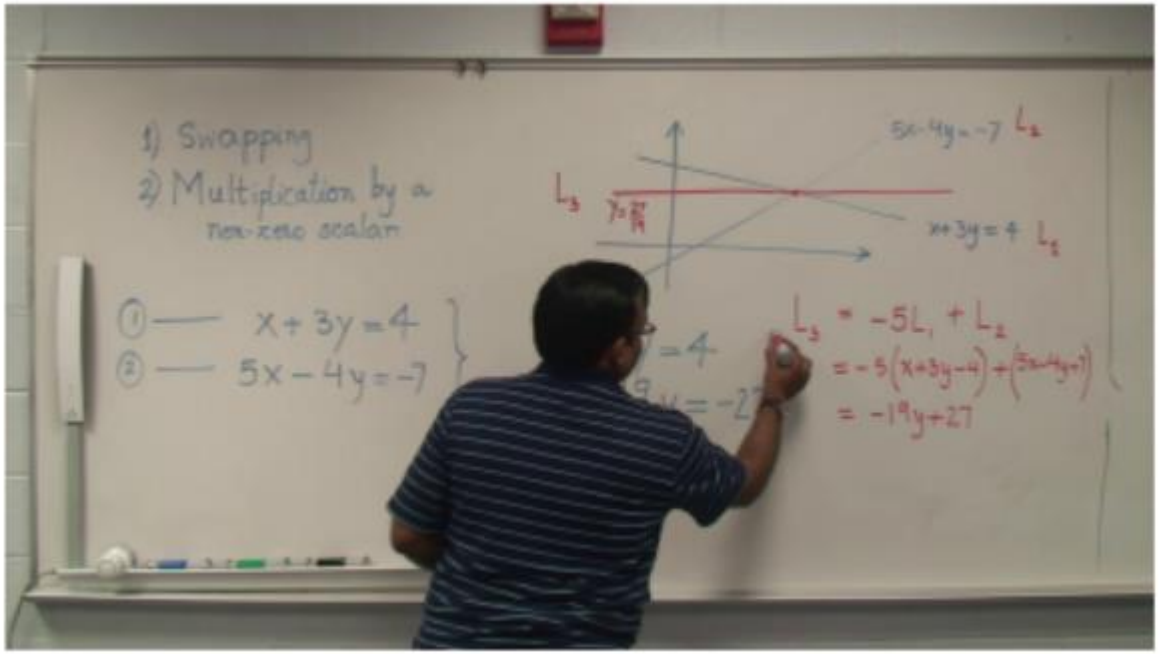
Jan. 24, 1997:

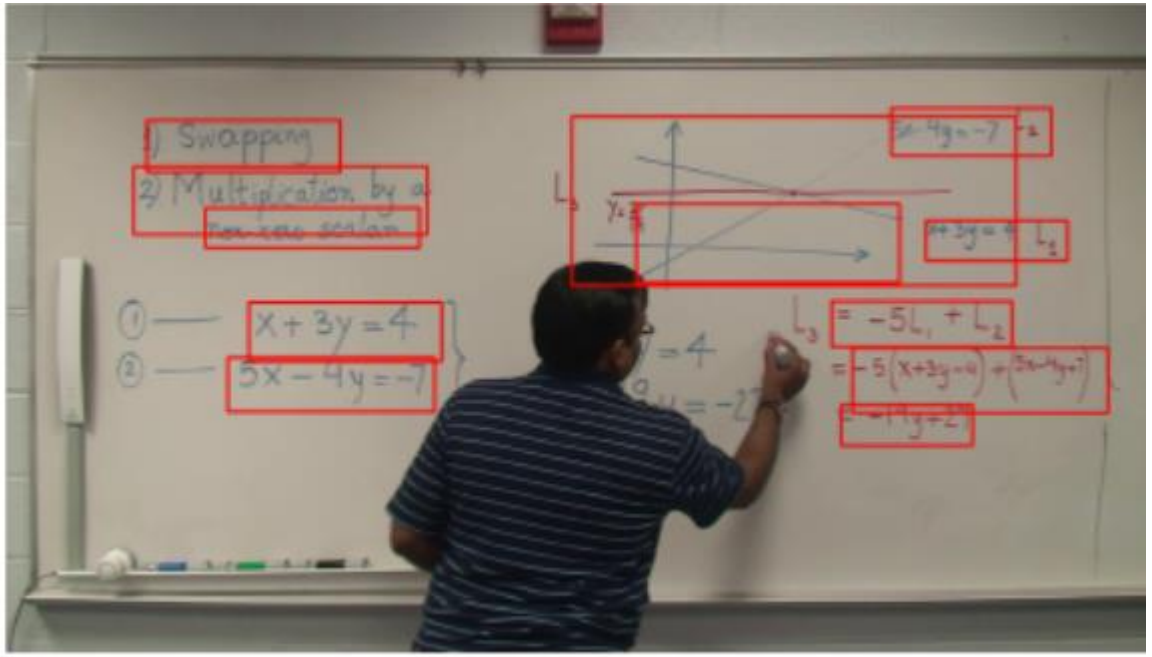
"This project represents a major step forward, not only for the Postal Service, but for technology in general," said Edward Kuebert, manager of image and telecommunications technology at USPS. "**It will do the seemingly impossible - help postal machines read handwritten mail.**"

Computing Community Consortium - March 25, 2009:

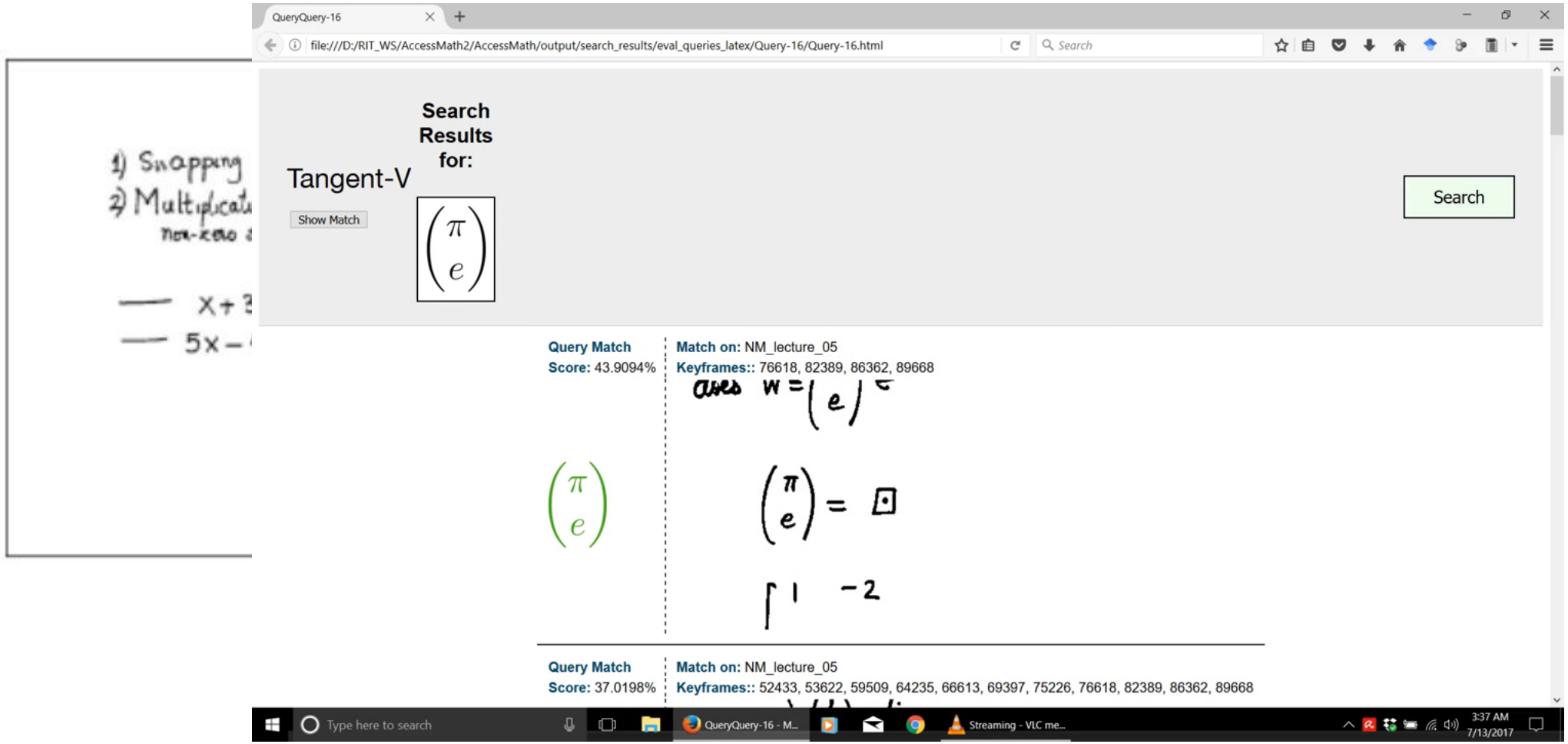
"Using a learning-based system developed at **SUNY Buffalo by Venu Govindaraju** and colleagues, 25 billion letters a year are processed automatically by the US postal service — bar-coded for precise delivery — **saving hundreds of millions of dollars…**"

Lecture Video Summarization and Search

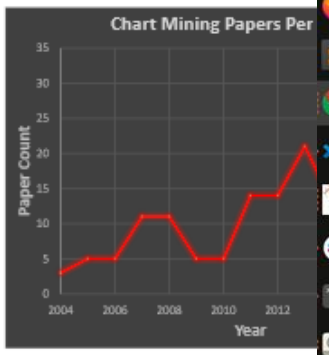




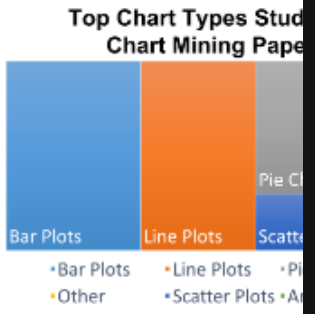
Lecture Video Summarization and Search



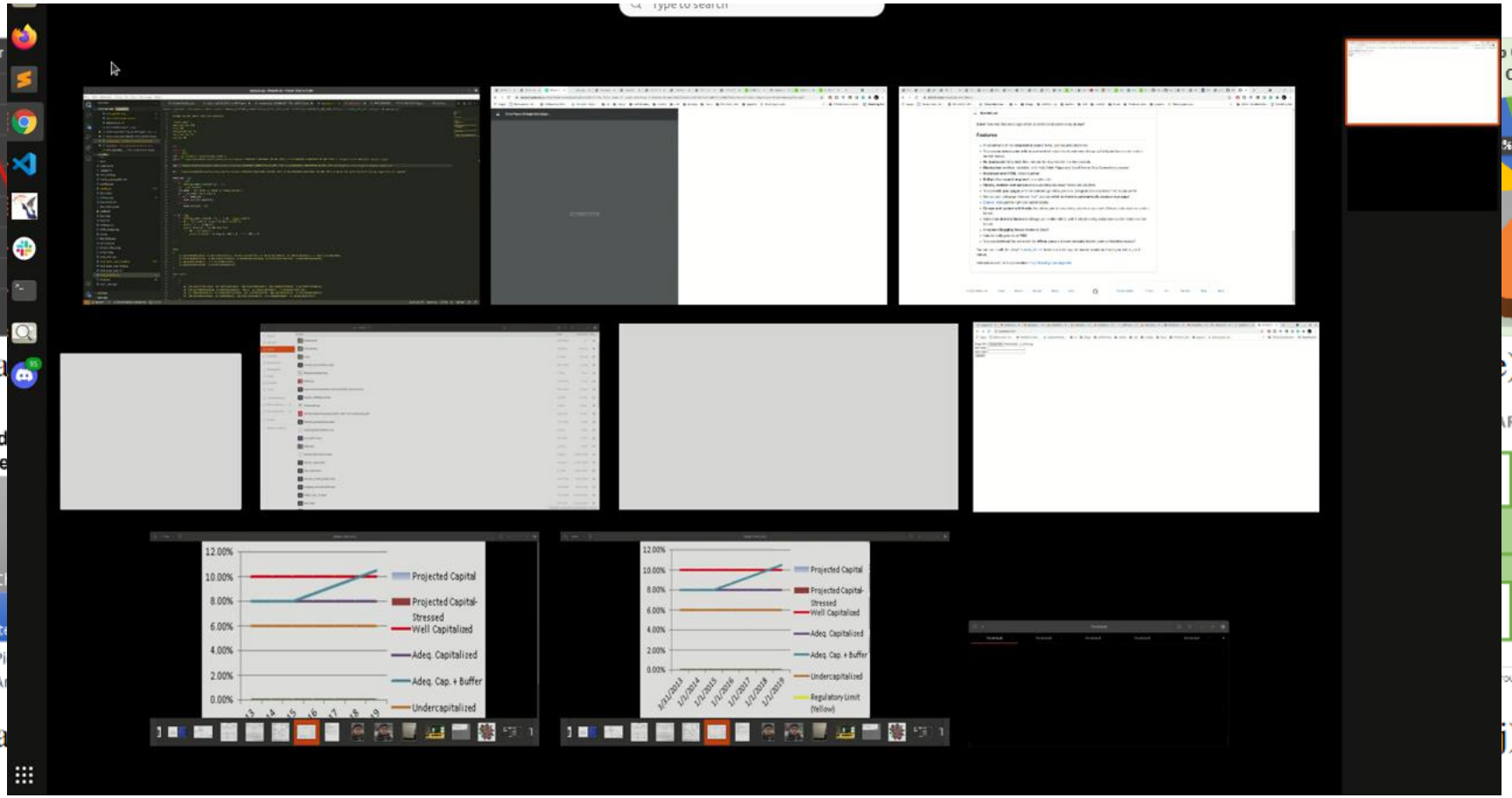
Infographics

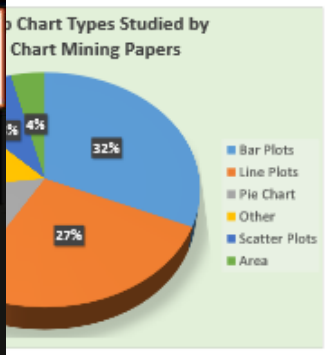


(a) Line Cha

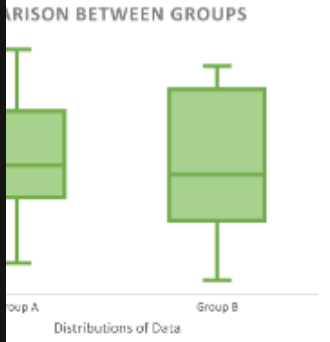


(f) Tree Ma



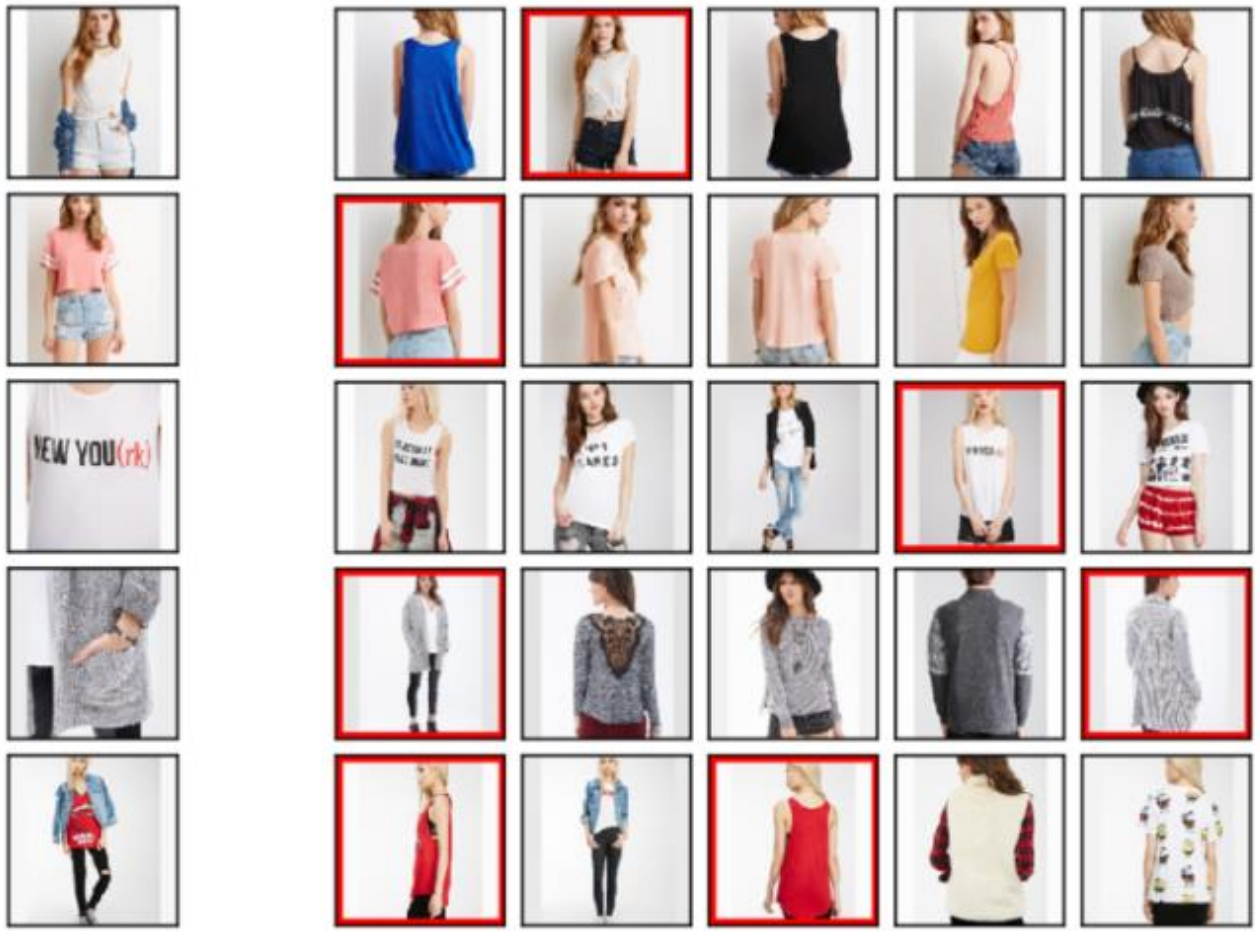


) Pie Chart

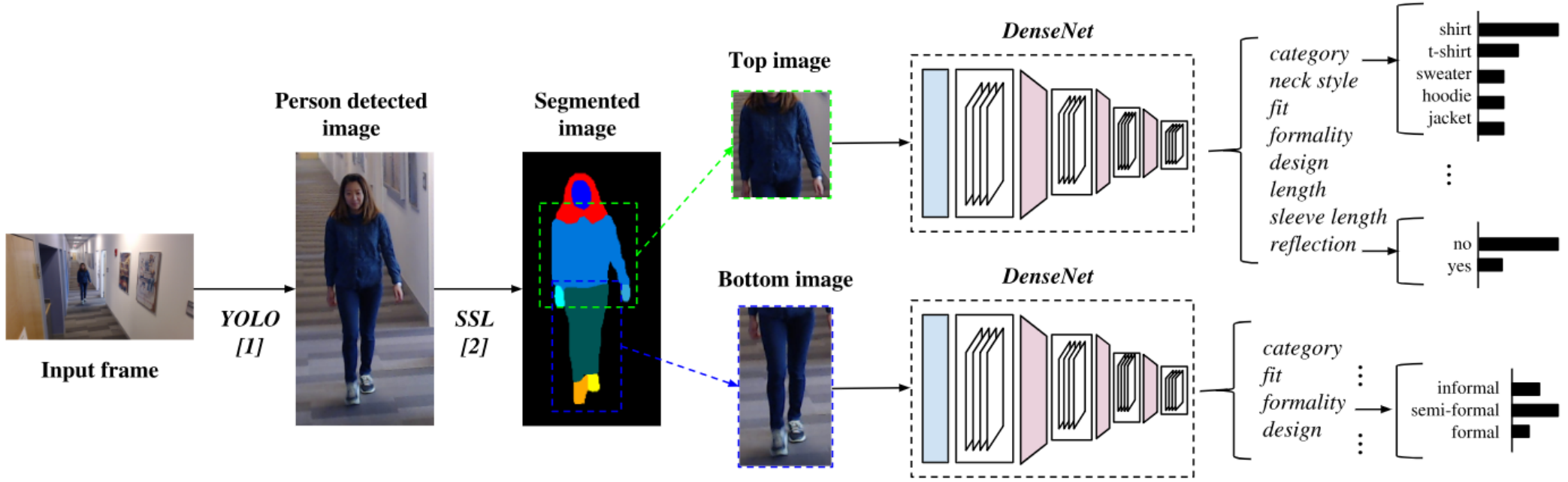


) Box Plot

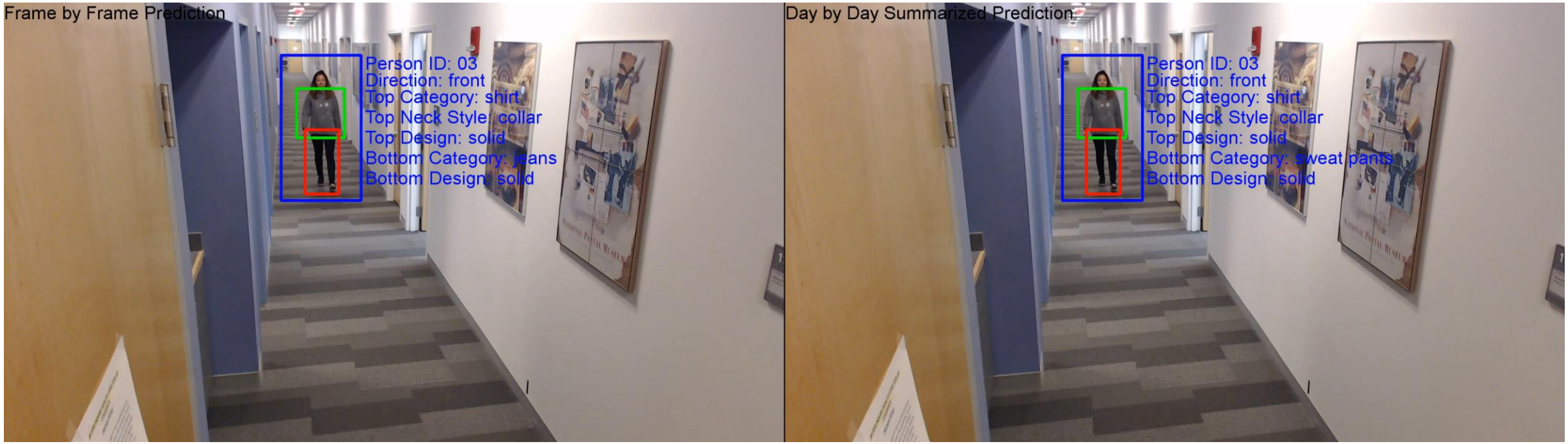
Image Retrieval



Clothing Attribute Detection and Wardrobe Model



Clothing Attribute Detection and Wardrobe Model



University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

COURSE CONTENT

What will you learn in this course?

- This will mainly be a project-oriented course that will introduce students to recent advancements in different computer vision tasks. The course will cover topics starting from the basics of deep learning to its application in solving a variety of computer vision problems.

- The primary objective of the course is to help students get an in-depth understanding of various real-world computer vision systems. This course will help students realize how modern deep learning techniques are used in current state-of-the-art computer vision tasks and also explore the challenges of designing, implementing, and benchmarking their custom solution for any of the computer vision tasks through a project.

What will you learn in this course?

- Deep Learning
  - Backpropagation
  - Optimizers, Initialization, Regularizations
  - CNN, RNN and LSTM's
- Object Recognition
  - Convolutional neural network for object recognition
  - AlexNet, GoogleNet, VGG,ResNet, DenseNet
  - Objective/loss functions for object recognition
  - Evaluation measures
- Object Detection
  - Convolutional neural network for object recognition
  - R-CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD
  - Objective/loss functions for object detection
  - Evaluation measures

What will you learn in this course ?

- Segmentation
  - Instance, Semantic and Panoptic Segmentation
  - U-Net, Dilated convolution and other CNN architecture for segmentation
  - Loss functions to handle class imbalance
- Other Computer vision tasks
  - Image Retrieval
  - Generative Models
  - Domain Adaptation
  - Knowledge Distillation
  - Self-Supervised Learning
  - Video Analytics
  - Face Detection and Recognition
  - Document Analysis

Programming language and Packages

- The course would primarily use python programming language.
- Python is beginner friendly, flexible and scalable object oriented programming language
- We will introduce Deep Learning frameworks such as TensorFlow, Keras and PyTorch.
- Students can use any of these packages for completing the assignment

TensorFlow

What will you NOT learn in this course ?

- The course would not focus on the following :-
  - Probabilistic Graphical Models.
  - Traditional Computer Vision techniques
  - Statistical Machine Learning
  - Natural Language Processing
  - Reinforcement Learning

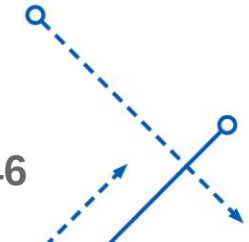
Grading Scheme

- Programming Assignment – 35 % of the grade
  - Assignment 1 – 9 %
  - Assignment 2 – 13%
  - Assignment 3 – 13%
- Quizzes – 10 % of the grade
  - Mostly through UBlearns. Around 5-10 quizzes
- Midterm – 15 % of the grade
  - Contains Multiple Choice, Numerical and Descriptive questions
- Project – 40 % of the grade
  - Project decided towards middle of the semester, two presentation.

I reserve the right to change any part of the syllabus or grading scheme

Course Details

- The Grades will be based on a curve with B+ being assigned to the median score.
- The link to the course website is https://cubs-ub.github.io/Computational_Vision/
- The link to piazza is piazza.com/buffalo/fall2021/cse673 . Please sign up
- Office hours and locations will be posted on piazza and course website.



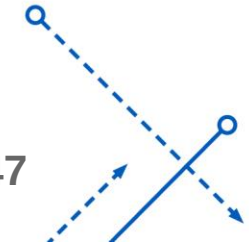
Miscellaneous

- Prerequisites for the course are:
  - Python programming
  - Linear Algebra
  - Probability and Statistics
  - Machine Learning / Computer Vision

- Midterm exam on “10/14/2020”

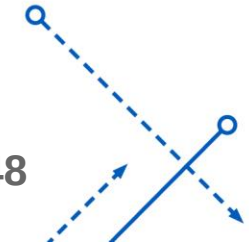
- The class will have a “no laptops, no cell phones” policy

- The Lecture will not be recorded



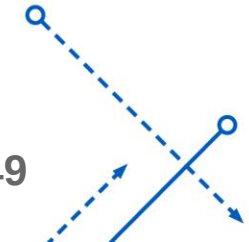
Academic Integrity

- **Plagiarism** in any form will earn you an F in the course.
- Please review the CSE academic integrity policy https://engineering.buffalo.edu/computer-science-engineering/information-for-students/academics/academic-integrity.html
- You can ask help from instructors.
- Copying code from any online resources like GitHub or sharing code among your classmates is academic integrity violation which results in you getting ‘F’ (grade) in the course
- Do not discuss implementation or help debug code for any of your classmates



Frequently Asked Questions

- Do I get a make up exam ?
  - No
- Can I get an incomplete in the course ?
  - No
- Are the assignments individual or group ?
  - The assignments will be group of 1-3 members depending upon the final enrollment.
- My partner copied from an online source. Will I be held responsible ?
  - Yes, you are responsible for the work that is submitted by the team

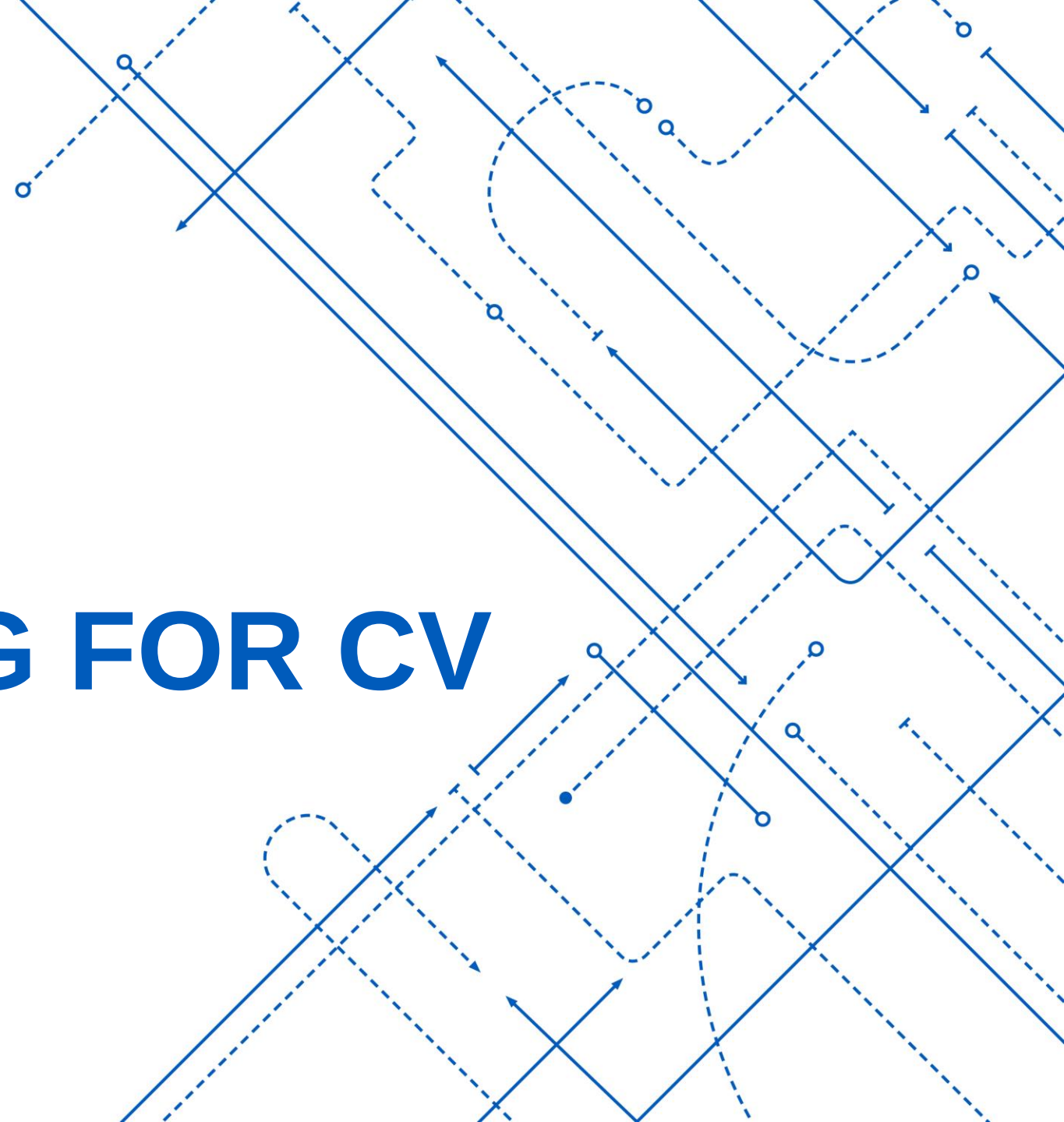


Frequently Asked Questions

- How much time will I have to dedicate for this course?
  - Depends on individual's background and ability, but about 15-20 hours per week on average

- I took ML | CV previously, I did not do well. Can I still take this course ?
  - Even though you can take the course, most of the contents assumes a strong understanding of ML and CV.

- What kind of help will I get during the programming assignment ?
  - The instructors will help you answer your conceptual questions
  - You will not get help in debugging your code.



DEEP LEARNING FOR CV

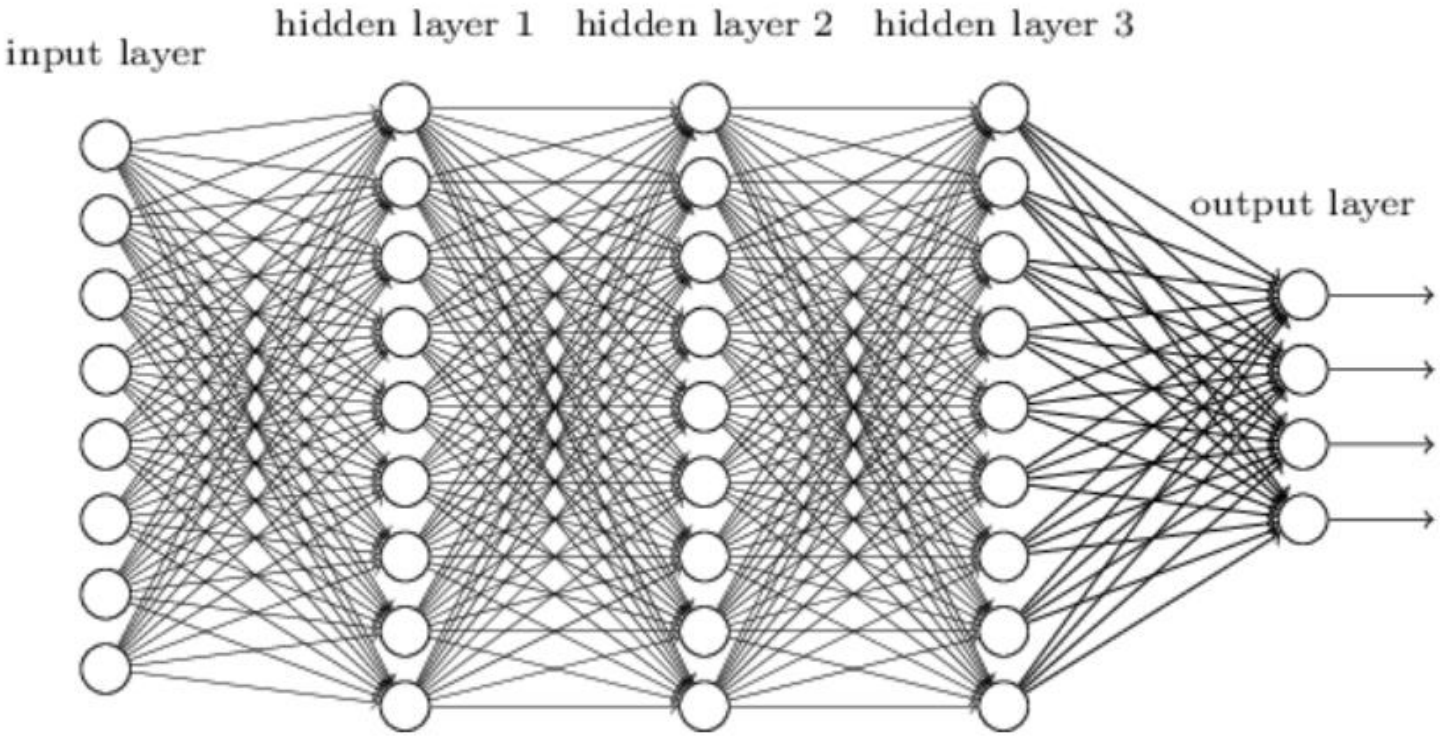
University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

ImageNet Dataset

- Introduced first in 2010, the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) evaluates algorithms for object detection and image classification at large scale.
- The ImageNet dataset contains 14,197,122 annotated images
- 21K classes or groups

Is neural networks a new concept?

- Idea of Artificial Neural Network is pretty old
- Backpropagation Algorithm was a 1986 paper by Hinton
- So why the sudden popularity ?
- People figured out how to use GPU's for training Neural Networks.




ANY
QUESTIONS
?

References

- https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/
- https://blog.microfocus.com/how-much-data-is-created-on-the-internet-each-day/
- https://techjury.net/blog/how-much-data-is-created-every-day/
- https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html
- https://www.ibm.com/topics/computer-vision
- https://www.youtube.com/watch?v=BxgFOQzWFI0
- http://kdavila.com/
- http://neuralnetworksanddeeplearning.com/chap5.html
- https://image-net.org/
- http://cs231n.stanford.edu/
- https://github.com/clovaai/CRAFT-pytorch/blob/master/figures/craft_example.gif
- https://www.youtube.com/watch?v=HS1wV9NMLr8
- https://ai-pool.com/d/could-you-explain-me-how-instance-segmentation-works