EE/CprE/SE 491 - sddec23-10

Developing a Deep Learning Model to Automatically Detect Microscale Objects in Images and Videos

Week 2 Report

02/06/2023 – 02/12/2023
Client : Professor. Santosh Pandey
Group number: 10

Team Members:

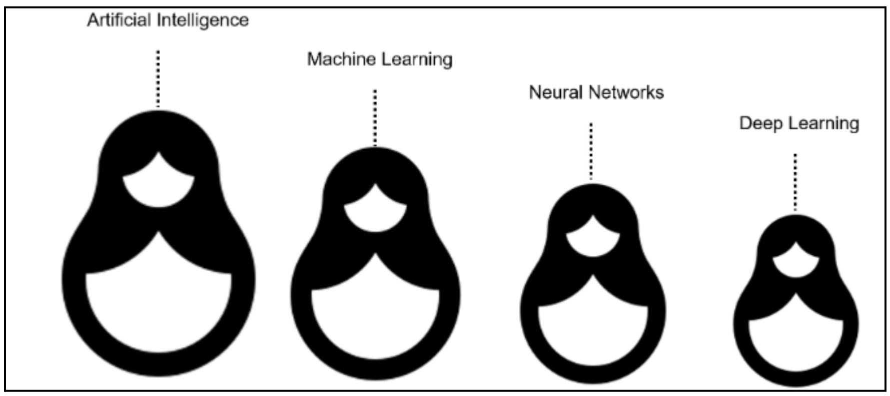
Katherine Moretina
Ethan Baranowski
Chris Cannon
Matthew Kim

Past week Accomplishments

Background on Artificial Intelligence, Machine Learning, and Neural Networks- Everyone

Artificial Intelligence, Machine Learning, and Neural Networks- https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks

- How do they relate to each other- each is a component of the prior term
  - If these terms were Russian nesting dolls, it would look something like:



- Artificial intelligence
  - The broadest term used to describe machines that mimic human intelligence
  - Predicts, automates, and optimizes human decision-making tasks
- Machine Learning- https://www.coursera.org/articles/types-of-machine-learning
  - Focused on using previous data to predict outcomes
  - Split into three types
    - Supervised learning- the machine is "supervised," meaning the user is feeding algorithm data to help the machine learn (data is labeled)
    - Unsupervised learning- unlabeled sets of data. The machine is looking for less obvious patterns
    - Reinforcement learning- learns with environment interactions and getting a positive or negative reward
- Neural Networks
  - Mimic the brain through a set of algorithms
  - Comprised of four main components- inputs, weights, bias/threshold, and an output
  - Each node, or "neuron" takes an input, assigns it a weight, and if the input and weight together meet the bias, or threshold, it will send an output signal to other nodes it is connected to.
- Deep learning
  - Technically, any neural network with more than three layers of "neurons" including the input and output layers.

- In practice, these tend to be black box solutions - it can be virtually impossible to tell what criteria the algorithm is using in the end.

YOLO, MobileNet, and SqueezeDet- Katie

- YOLO- https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/#:~:text=What%20is%20YOLO%3F,probabilities%20of%20the%20detected%20images.
  - Overview
    - Stands for "You Only Look Once." Real-time object detection algorithm used in commercial products by large tech companies.
    - Multiple variants of YOLO have been released, each performing faster than the last
    - Regression problem: provides class probabilities
  - Deep Neural Network based object detection approach
  - Fast algorithm because it predicts in real time, entire image is done in a single algorithm run
  - High accuracy with minimal background error
  - Not specialized in small objects
  - Study from https://ieeexplore.ieee.org/document/9554234 states 99.3% accuracy when Intersection over Union threshold is 50%. The training set was 600 images and tested with 400 images (YOLOv5)
    - The image resolution was 1280x659 pixels and in .png format
  - Time varies depending on the version, but most testing has been done with v5
    - The average detection time from the IEEE article averaged 0.138 sec
- MobileNet- https://www.analyticsvidhya.com/blog/2022/09/object-detection-using-yolo-and-mobilenet-ssd/#:~:text=Mobilenet%20SSD%20is%20an%20object,detection%20optimized%20for%20mobile%20devices.
  - A lot of resources I'm finding say MobileNet is used in conjunction with SSD
  - Overview
    - Models called for mobile and embedded applications
    - Uses standard convolution into depth and point convolution
    - Splits normal convolution that is one layer into two
    - Depth convolution applies a filter to two inputs. The point convolution generates convolution to combine outputs of depth convolution
    - SSD provides localization while MobileNet provides classification
  - A subset of convolutional neural networks
  - The article found compares MoblieNet SSD to YOLO. Conclusions were;
    - YOLO had better accuracy. Mobile net provided faster detection
    - YOLO is better for small objects
  - https://18it107.medium.com/object-detection-using-mobilenet-ssd-8ddf64d5de9a
    - Provided good information on SSD and MobileNet combined

- **SDD300:** 59 FPS with mAP 74.3%
- **SSD512:** 22FPS with mAP 76.9%
- **Faster R-CNN:** 7 FPS with mAP 73.2%
- **YOLO:** 45 FPS with mAP 63.4%

Method	mAP	FPS	batch size	# Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	1	~ 6000	$\sim 1000 \times 600$
Fast YOLO	52.7	155	1	98	448×448
YOLO (VGG16)	66.4	21	1	98	448×448
SSD300	74.3	46	1	8732	300×300
SSD512	76.8	19	1	24564	512×512
SSD300	74.3	59	8	8732	300×300
SSD512	76.8	22	8	24564	512×512

Performance

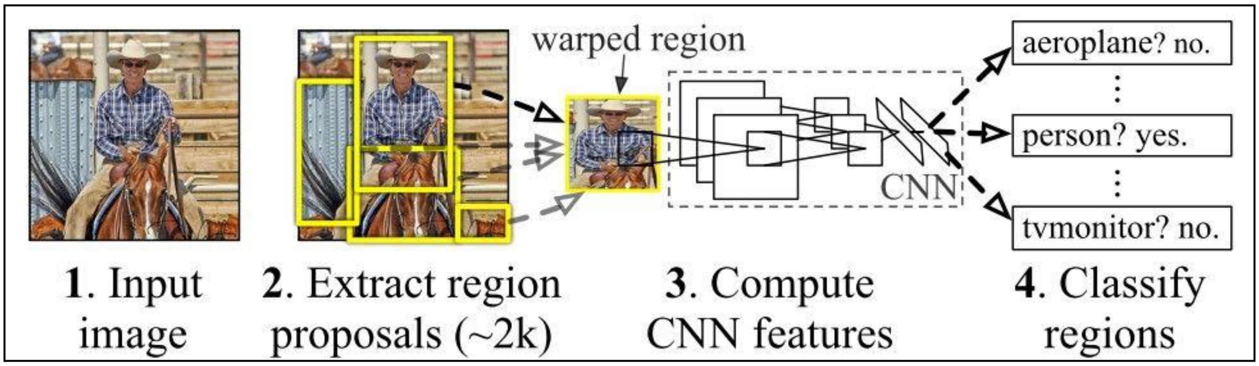
- This article found higher accuracy with combined MobileNet and SSD than with YOLO

- SqueezeDet- https://ieeexplore.ieee.org/abstract/document/8014794
  - Overview
    - Deep neural network, single shot detector algorithm
    - Developed for autonomous driving
    - Convolutional layers are used to extract feature maps but also as the output layer to compute bounding boxes and class probabilities
    - Single forward pass neural network allows SqueezeDet to work extremely fast
    - Detection pipeline inspired by YOLO
  - 57.2 frames per second with the original model and 32.1 with SqueezeDet+
  - 7281 images (of 1242x375) split in half- half training and half validation
  - Only was able to find one article and one github page on this method so a lot of information is missing

R-CNN, FAST RCNN, Faster RCNN - Matthew

- R-CNN (Region-Convolutional Neural Network)
  - Used CNN to make higher object detection performance.
  - Algorithm
    - Use selective search algorithm to generate Region proposal from input image
    - Warp Region proposal in fixed size to use as CNN input
    - From the feature map generated by CNN, get the vector.
    - Use that vector in different learning classes.
  - Cons
    - CNN takes a huge amount of time for learning.
    - It requires a large amount of space.

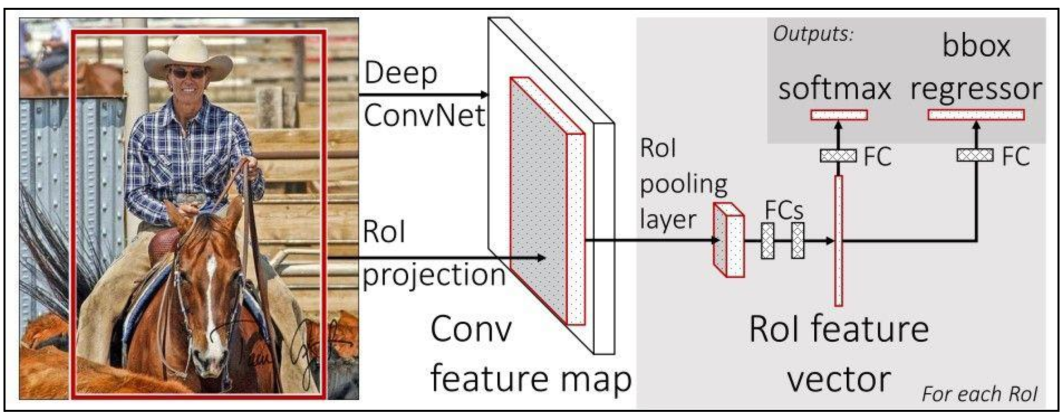
        - Has slow object detection speed.
        - For the fixed size, it needs wrap, which lost of image can be happen.



[https://blog.paperspace.com/faster-r-cnn-explained-object-detection/#:~:text=Faster%20R%2DCNN%20is%20a,the%20locations%20of%20different%20objects.]

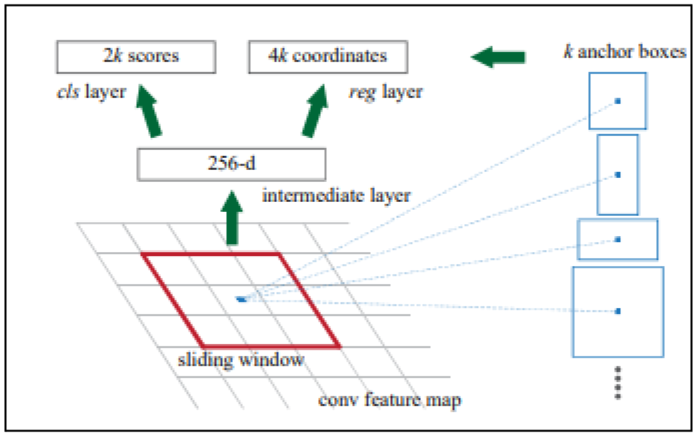
- Fast R-CNN

Fast R-CNN paper [https://arxiv.org/pdf/1504.08083.pdf]
    - Algorithm designed to cover disadvantages of using R-CNN and SPPnet.
    - Fast R-CNN uses RoI Pooling for End-To-End learning.
    - Still uses selective search, which calculation is slow.



[https://blog.paperspace.com/faster-r-cnn-explained-object-detection/#:~:text=Faster%20R%2DCNN%20is%20a,the%20locations%20of%20different%20objects.]

- Faster R-CNN :
- Faster R-CNN paper[https://arxiv.org/pdf/1506.01497.pdf]
- Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks
- https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf
    - Improved model from fast R-CNN.
    - Does not use selective search anymore, but uses Region Proposal Network.
        - Possible to use GPU for RoI calculation.



[https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf]

Data Labeling Software - Matthew

Data labeling?

For machine learning and deep learning modeling processes, data requires specific value. Since a machine learning program needs more than a thousand pieces of data, a programmer needs to spend a large amount of time labeling each image. Therefore, labeling programs will reduce sufficient time.

- Amazon SageMaker : https://aws.amazon.com/ko/sagemaker/data-labeling/?sagemaker-data-wrangler-whats-new.sort-by=item.additionalFields.postDateTime&sagemaker-data-wrangler-whats-new.sort-order=desc
    - Automatic data labeling service made by Amazon.
    - Easy to make data sets.
    - Offer low quality detection (Plus option)
    - Increased visibility of data labeling operations
- YOLO Mark : https://github.com/AlexeyAB/Yolo_mark
    - Tools made for Yolo Neural network.
    - Mark as rectangular box.
- Label Box : https://github.com/Labelbox/Labelbox
    - Fastest data annotation tool.
    - Labeling tool to offer an iterate workflow process for accurate data labeling and to create optimized datasets.
    - Made for machine learning programs. -> Easy to apply in machine learning programs.
    - It can export as csv, json, COCO, VOC, TFRecord. (generally used data set formats). Easy to change format without extra edit.
    - Cons: It uses the web to upload data to create annotations.
        - Only possible to access if there is internet connection
        - Low data security

- CVAT : https://github.com/opencv/cvat
  - Open-source labeling tool for computer vision.
  - Supports image and video annotations.
  - Web based (Google Chrome) tool.
  - Usually used for object detection, image segmentation and image classification.
  - Supports automatic labeling.
  - Cons: takes time to label.

Individual Contributions

Member	Tasks Completed	Hours This Week	Total Hours
Katherine Moretina	Read articles on YOLO, MobileNet, and SqueezeDet. Gained background information on AI, ML and Neural Networks in preparation for algorithm research. Brainstormed criteria for algorithm evaluation	3	7
Matthew Kim	Studied different types of machine learning algorithms, specially Yolo and R-CNN. Read the articles provided by the graduate assistant.	2	6
Chris Cannon	Assisted Ethan with setting up project management. Continued reading articles provided by grad student, brushed up on AI/ML, began research on Pyramid Networks and SSD.	3	5
Ethan	Set up gitlab agile style board for task management and assignment. Began guided research on types of machine learning algorithms and applications to small object detection. Set up team meeting schedule.	4	8

Plans for Coming Week

- Compile research results of Neural Networks, Object Detection, and Machine Learning.
- Discuss evaluation methods for ranking the researched algorithms.
- Continue discussion on whether to implement multiple algorithms or single most applicable algorithm.
- Group research on the highest ranked algorithms.
- Begin discussion of labelling data with Yunsoo Park.
- Continue reviewing code given to us by the graduate assistant

- Gain access to Soybean image repository.
- Have Yunsoo Park walk us through coding on the lab computer.
- Setup Jupyter Notebooks server for student collaboration
- Compare current algorithm findings- this week we will decide on a top 3-5 algorithms we want to use and do more research on those