FACULTY OF ENGINEERING
UNIVERSITY OF ALBERTA

Laboratory of Intelligent Manufacturing, Design and Automation Hybrid is Smart ☺

LIMDA Hybrid is Smart!

Mechanical Engineering

women in scholarship, engineering, science & technology

MOTOROLA

VITA SCIENTIA AMICITIA

A Review of Intelligent Repair Processes for End-of-Life Components

Authors: Eman Shayeb
Ms. Habiba Imam
Ms. Rabiya Abbasi
Dr. Rafiq Ahmad *

moving life forward

An Introduction to Computer Vision

What is Computer Vision?

- Computer vision is a field related to but separate than AI that aims to create machines that are able to perceive the world in the same way that humans do. [1]
- ConvNets are important because they currently serve as the basis of development for computer vision [2]

Where Can I Find Computer Vision In my Daily Life? [3]

- Facial Recognition
- Image classification
- Recommendation Systems
- Natural Language Processing

Convolutional Neural Networks

- ConvNets are Deep Learning Algorithms that can differentiate between parts of an image by attaching weights/biases (significance) to certain parts of said image. [4]
  - The ConvNet is built similarly to the visual cortex of the human brain because of its connectivity and behaviors in the Receptive Field. [5]

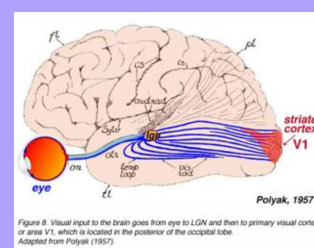
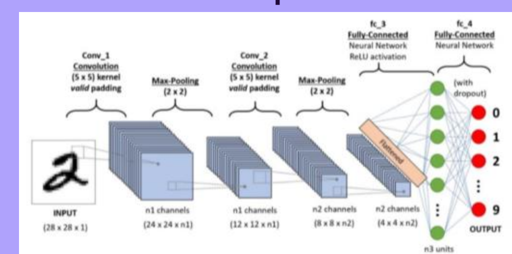


Figure 8. Visual input to the brain goes from eye to LGN and then to primary visual cortex, or area V1, which is located in the posterior of the occipital lobe. Adapted from Polyak (1957).



- ConvNets can be trained to recognize features on their own. [6]

1 Saha, Sumit. “A Comprehensive Guide to Convolutional Neural Networks-the ELI5 Way.” *Medium*, Towards Data Science, 17 Dec. 2018, towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.
2 Ibid
3 ibid
4 ibid
5 Schmolesky, Matthew. “The Primary Visual Cortex by Matthew Schmolesky – Webvision.” *Webvision*, webvision.med.utah.edu/book/part-ix-brain-visual-areas/the-primary-visual-cortex. Accessed 29 July 2020.
6 Saha, Sumit. “A Comprehensive Guide to Convolutional Neural Networks-the ELI5 Way.” *Medium*, Towards Data Science, 17 Dec. 2018, towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.

Eman Shayeb
Realemanshayeb@gmail.com
Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

Project Description

This presentation serves as an exploration of the different kinds of detection models, and their strengths and weaknesses. It will also include applications and functions of said deep learning detection models.

It also serves as justification for the use of Faster R-CNN rather than other architectures for damage detection

Project

This computer vision project uses the Faster R-CNN architecture for damage detection.

Main goal: train Faster R-CNN architecture to detect damage on a cylindrical shape.
Damage will be repaired, and machine will be able to automatically stop when cylinder is fully repaired.

This can be applied to many areas such as detection of different types of damage on different types of objects:

- **Damage detection and repair on pipes (water pipes, oil pipes etc.)**
- **Damage detection on cylindrical building pillars**
- **Other mechanical components**

The Architecture is currently trained to detect prominent areas of damage, but Faster R-CNN can be trained to smaller types of corrosion, cracks, and dents

Research Motivation and Methodology

Motivation

Damage detection architectures can be applied in many different ways to improve daily life + manufacturing projects. One of the most important factors is being able to identify which architectures and programs will allow for the most efficient damage detection.

Methodology

To gather the information for this particular project, the keywords used were “damage detection”, “computer vision”, “convolutional neural networks”.

Due to the virtual nature of this research, digital libraries and databases were the main source of information for it. This is also the reason why my own is largely absent from this project.

The emergence of Convolutional Network Based Detection

CNN's were quite popular in the 1990's, until they fell off during the early 2000s because of the emergence of support vector machines

Krizhevsky et al. released a paper in 2012 that demonstrated CNN's higher image classification accuracy, thus reviving interest in the use of CNN's

Huang et al., Jonathan. "Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors." *Arxiv*, 25 Apr. 2017, arxiv.org/pdf/1611.10012.pdf.

Eman Shayeb
Realemanshayeb@gmail.com
Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

Traditional Object Detection Models

consist of 3 stages:

key: +pro
-con

Informative Region Selection

+scans whole image with multi-scale sliding window.

- is computationally expensive

- too many redundant windows

Feature Extraction

+extracting features provides robust representation

- different backgrounds and illumination mean that creating a descriptor that applies to *all* objects is difficult

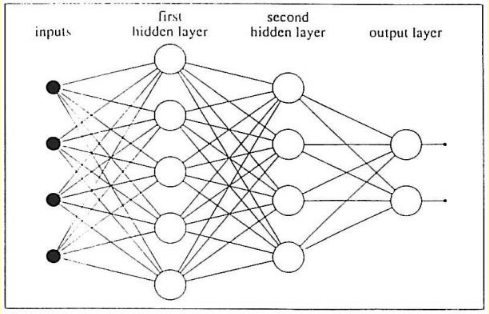
Classification

+separates object from surrounding area

Stewart, Matthew. "Simple Introduction to Convolutional Neural Networks." *Towards Science*, 26 Feb. 2019, towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac.

What Makes CNN's better than MLP's?

Convolutional Neural Networks (CNN)	Multilayer Perceptron (MLP)
• Linear transformation invariant • depth is created (nearby pixels are more strongly related than far ones) • usage of filters allows for feature identification • It is very unlikely for the CNN to learn the same feature because filters are random numbers that continuously update as the CNN is trained. (unless # of filters is really large)	• Not Linear transformation invariant • Loss of spatial information, this means that images would not be able to be defined as clearly • Many weights and biases that make MLP's more computationally heavy than CNN's



Kirkegaard, Poul Henning. "Use of a Neural Network for Damage Detection and Location in a Steel Member - Danish National Research Database." *Forsknings Database*, 29 July 1992, www.forskningsdatabasen.dk/en/catalog/2389380154.

Eman Shayeb
Realemanshayeb@gmail.com
Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

CNN Architecture

layers called **feature maps**

Typical CNNs have...

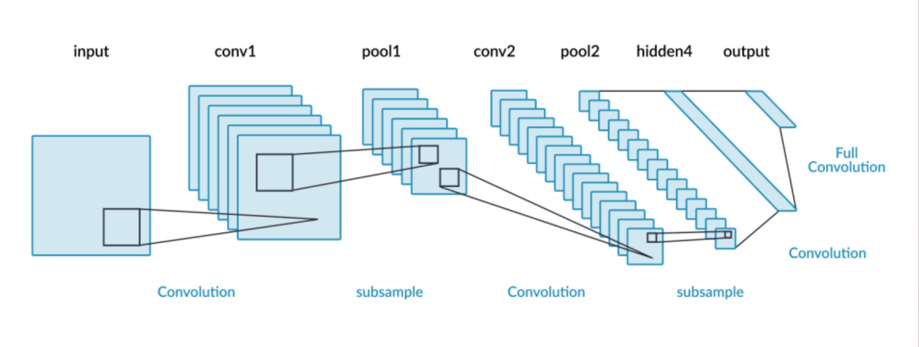
Receptive field -- each neuron is connected to neurons of adjacent layers.

input layer -- 3D matrix of pixel intensities for different colour channels

internal layer -- multichannel image

Ren, Shaoqing. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *ArXiv.Org*, 4 June 2015, arxiv.org/abs/1506.01497.

"Convolutional Neural Network Architecture: Forging Pathways to the Future." *MissingLink.Ai*, missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future. Accessed 29 July 2020.



Terms:

Convolution layer - creates a prediction of each feature of the image through the creation of a feature map. This is done through a filter that scans the image, pixel by pixel.

Pooling layer – After the conv layer creates the information from the image, the pooling layer will keep the necessary information.

Fully Connected Input Layer – flattens previous outputs to form a single vector, which will be input into the next layer.

Fully connected output layer – Final touches; determines final probabilities for classifying an image.

R-CNN and Fast R-CNN

key: +pro
-con

R-CNN [9]

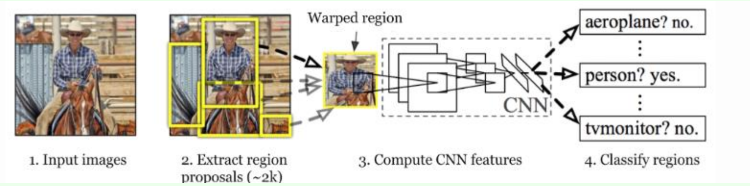
This was an object detection model introduced in a paper by Ross Girshek et al. in 2014.

+The model essentially cuts the image into pieces for more efficient detection/localization. (segmentation)

-2000 region proposals per image

-takes about 49 seconds to detect a single image

-computationally expensive and uses up a lot of disk space



Fast R-CNN [10]

This is an improved object detection model introduced in a paper by Ross Girshek et al. in 2015

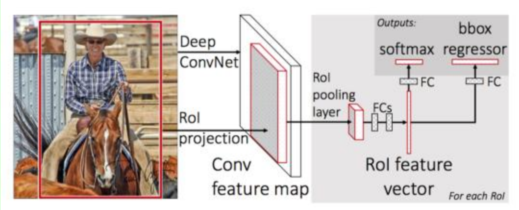
+ All region proposals are input together as one forward propagation

+ combines things like ConvNet and classification layers which saves disk space by avoiding the need to save

+ training/detection time are decreased from R-CNN counter part

+ from 84 hours to 8.75 hours

- selective search takes up most of the computation time

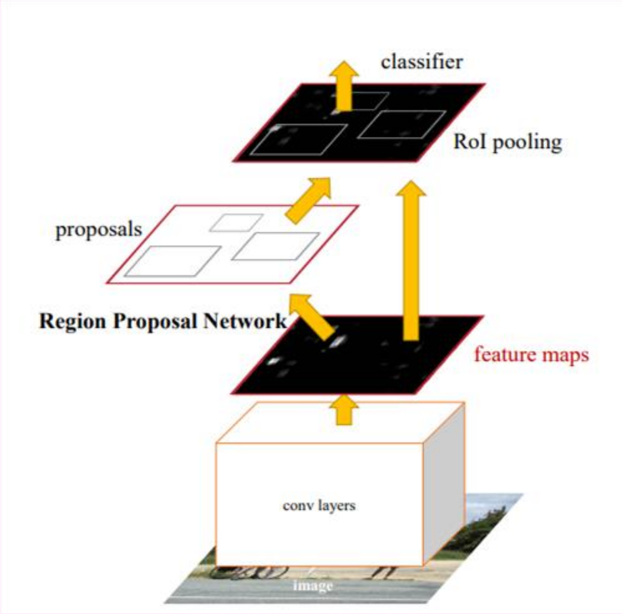


Girshick, Ross. "Fast R-CNN." *ArXiv.Org*, 30 Apr. 2015, arxiv.org/abs/1504.08083.
"Girshick, Ross Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." *ArXiv.Org*, 11 Nov. 2013, arxiv.org/abs/1311.2524.

Eman Shayeb
Realemanshayeb@gmail.com

Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

Faster R-CNN



Faster R-CNN detection process:

1. Insert image.
2. Image passes through a trained CNN to create a feature map.
3. Region Proposal Network will use the features computed (in step 2) to create bounding boxes.
4. Create list of probabilities for potential objects + their location in image
5. Apply Region of Interest Pooling – this picks out features that relate to (relevant) objects
6. R-CNN kicks in and classifies objects in bounding boxes

Faster R-CNN

- Can learn informative object representations without designing features manually
- Not as computationally heavy as other architectures
- Object detection architecture that functions very close to real time
- Learns with deeper architecture
- increased expressive capabilities
- Introduces Regional Proposal Networks to Fast R-CNN

Ren, Shaoqing. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *ArXiv.Org*, 4 June 2015, arxiv.org/abs/1506.01497.

Eman Shayeb
Realemanshayeb@gmail.com
Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

More Detail

Many detection architectures have been based on Faster R-CNN, and other architectures could arguably have more “advanced” features, but that is not necessary for this particular project because the damage is usually large without much variance -- making it easier for the system to detect.

- Example 1: MASK R-CNN is based on Faster R-CNN, but it adds onto the pre-existing branch for bounding boxes with branch for instance/semantic segmentation *which is not required for this project.*

He, Kaiming. “Mask R-CNN.” *Openaccess*, 11 Nov. 2013, openaccess.thecvf.com/content_ICCV_2017/papers/He_Mask_R-CNN_ICCV_2017_paper.pdf.

Eman Shayeb Realemanshayeb@gmail.com
Dr Rafiq Ahmad https://sites.ualberta.ca/~rafiq1/

Analysis

Key: minimal issues 🟢 Runs into some issues 🟡 Runs into many issues 🔴

	REAL TIME IMAGE PROCESSING	DAMAGE QUANTIFICATION	DAMAGE DETECTION	DAMGE IDENTIFICATION		
				CORROSION	CRACK	DENT
R-CNN	🔴	🟢	🟡	🟡	🟢	🟢
FAST R-CNN	🟡	🟢	🟢	🟡	🟢	🟢
FASTER R-CNN	🟢	🟢	🟢	🟡	🟢	🟢
MASK R-CNN	🟢	🟢	🟢	🟢	🟢	🟢
YOLO	🟢	🟡	🟡	🟡	🟡	🟡
SSD	🟢	🟢	🟢	🟢	🟢	🟢
R-FCN	🟢	🟢	🟢	🟢	🟢	🟢

- https://www.forskningsdatabasen.dk/en/catalog/2389380154https://arxiv.org/pdf/1811.04535.pdf
- https://www-spiedigitallibrary-org.login.ezproxy.library.ualberta.ca/conference-proceedings-of-spie/10598/2295954/Deep-faster-R-CNN-based-automated-detection-and-localization-of/10.1117/12.2295954.full?SSO=1
- https://onlinelibrary-wiley-com.login.ezproxy.library.ualberta.ca/doi/epdf/10.1002/stc.2381
- https://openaccess.thecvf.com/content_ICCV_2017/papers/He_Mask_R-CNN_ICCV_2017_paper.pdf
- https://www-sciencedirect-com.login.ezproxy.library.ualberta.ca/science/article/pii/S0926580519302663
- https://data.smar-conferences.org/SMAR_2019_Proceedings/Inhalt/th.3.a.6.pdf
- https://search-proquest-com.login.ezproxy.library.ualberta.ca/docview/2388649388?accountid=14474
- https://onlinelibrary-wiley-com.login.ezproxy.library.ualberta.ca/doi/full/10.1111/mice.12367
- https://ieeexplore.ieee.org/abstract/document/8937176/keywords#keywords
- https://www.researchgate.net/profile/Young-Jin_Cha/publication/321341582_Autonomous_Structural_Visual_Inspection_Using_Region-Based_Deep_Learning_for_Detecting_Multiple_Damage_Types/links/5a1dbbaca6fdccc6b7f848d0/Autonomous-Structural-Visual-Inspection-Using-Region-Based-Deep-Learning-for-Detecting-Multiple-Damage-Types.pdf
- https://arxiv.org/pdf/1406.4729.pdf
- https://ieeexplore-ieee-org.login.ezproxy.library.ualberta.ca/stamp/stamp.jsp?tp=&arnumber=8906437
- file:///C:/Users/Eman%20Shayeb/Downloads/applsci-08-01678.pdf
- https://arxiv.org/pdf/1801.09454.pdf

Eman Shayeb
Realemanshayeb@gmail.com

Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

Speed/Accuracy Trade-Off

A study done by Jonathan Huang et al. in 2017 compares multiple trade-offs for 3 main convolutional object detectors, Faster R-CNN, R-FCN, and SSD. It is important to note that SSD and R-FCN are based on Faster R-CNN.

	Faster R-CNN	R-FCN	SSD
Sensitivity to feature extraction quality	More sensitive than SSD	More sensitive than SSD	Less sensitive than Faster R-CNN and R-FCN
Speed and accuracy	Slower than the other two, but more accurate	Faster but less accurate than Faster R-CNN	Faster but less accurate than Faster R-CNN
Object Size	Scales according to size of object, computes more pixels	Scales according to size of object, computes more pixels	Resizes images, computes less pixels
Small Object	Performs well	Performs well	Performs very poorly
Large Object	Performs well	Performs well	Performs well
Image Resolution	Performance accuracy decreases in low resolution, and vice versa	Performance accuracy decreases in low resolution, and vice versa	Performance accuracy decreases in low resolution, and vice versa
Proposal Number	Less proposals will speed it up and maintain accuracy	Less proposals will speed it up and maintain accuracy	Less proposals will speed it up and maintain accuracy

Sources

- https://arxiv.org/pdf/1611.10012.pdf
- https://ieeexplore-ieee-org.login.ezproxy.library.ualberta.ca/stamp/stamp.jsp?tp=&arnumber=8750729

Eman Shayeb
Realemanshayeb@gmail.com

Dr Rafiq Ahmad
https://sites.ualberta.ca/~rafiq1/

Other Real Life Applications

Traffic Sign Detection (Faster R-CNN)

- Helps with unmanned/automatic driving
- A study done by Faming Shao et al. uses faster R-CNN to reduce computing time for traffic sign detection

Shao, Faming. "Improved Faster R-CNN Traffic Sign Detection Based on a Second Region of Interest and Highly Possible Regions Proposal Network." *PubMed Central (PMC)*, 1 May 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6567367.

Smart Nursery (Faster R-CNN)

- Somnuk Phon-Amnuaisuk et al. explored the development of a Smart Nursery with Faster R-CNN
- This includes detecting the baby and what the baby is doing (sleeping, crawling, playing etc.)

---. "Exploring the Applications of Faster R-CNN and Single-Shot Multi-Box Detection in a Smart Nursery Domain." *ArXiv.Org*, 27 Aug. 2018, arxiv.org/abs/1808.08675.

Vehicle Detection (YOLO and SSD)

- Mehdi Masmoudi et al. explored various detection models that were able to detect vehicles in real time
- Their findings can be used for increasing the efficiency vehicle/object detection in fully automated driving

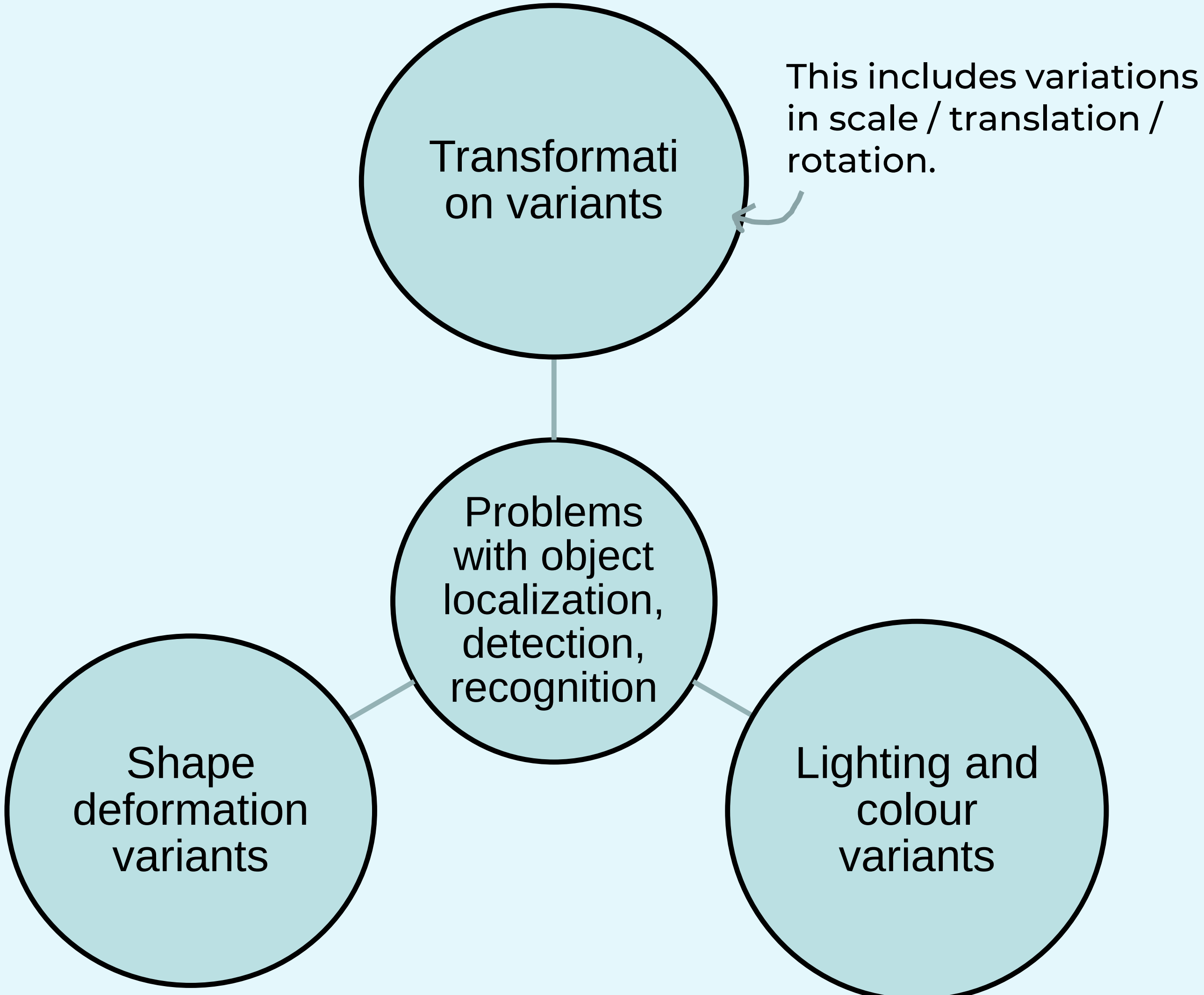
Masmoudi, Mehdi. "Object Detection Learning Techniques for Autonomous Vehicle Applications." *Ieeexplore*, 2019, ezpa.library.ualberta.ca/ezpAuthen.cgi?url=https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8906437.

Aerial Vehicle Imagery (Faster R-CNN, R-FCN, SSD, and YOLO)

- Suk-Ju Hong et al. published a study that used multiple detection models to track and count birds in their natural habitats
- Their development of these models can be applied to future aerial imagery with varying resolutions and flight altitudes.

Hong, Suk-Ju. "Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery." *MDPI*, 6 Apr. 2019, www.mdpi.com/1424-8220/19/7/1651/htm.

Common Issues of Computer Vision



Phon-Amnuaisuk, Somnuk. “Exploring the Applications of Faster R-CNN and Single-Shot Multi-Box Detection in a Smart Nursery Domain.” *Arixv.Org*, 27 Aug. 2018, arxiv.org/pdf/1808.08675.pdf.

Conclusion

- CNN's form the basis of modern computer vision technologies and regained popularity in 2012
- Traditional Object detection models, like MLP's, have 3 distinct stages: Region Selection, Feature Extraction, Classification
- R-CNN → combine ConvNet and classification layers to save disk space → Fast R-CNN → adjusting selective search method → Faster R-CNN
- Faster R-CNN is used for this particular project because of its speed and accuracy, but also relative simplicity compared to models like MASK R-CNN
- These models can be applied in many areas like traffic detection for unmanned driving to monitoring animals to creating smart nurseries
- There tend to be 3 main variants that must be overcome by computer vision: Transformation, shape, lighting/colour

I really enjoyed doing research on computer vision this Summer! I will definitely continue exploring this field in the future, and I am excited to see what new technologies develop over time!

Acknowledgements

- I would like to thank my sponsors, the Edmonton Chapter of Beta Sigma Phi and Motorola – without them I would not have been able to participate in the program in the first place.

- I would also like to thank WISEST for giving me this once in a lifetime opportunity to participate in such a unique program, and for adapting so seamlessly to the virtual program (it almost feels like I was on campus).

- I would additionally like to thank the LIMDA team for helping me throughout my research – particularly Dr Rafiq Ahmad, Ms. Habiba Iman, and Ms. Rabiya Abbasi for supervising and aiding me throughout this process!

- Thank you for a wonderful Summer of making connections and learning a ton!

References

- cha, yoong-jin. “Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types.” *Researchgate*, 2017, www.researchgate.net/profile/Young-Jin_Cha/publication/321341582_Autonomous_Structural_Visual_Inspection_Using_Region-Based_Deep_Learning_for_Detecting_Multiple_Damage_Types/links/5a1dbbaca6fdccc6b7f848d0/Autonomous-Structural-Visual-Inspection-Using-Region-Based-Deep-Learning-for-Detecting-Multiple-Damage-Types.pdf.
- Cho, Soojin, and Byunghyun Kim. “Mage-based Concrete Crack Assessment Using Mask Andregion-based Convolutional Neural Network.” *Onlinelibrary-Wiley*, 28 Apr. 2019, onlinelibrary-wiley-com.login.ezproxy.library.ualberta.ca/doi/epdf/10.1002/stc.2381.
- Deng, Lu. “Region-Based CNN Method with Deformable Modules for Visually Classifying Concrete Cracks.” *Proquest*, 7 Apr. 2020, search-proquest-com.login.ezproxy.library.ualberta.ca/docview/2388649388?accountid=14474.
- Fujia, Wei. “Instance-Level Recognition and Quantification for Concrete Surface Bughole Based on Deep Learning.” *Science Direct*, Nov. 2019, www-sciencedirect-com.login.ezproxy.library.ualberta.ca/science/article/pii/S0926580519302663.
- Gayhun, Suh, and Cha Young-Jin. “Deep Faster R-CNN-Based Automated Detection and Localization of Multiple Types of Damage.” *Spiedigitallibrary*, 27 Mar. 2018, ezpa.library.ualberta.ca/ezpAuthen.cgi?url=https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10598/2295954/Deep-faster-R-CNN-based-automated-detection-and-localization-of/10.1117/12.2295954.full?SSO=1.
- Girshick, Ross. “Fast R-CNN.” *ArXiv.Org*, 30 Apr. 2015, arxiv.org/abs/1504.08083.
- ---. “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation.” *ArXiv.Org*, 11 Nov. 2013, arxiv.org/abs/1311.2524.
- He, Kaiming. “Mask R-CNN.” *Openaccess*, 11 Nov. 2013, openaccess.thecvf.com/content_ICCV_2017/papers/He_Mask_R-CNN_ICCV_2017_paper.pdf.
- ---. “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition.” *Arxiv.Org*, 23 Apr. 2015, arxiv.org/pdf/1406.4729.pdf.
- Hong, Suk-Ju. “Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery.” *MDPI*, 6 Apr. 2019, www.mdpi.com/1424-8220/19/7/1651/htm.
- Huang et al., Jonathan. “Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors.” *Arxiv*, 25 Apr. 2017, arxiv.org/pdf/1611.10012.pdf.
- Kirkegaard, Poul Henning. “Use of a Neural Network for Damage Detection and Location in a Steel Member - Danish National Research Database.” *Forsknings Database*, 29 July 1992, www.forskningsdatabasen.dk/en/catalog/2389380154.
- Li, Yiting. “Research on a Surface Defect Detection Algorithm Based on MobileNet-SSD.” *MDPI*, 17 Sept. 2018, www.mdpi.com/2076-3417/8/9/1678.
- Maeda, Hiroya. “Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone.” *ArXiv.Org*, 29 Jan. 2018, arxiv.org/abs/1801.09454.
- Masmoudi, Mehdi. “Object Detection Learning Techniques for Autonomous Vehicle Applications.” *Ieeexplore*, 2019, ezpa.library.ualberta.ca/ezpAuthen.cgi?url=https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8906437.
- ---. “Object Detection Learning Techniques for Autonomous Vehicle Applications.” *Ieeexplore*, 2019, ieeexplore-ieee-org.login.ezproxy.library.ualberta.ca/stamp/stamp.jsp?tp=&arnumber=8906437.
- Phon-Amnuaisuk, Somnuk. “Exploring the Applications of Faster R-CNN and Single-Shot Multi-Box Detection in a Smart Nursery Domain.” *Arixv.Org*, 27 Aug. 2018, arxiv.org/pdf/1808.08675.pdf.
- Ren, Shaoqing. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” *ArXiv.Org*, 4 June 2015, arxiv.org/abs/1506.01497.
- Schmolesky, Matthew. “The Primary Visual Cortex by Matthew Schmolesky – Webvision.” *Webvision*, webvision.med.utah.edu/book/part-ix-brain-visual-areas/the-primary-visual-cortex. Accessed 29 July 2020.
- Shao, Faming. “Improved Faster R-CNN Traffic Sign Detection Based on a Second Region of Interest and Highly Possible Regions Proposal Network.” *PubMed Central (PMC)*, 1 May 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6567367.
- Stewart, Matthew. “Simple Introduction to Convolutional Neural Networks.” *Towards Science*, 26 Feb. 2019, towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac.
- Ukhwah, Ernin Nestwatul. “Asphalt Pavement Pothole Detection Using Deep Learning Method Based on YOLO Neural Network - IEEE Conference Publication.” *IEEE XPLORE*, 1 Aug. 2019, ieeexplore.ieee.org/abstract/document/8937176/keywords#keywords.
- Wu, Zheng Yi. “Deep Learning-Based Defect Detection and Assessment for Engineering Structures.” *Smar Conferences*, 2019, data.smar-conferences.org/SMAR_2019_Proceedings/Inhalt/th.3.a.6.pdf.
- Xue, Yadong. “A Fast Detection Method via Region-Based Fully Convolutional Neural Networks for Shield Tunnel Lining Defects.” *Wiley Online Library*, 23 Mar. 2018, onlinelibrary-wiley-com.login.ezproxy.library.ualberta.ca/doi/full/10.1111/mice.12367.

Thank you ☺

Any Questions?

Eman Shayeb
Realemanshayeb@gmail.com

Dr Rafiq Ahmad