AI computation research

Seok-Bum Ko, Ph.D., P.Eng.
Professor, Department of Electrical and Computer Engineering
Professor, Division of Biomedical Engineering
University of Saskatchewan, Canada
November 22, 2023

University of Saskatchewan

Where is Saskatchewan?

University of Saskatchewan

UNIVERSITY OF SASKATCHEWAN

- U15.ca
- 2 Nobel prize winners
- World leading researches: GIFS, GIWS, CLS, VIDO, etc
- www.usask.ca
- Andrew Donald Booth (1961-1972, UofS), https://history.computer.org/pioneers/booth-ad.html

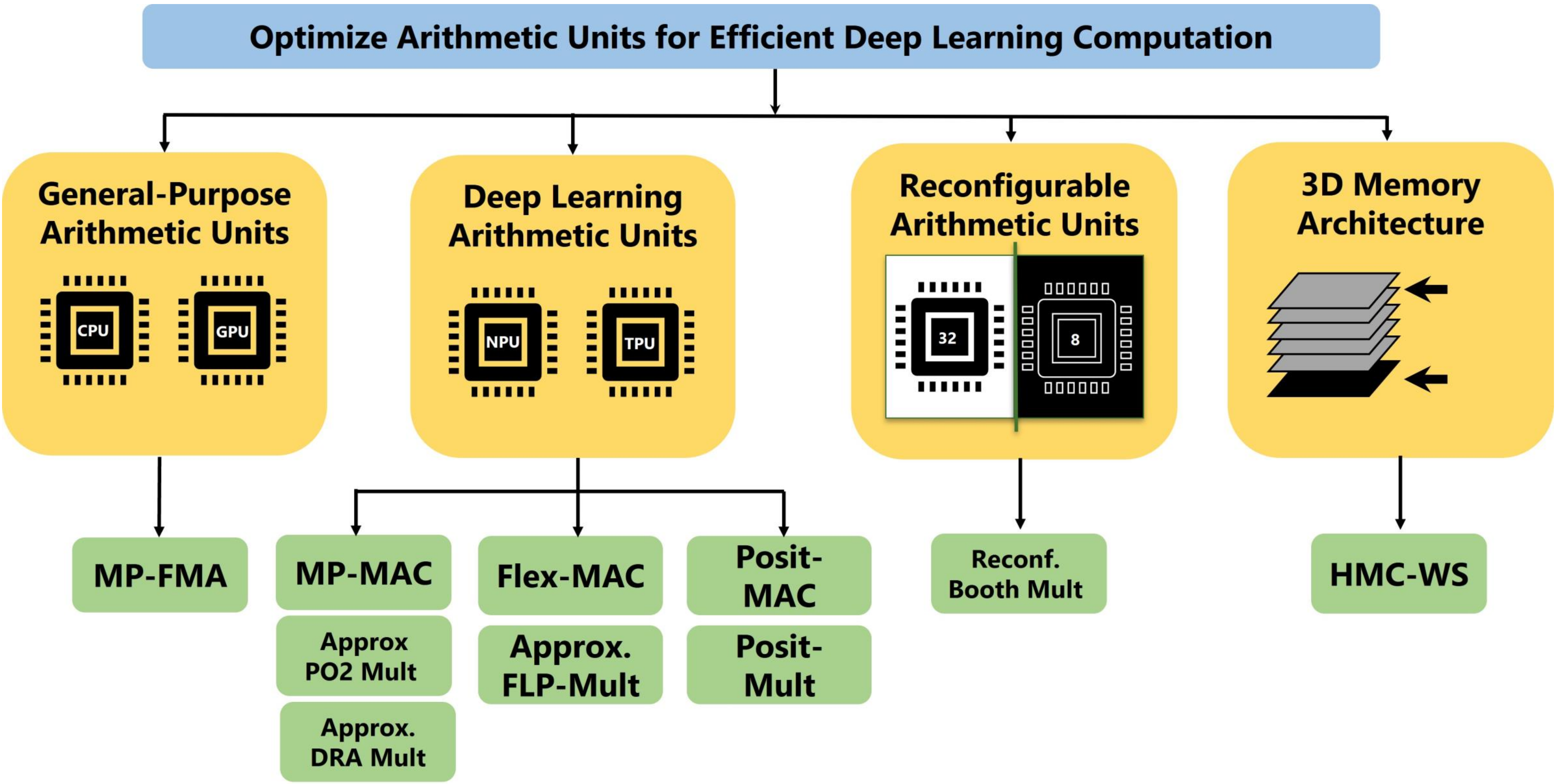
- Seokbum Ko, Ph.D., P.Eng.
  – Professor, Dept of ECE
  – Professor, Div. of Biomedical Engineering
  – Research interests: Computer architecture, FPGA/ASIC implementation of compute intensive applications, deep learning and biomedical engineering
  – Associate Editors, IEEE Transactions on VLSI, IEEE Transactions on Circuits and Systems I, IEEE Access, IET Computers & Digital Techniques
  – IEEE CAS Technical Committee, IEEE P3109, IEEE754-2029, IEEE Senior member

Contents

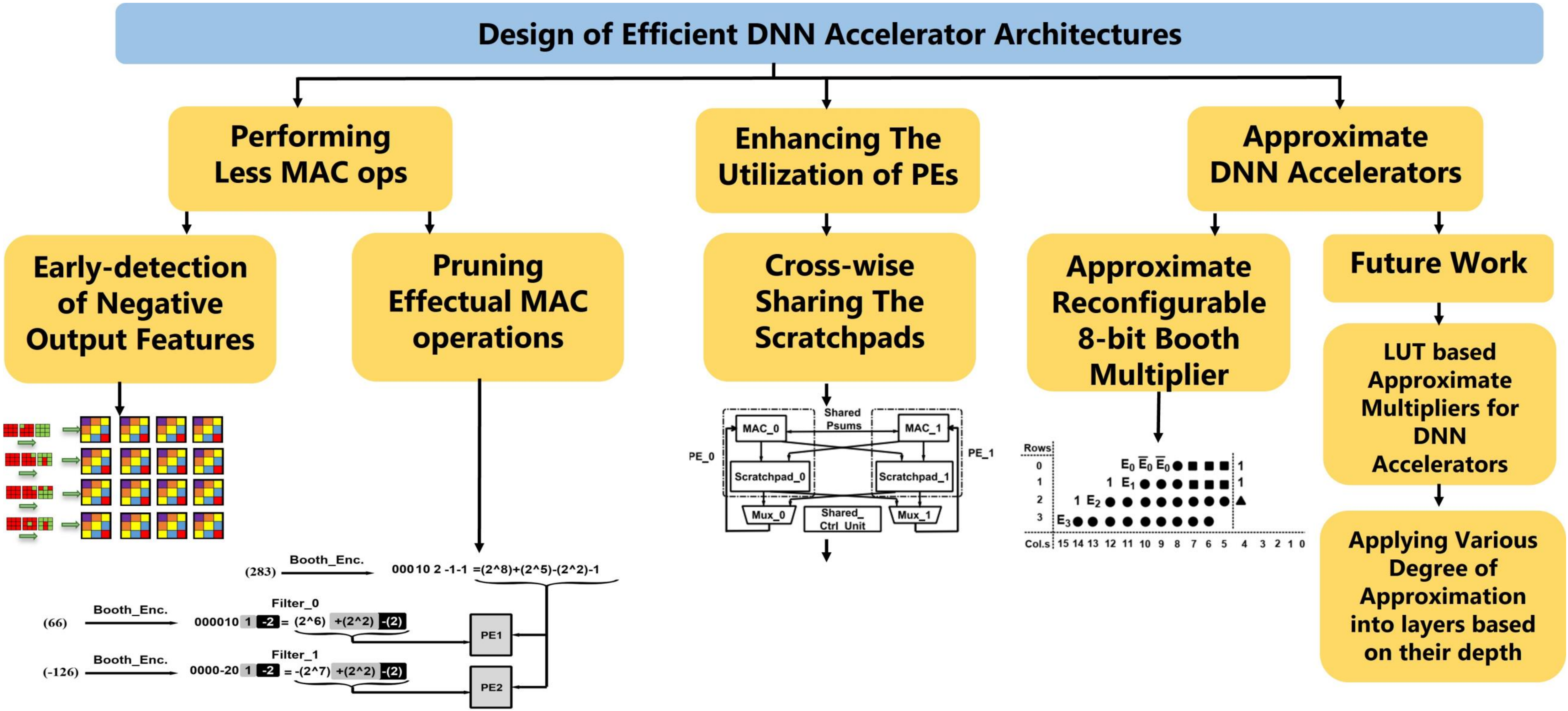
- **Hardware Design for Deep Learning**
  - Hardware Arithmetic Units for Deep Learning
  - FPGA-Based Convolutional Neural Network Accelerator
  - Approximate Computing
  - RISC-V Processor
  - Accelerating Deep Neural Network
  - Efficient Spiking Neural Network Processing

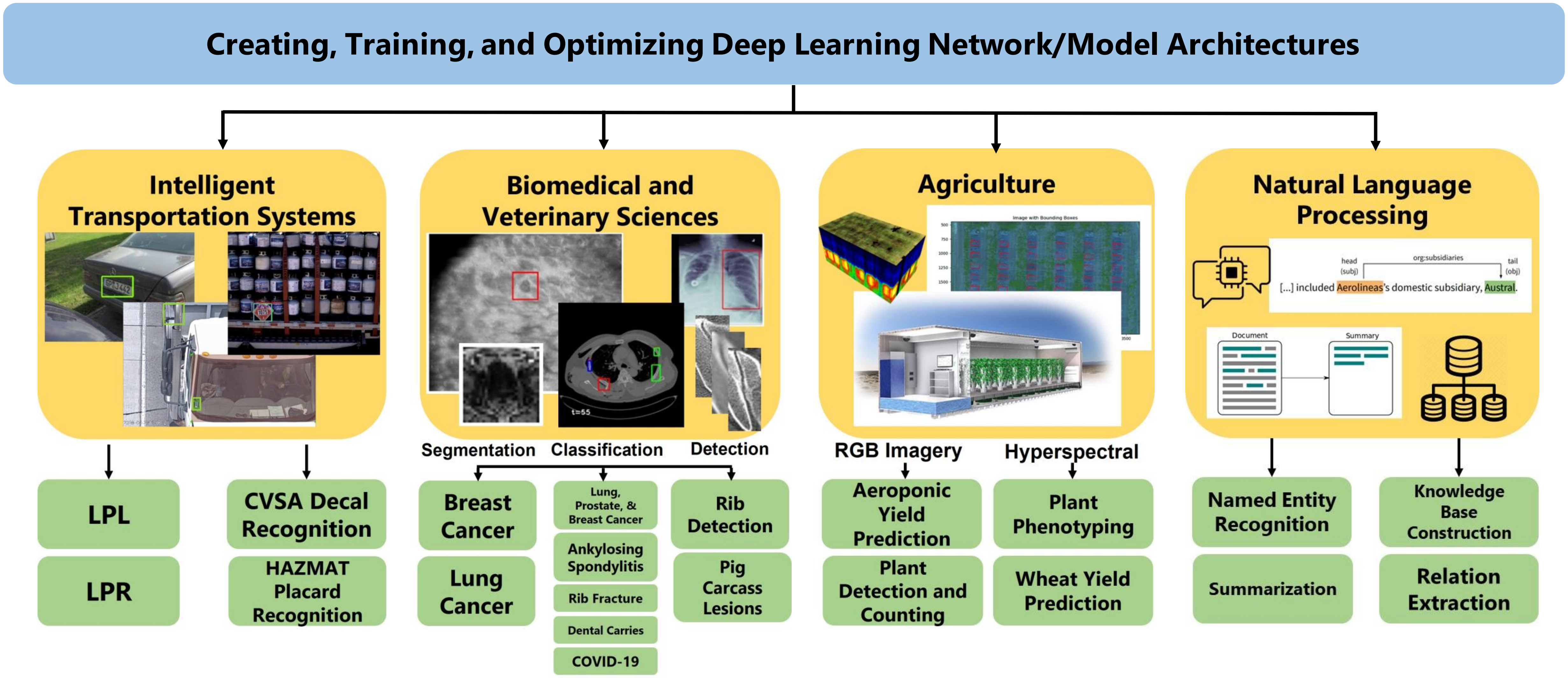
- **Security and Cryptography**
  - High throughput and area-efficient implementation of AES
  - Area efficient nano-AES implementation for IoT devices
  - Post-Quantum Cryptosystem for IoT Resource-Constrained Devices

- **Deep learning applications**
  - Retinal Blood Vessel Segmentation
  - Deep Learning for Classification of Small (≤ 2 cm) pulmonary nodules on CT
  - DL for Improved Pulmonary Nodule Classification on CT
  - Breast Cancer Classification in ABUS using Multi-view CNN with Transfer Learning
  - Early Detection of Ankylosing Spondylitis using Machine and Deep Learning
  - Rib Tracking, Labelling, and Classification from Axial CT Scans
  - Deep Learning for Fast Detection of Explosives
  - Yield Prediction in Aeroponics
  - Plant phenotyping from Hyperspectral Images
  - License Plate Localization and Recognition
  - Commercial Vehicle Characteristics
  - Super Resolution with Deep Learning
  - Biomedical Signal Processing
  - DL for Dental Caries Classification
  - Covid-19 Diagnosis in CXRs
  - Academic Article Information Extraction



Optimize Arithmetic Units for Efficient Deep Learning Computation
General-Purpose Arithmetic Units
CPU
GPU
Deep Learning Arithmetic Units
NPU
TPU
Reconfigurable Arithmetic Units
32
8
3D Memory Architecture
MP-FMA
MP-MAC
Approx PO2 Mult
Approx. DRA Mult
Flex-MAC
Approx. FLP-Mult
Posit-MAC
Posit-Mult
Reconf. Booth Mult
HMC-WS

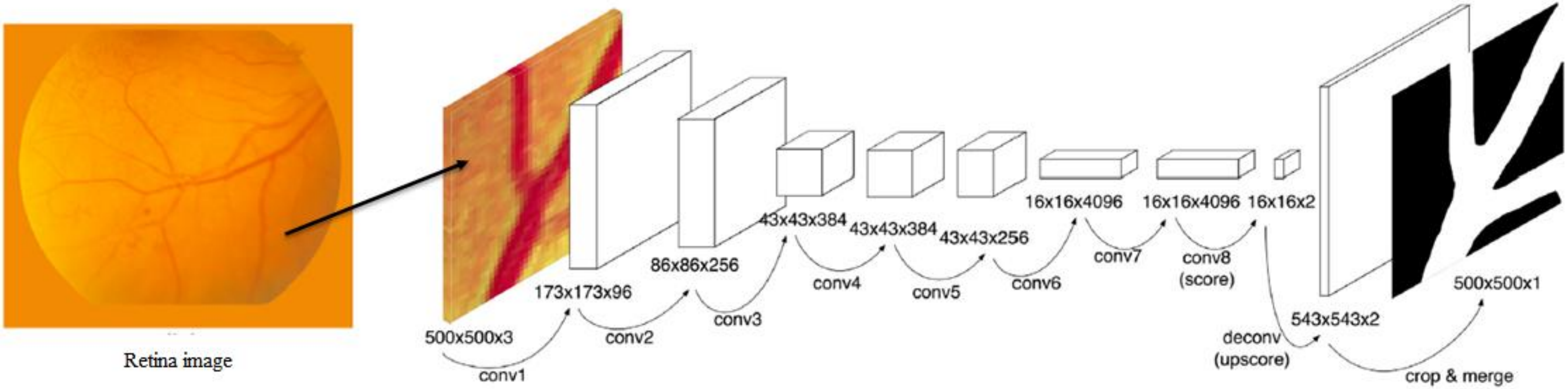

Design of Efficient DNN Accelerator Architectures
Performing Less MAC ops
Enhancing The Utilization of PEs
Approximate DNN Accelerators
Early-detection of Negative Output Features
Pruning Effectual MAC operations
Cross-wise Sharing The Scratchpads
Approximate Reconfigurable 8-bit Booth Multiplier
Future Work
LUT based Approximate Multipliers for DNN Accelerators
Applying Various Degree of Approximation into layers based on their depth
(283) Booth_Enc. 00010 2 -1-1 =(2^8)+(2^5)-(2^2)-1
(66) Booth_Enc. Filter_0 000010 1 -2 = (2^6) +(2^2) -(2)
(-126) Booth_Enc. Filter_1 0000-20 1 -2 = -(2^7) +(2^2) -(2)
PE1
PE2
Shared Psums
MAC_0
MAC_1
PE_0
PE_1
Scratchpad_0
Scratchpad_1
Mux_0
Mux_1
Shared_ Ctrl_Unit
Rows
0
1
2
3
Col.s 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0



Creating, Training, and Optimizing Deep Learning Network/Model Architectures
Intelligent Transportation Systems
LPL
LPR
CVSA Decal Recognition
HAZMAT Placard Recognition
Biomedical and Veterinary Sciences
Segmentation
Classification
Detection
Breast Cancer
Lung Cancer
Lung, Prostate, & Breast Cancer
Ankylosing Spondylitis
Rib Fracture
Dental Carries
COVID-19
Rib Detection
Pig Carcass Lesions
Agriculture
Image with Bounding Boxes
RGB Imagery
Hyperspectral
Aeroponic Yield Prediction
Plant Detection and Counting
Plant Phenotyping
Wheat Yield Prediction
Natural Language Processing
head (subj)
org:subsidiaries
tail (obj)
[...] included Aerolineas's domestic subsidiary, Austral.
Document
Summary
Named Entity Recognition
Summarization
Knowledge Base Construction
Relation Extraction

Deep Learning for Retinal Imaging

Retinal blood vessel segmentation using fully convolutional network with transfer learning



- A full-size retina image is divided into multiple 50 x 50 image patches.
- Each image patch is fed into fully convolutional AlexNet for retinal vessel segmentation.

Z. Jiang, H. Zhang, Y. Wang and S. Ko*, 2018. "Retinal blood vessel segmentation using fully convolutional network with transfer learning," *Elsevier Computerized Medical Imaging and Graphics*. Vol. 68, pp. 1-15.

Retinal blood vessel segmentation using fully convolutional network with transfer learning

Performance comparison with the related deep learning based works on DRIVE database.

Algorithm	Accuracy	Sensitivity	Specificity	AUC
Wang et al. (2015)[1]	0.9533	0.8173	0.9733	0.9475
Soomro et al. (2017)[2]	0.9480	0.7460	0.9170	0.8310
Li et al. (2016)[3]	0.9527	0.7569	0.9816	0.9738
Proposed method	0.9624	0.7540	0.9825	0.9810

Performance comparison with the related deep learning based works on STARE database.

Algorithm	Accuracy	Sensitivity	Specificity	AUC
Wang et al. (2015)[1]	0.9621	0.8104	0.9791	0.9751
Soomro et al. (2017)[2]	0.9470	0.7480	0.9220	0.8350
Li et al. (2016)[3]	0.9628	0.7726	0.9844	0.9879
Proposed method	**0.9734**	**0.8352**	**0.9846**	**0.9900**

[1]Wang, S., Yin, Y., Cao, G., Wei, B., Zheng, Y., Yang, G., 2015. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing* 149 (PB), 708–717.

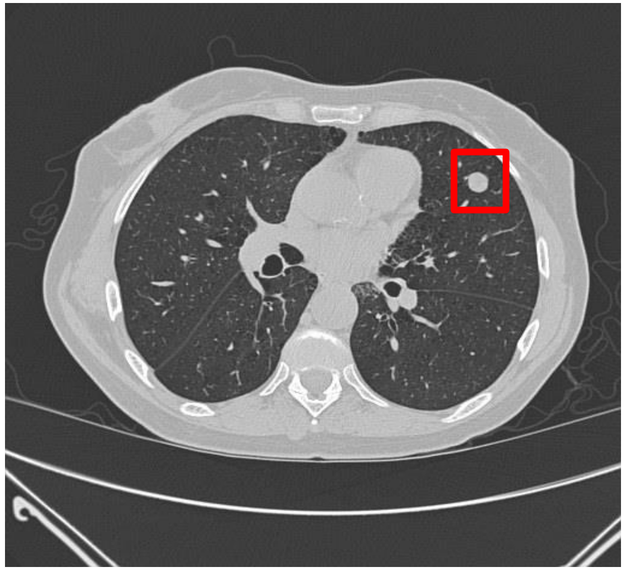
[2]Soomro, T.A., Afifi, A.J., Gao, J., Hellwich, O., Khan, M.A.U., Paul, M., Zheng, L., 2017. Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network. *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* 1–8.

[3]Li, Q., Feng, B., Xie, L., Liang, P., Zhang, H., Wang, T., 2016. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Trans. Med. Imaging* 35 (1), 109–118.

UNIVERSITY OF SASKATCHEWAN

Lung Nodule Classification on Computed Tomography

- Computed Tomography (CT) is helpful in reducing the lung cancer
  - High sensitivity to detect lung nodule
  - Easy to locate lung nodule
- Different types of nodules have ***similar visual representation***

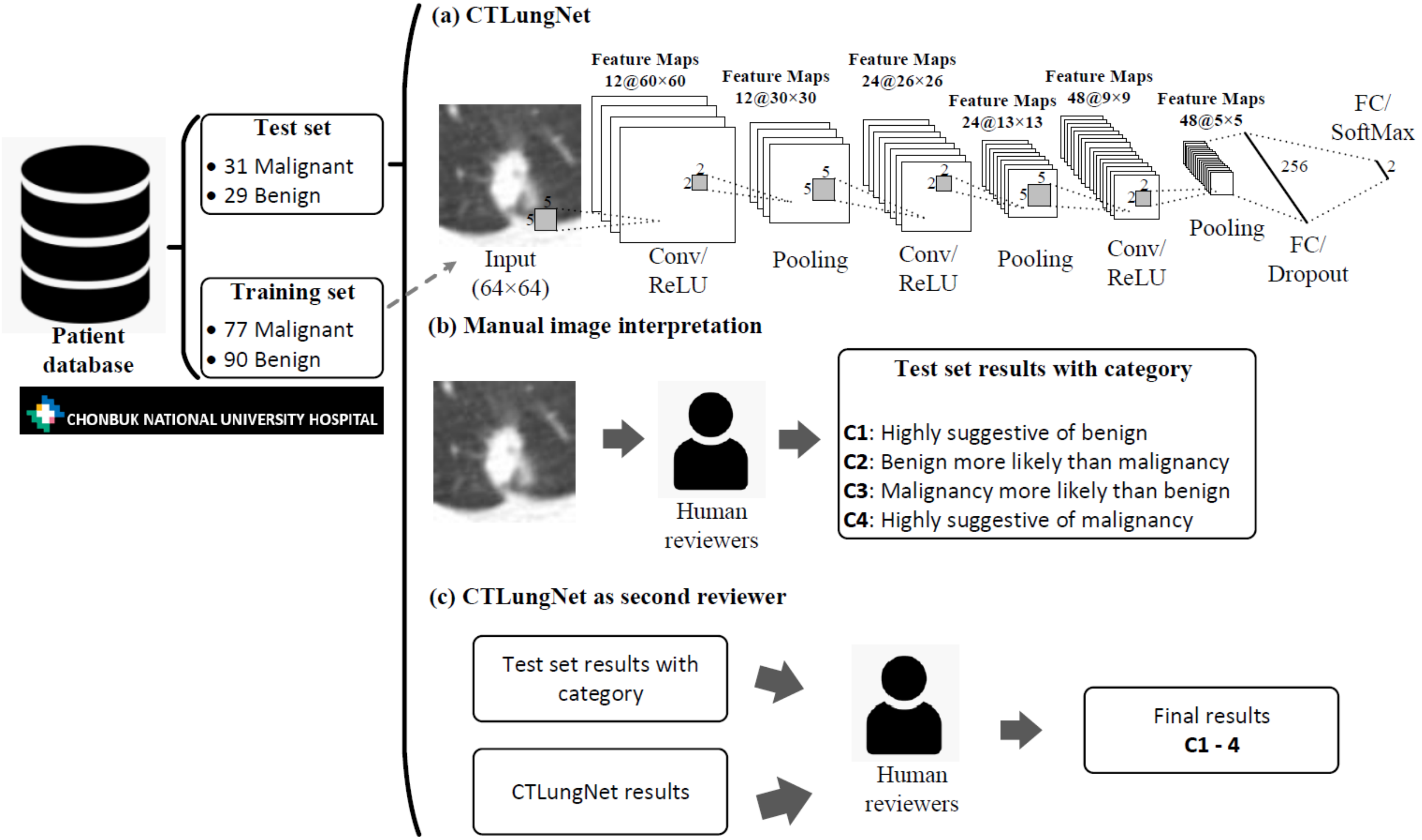


Benign nodule

Malignant nodule

*According to seminal National Lung Screening Tail, screening CT scan reduce lung cancer death in **20%**, but **96.4%** of nodule findings showed a false positive*

Deep Learning for Classification of Small (≤ 2 cm) pulmonary nodules on CT Imaging



- CTLungNet
  - 3 Conv layers and 2 FC layers
  - 10-times-faster processing time (0.90 sec/slice) compared to AlexNet (8.79 sec/slice)
  - 180-times-reduced number of learnable parameters (0.34 million) compared to AlexNet (61 million)

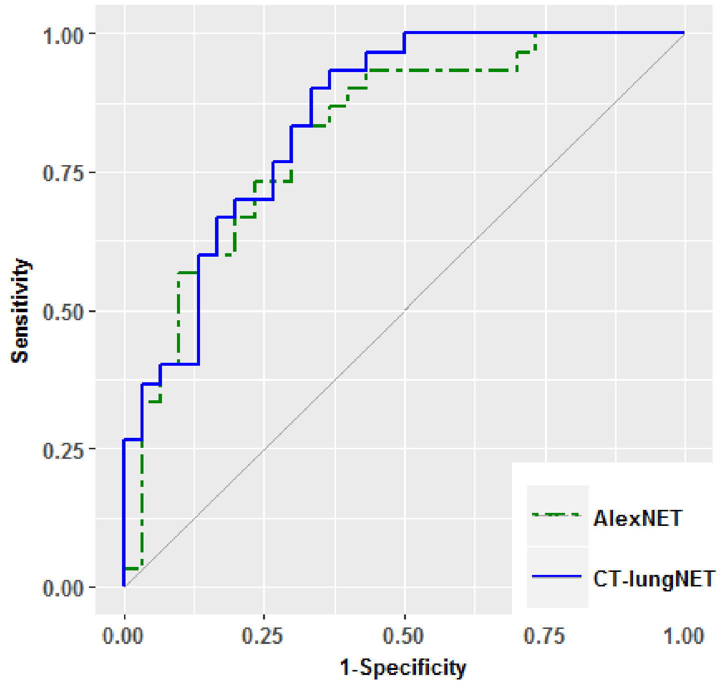
K. Chae, G. Jin, S. Ko, Y. Wang, H. Zhang, E. Choi, H. Choi, 2020. "Deep Learning for Classification of A Small (≤2cm) Pulmonary Nodule on CT Imaging: A Preliminary Study," accepted to *Elsevier Academic Radiology*, Vol. 27, Iss. 4, e55-e63.

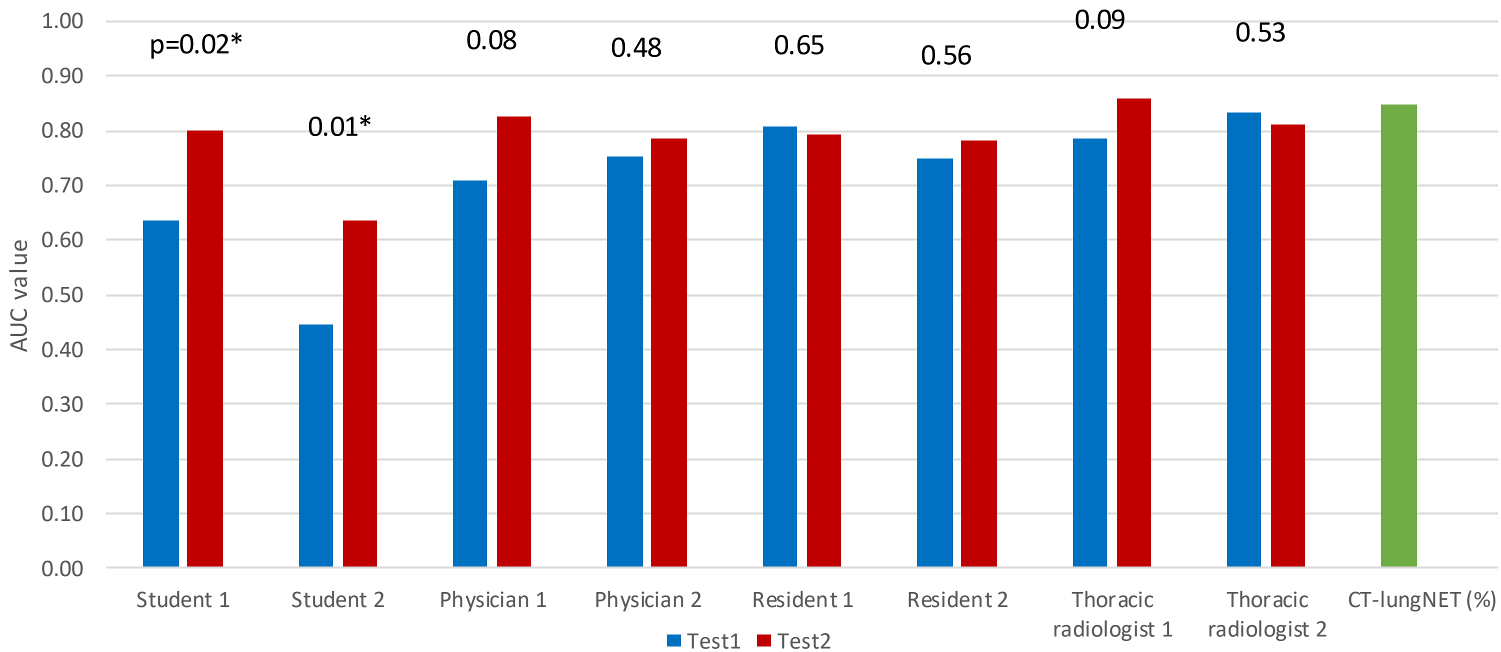
Deep Learning for Computed Tomography

UNIVERSITY OF SASKATCHEWAN

Comparison with AlexNet and Human Reviewers

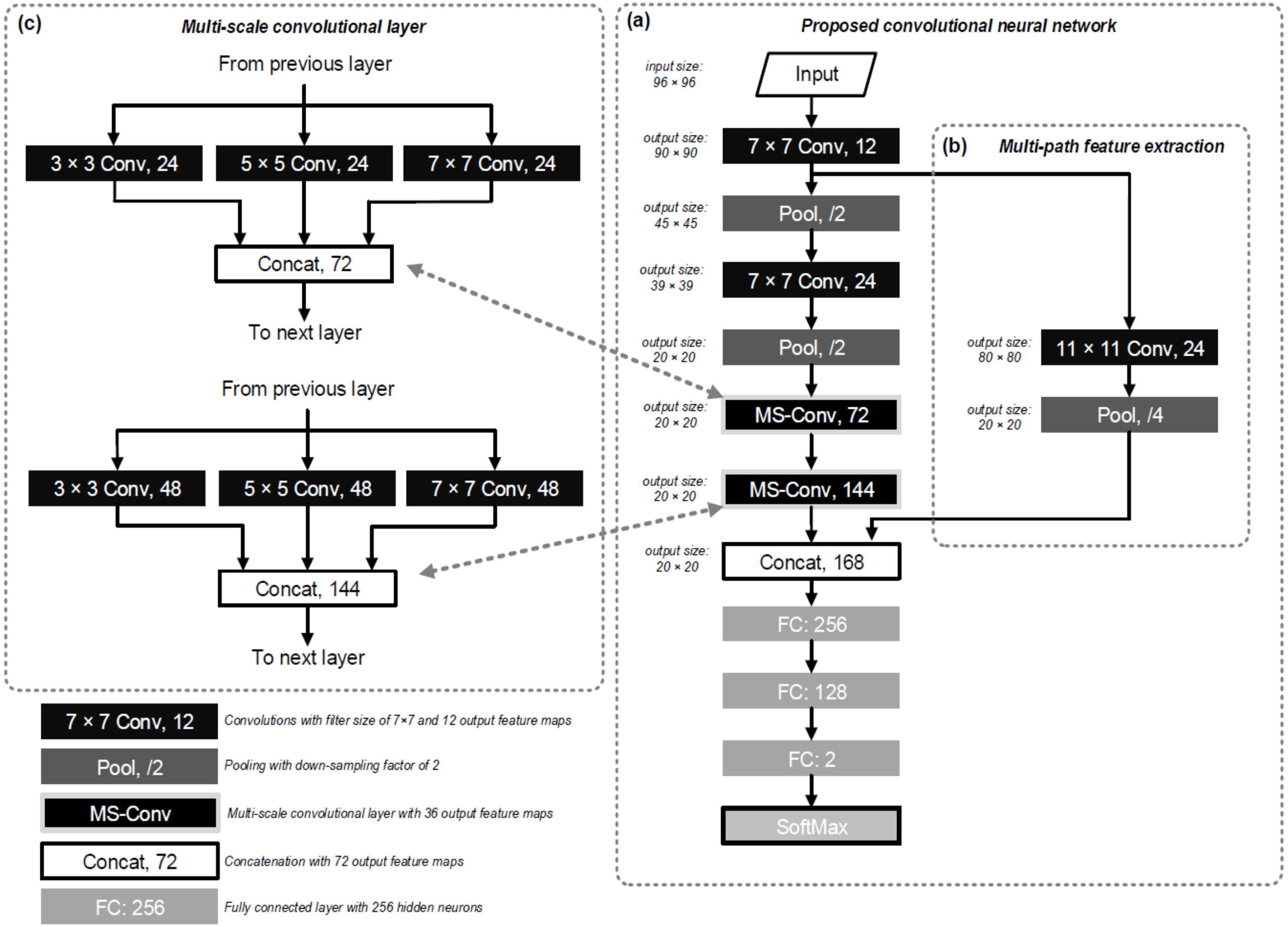
- CTLungNet showed higher AUC value (0.85 vs 0.82) in ROC curves.
- CTLungNet, as a second reviewer, significantly improved performances of radiologists with less experiences (e.g., two students and physicians).





Note: *Test 1, independent observer performance test without aid of CTLungNet; Test 2, observer performance test with reviewing the CTLungNet's malignancy prediction rate.*

Novel Convolutional Neural Network Architecture for Improved Pulmonary Nodule Classification on CT



- Multi-path feature extraction
  - Concatenating feature maps of different levels of layers
  - Combine more robust features with respect to fine global (e.g., contour of nodule) and sparse local (e.g., unique pattern) structures
- Multi-scale convolutional layer
  - Convolving with three filters with different scales
  - Extract more local sparse structures with respect to different sizes

Comparison with Pervious Works

	Method	Accuracy (%)	Sensitivity	Specificity	AUC
Li et al. [1]	CNN with 1 single-scale convolutional layer	80.15	0.789	0.818	0.854
Zhao et al. [2]	CNN with 2 single-scale convolutional layers	84.97	0.843	0.858	0.902
Tajbakhsh et al. [3]	Transfer learning on pre-trained AlexNet	80.58	0.821	0.787	0.855
Shin et al. [4]	Transfer learning on pre-trained GoogLeNet	86.68	**0.906**	0.798	0.933
Shen et al. [5]	Multi-cropped CNN	86.77	0.846	0.895	0.940
Proposed Method	Multi-scale + multi-path	**90.38**	0.887	**0.924**	**0.948**

- All CNNs were trained and evaluated under LUNGx Challenger database [6] with 5 folder cross-validation.

Y. Wang, H. Zhang, K. Chae, Y. Choi, G. Jin, and S. Ko, 2020. "Novel convolutional neural network architecture for improved pulmonary nodule classification on computed tomography," accepted to *Springer Multidimensional Systems and Signal Processing*, Vol. 31, pp. 1163-1183.

References:

[1] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, M. Chen, "Medical image classification with convolutional neural network," *13th International Conference on Control Automation Robotics Vision (ICARCV)*, 2014, pp. 844-848.

[2] X. Zhao, L. Liu, S. Qi, Y. Teng, J. Li, W. Qian, "Agile convolutional neural network for pulmonary nodule classification using CT images," *International Journal of Computer Assisted Radiology and Surgery* 13 (4) (2018) pp. 585-595.

[3] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, J. Liang, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Transactions on Medical Imaging* 35 (5) (2016) pp. 1299-1312.

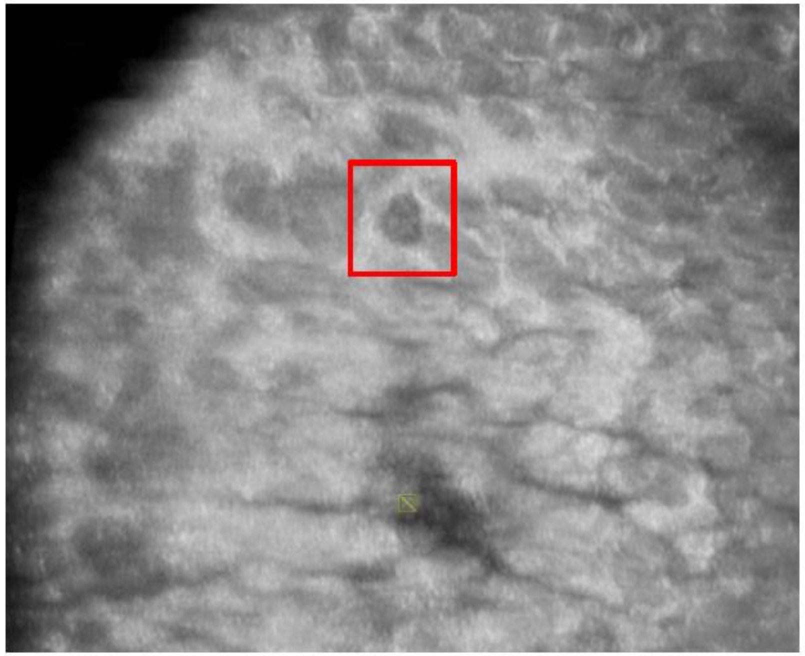
[4] H. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Transactions on Medical Imaging* 35 (5) (2016) pp. 1285-1298.

[5] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, J. Tian, "Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification," *Pattern Recognition* 61 (2017) pp. 663-673.

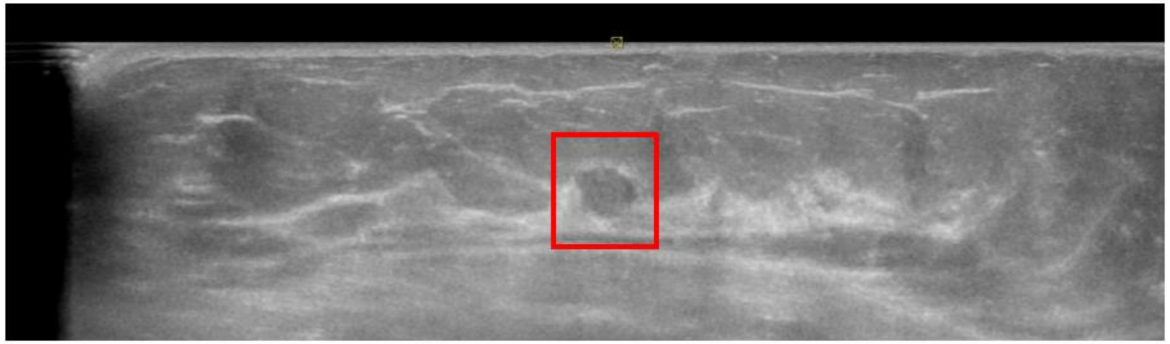
[6] S. G. Armato, K. Drukker, F. Li, L. Hadjiiski, G. D. Tourassi, J. S. Kirby, L. P. Clarke, R. M. Engelmann, M. L. Giger, G. Redmond, K. Farahani, "LUNGx Challenge for computerized lung nodule classification," *Journal of Medical Imaging* 3 (2016) pp. 3-9.

Breast Lesion Classification on Automated Breast Ultrasound

- Automated breast ultrasound (ABUS) is nonradioactive and used as a supplemental screening for breast lesion detection.
  - Compared with mammography, screening ABUS still takes a ***significantly longer time***.



Coronal view



Transverse view

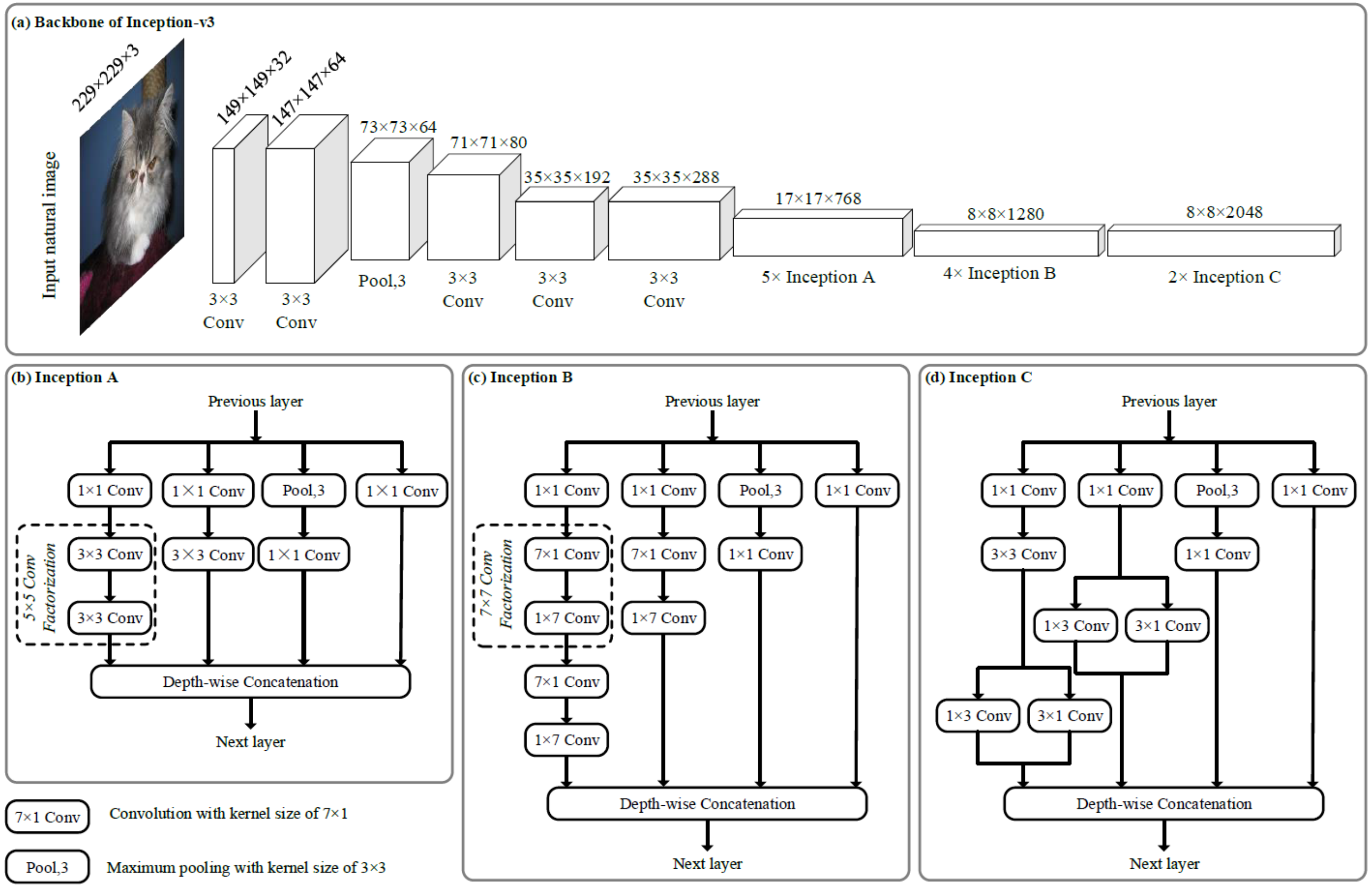
Example of ABUS image for screening in 50-year-old woman. A benign nodule is located in both coronal and transverse views and enclosed by red rectangular boxes.

Y. Wang, E. Choi, H. Zhang, G. Jin and S. Ko, 2020. "Breast cancer classification in automated breast ultrasound using multi-view CNN with transfer learning," *Elsevier Ultrasound in Medicine & Biology*, Vol. 46, Iss. 5, pp. 1119-1132.

Deep Learning for Ultrasound Imaging

UNIVERSITY OF SASKATCHEWAN

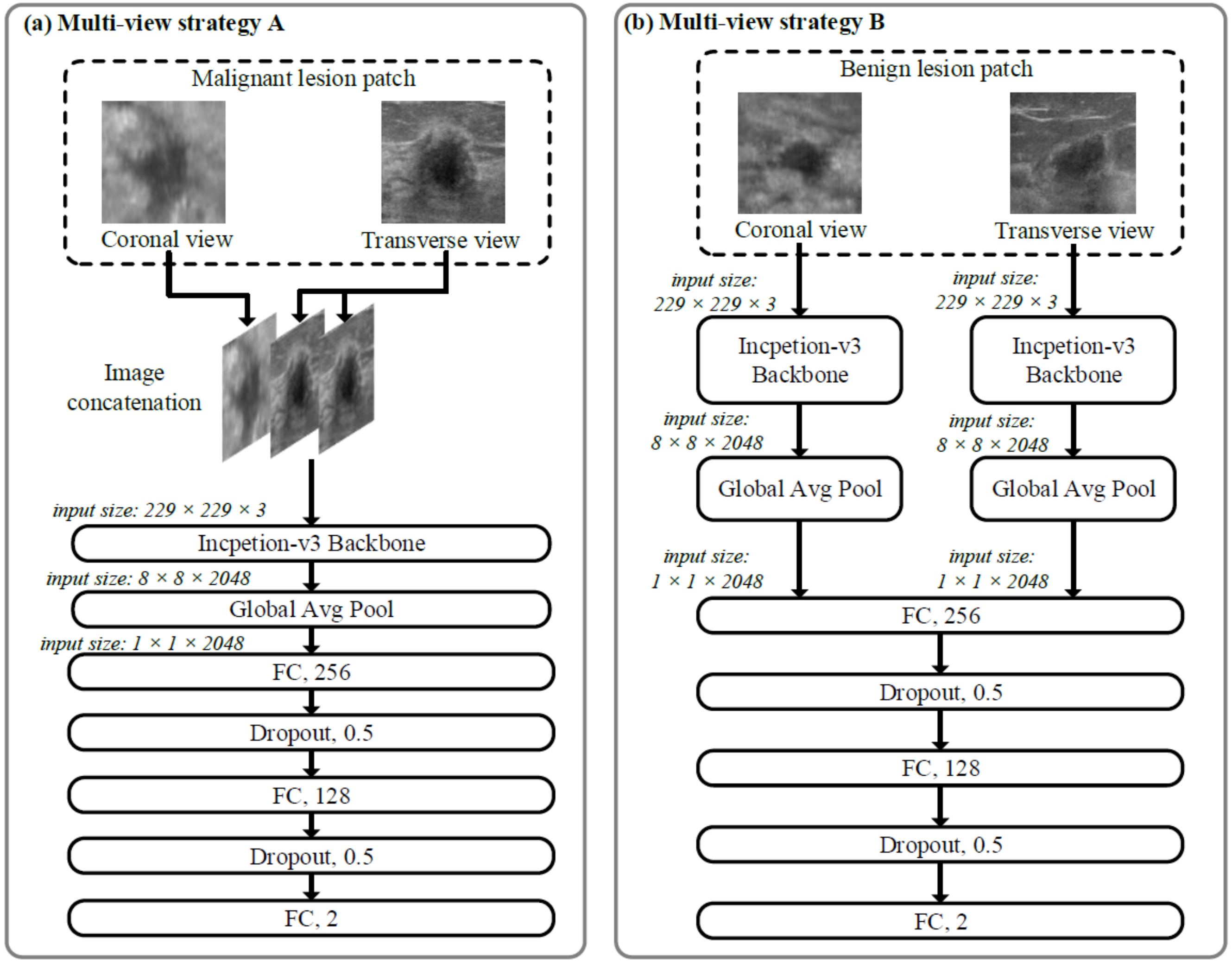
Transfer learning on Inception v-3



- Inception-v3
  - Widely used for natural image classification task
  - Achieved 78.1% accuracy on ImageNet
- Transfer learning on Inception-v3
  - Without design CNN architecture
    - More complex hyper-parameters tuning
    - Trial-and-error to determine number of layers
  - Inherit feature extraction power from natural images to medical imaging

Deep Learning for Ultrasound Imaging

Breast Cancer Classification in ABUS using Multi-view CNN with Transfer Learning



- Multi-view strategy A
  - Concatenating different views of lesion patches
  - Use single Inception-v3 for feature extraction
- Multi-view strategy B
  - Dual paths feature extractions
  - Concatenating extracted features by Fully-connected layer

UNIVERSITY OF SASKATCHEWAN

Comparison with other CNN architectures

Method		Sensitivity	Specificity	AUC
View strategy	**Backbone**			
Coronal view	Inception-v3	0.831	0.800	0.891
Transverse view	Inception-v3	0.843	0.844	0.953
Multi-view strategy A	ResNet-50	0.809	0.830	0.928
	Inception-ResNet-v2	0.870	0.875	0.976
	Inception-v3	**0.885**	**0.889**	**0.982**
Multi-view strategy B	Inception-v3	0.853	0.867	0.959

CNN-based Carries Identification on Micro-CT Image

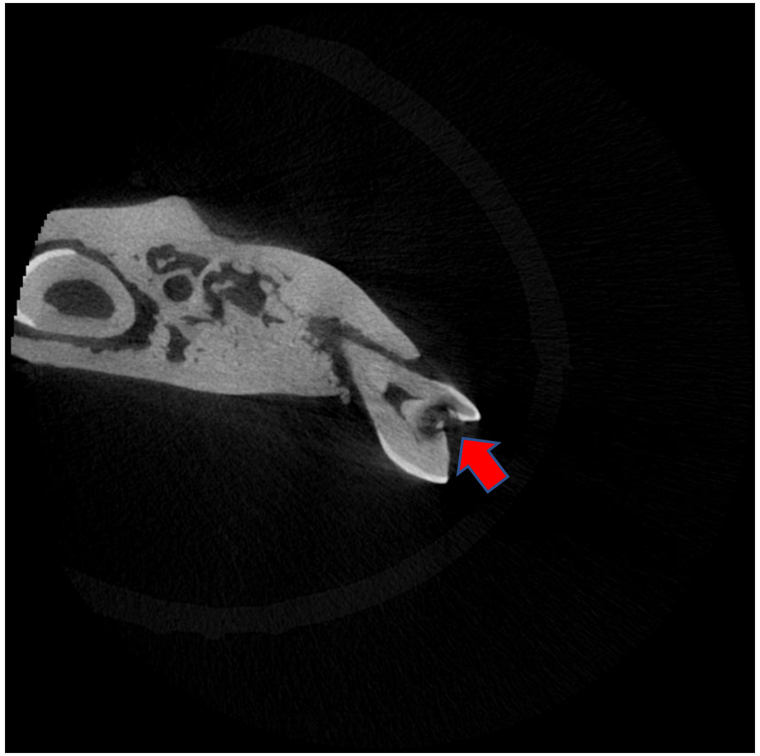


Illustration of Micro-CT slice image. Red arrow indicates a caries.

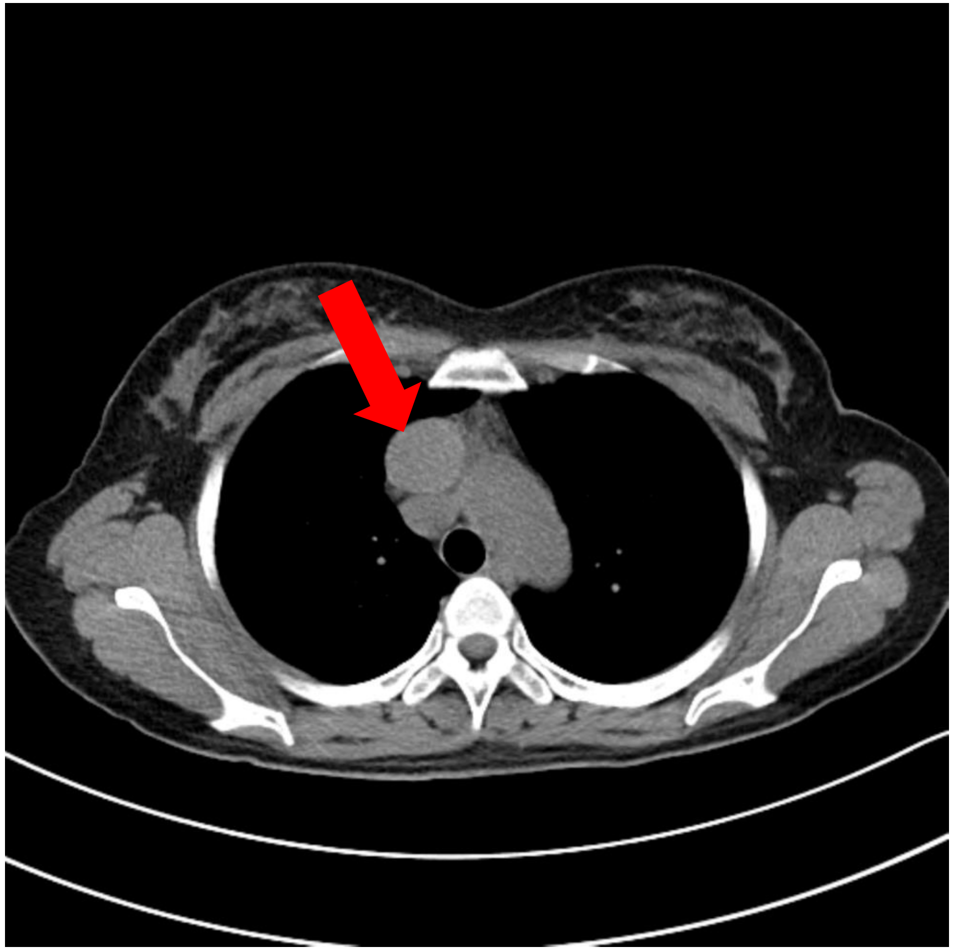
- Micro-CT is a most advanced modality used by dentists to diagnose carries.
  - Provides a 3D view of the entire teeth.
  - Improved carries detection rate compared to X-ray.
- **Problem**: screening on Micro-CT is still a time-consuming time due to each Micro-CT Scan produces hundreds of slice images.
- **Solution**: A CNN-based CADx system is developed to classify slice images of a Micro-CT scan into normal (Without carries) or abnormal (Contains carries).

Preliminary Results

- One Micro-CT cases was collected from department of health science at University of Saskatchewan:
  - Consists of 590 micro-CT slice image.
  - Two carries were identified in 40 of the 590 slice images.
- Transfer learning is applied to train three different CNN architectures:
  - Inceptionv3
  - DenseNet
  - ResNet
- ResNet achieved the best classification performance compared to other two.

	Sensitivity	**Specificity**
ResNet backbone	80.00% (8/10)	88.98% (105/118)

Anterior Mediastinal Lesion

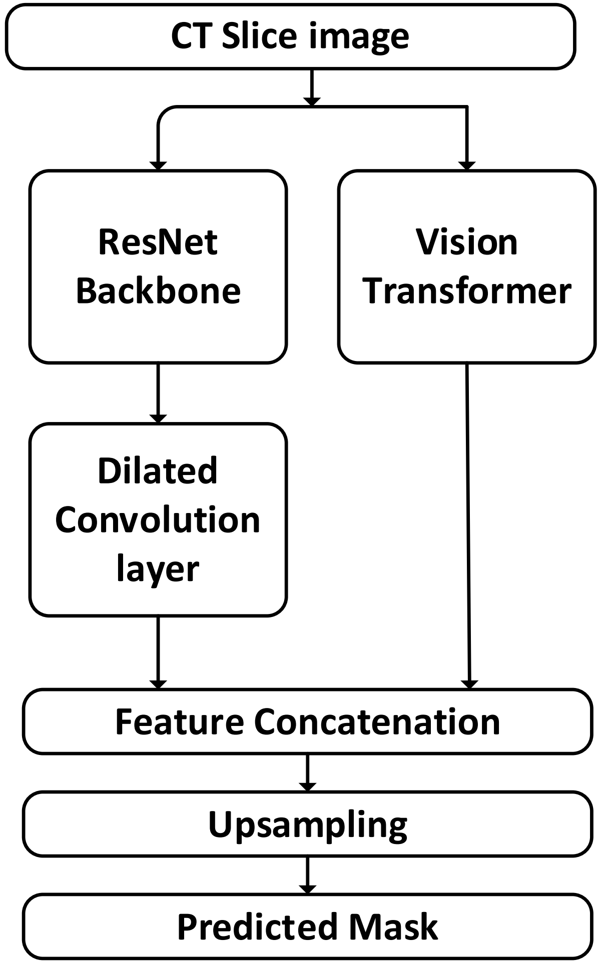


- A common disease in the chest.
- CT is widely used in the diagnosis of mediastinal diseases.
- **Problem**: difficult to distinguish lesions in CT images because of image artifact, intensity inhomogeneity, and their similarity with other tissues.
- **Solution**: segmented lesion can provide radiologists a method to better subtract the features of the lesions, thereby improving the accuracy of diagnosis.

Deep Learning on Anterior Mediastinal Lesion Diagnosis

UNIVERSITY OF SASKATCHEWAN

Novel CNN for Mediastinal Lesion Segmentation



- A total of 185 CT scans was collected from JBNU hospital.
- A CNN-based architecture is developed to segment lesion from CT imaging.
  - Multi-feature learning via ResNet and vision transformer.

- Preliminary Results:

	Dice coefficient	Sensitivity	Specificity
3D-ResUNet [1]	87.73%	N/A	N/A
Proposed Model	85.41%	0.854	0.8523

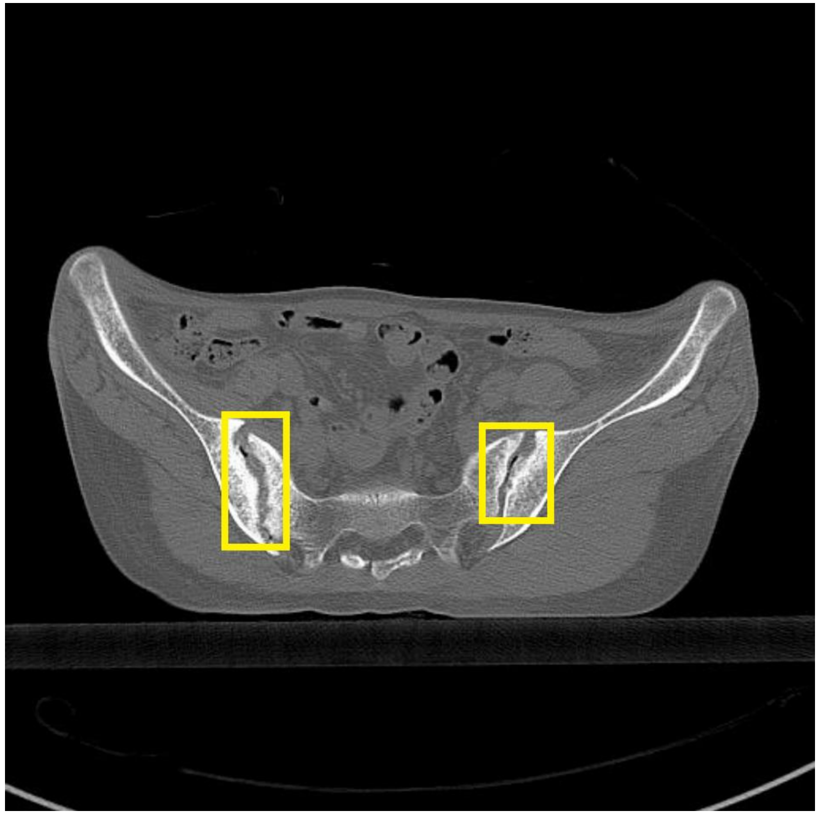
[1] Huang, S., Han, X., Fan, J., Chen, J., Du, L., Gao, W., Liu, B., Chen, Y., Liu, X., Wang, Y., Ai, D., Ma, G., & Yang, J. (2021). Anterior Mediastinal Lesion Segmentation Based on Two-Stage 3D ResUNet With Attention Gates and Lung Segmentation. *Frontiers in Oncology, 10.*

What is Ankylosing Spondylitis (AS)?

- AS is an arthritis that affects the spine, and sacroiliac (SI) joints in the pelvic region.
- Not well known outside MSK radiologists and has an average diagnostic delay of **7-10** years.
  - Continued patient suffering
  - Strain on healthcare system, 12k (adjusted for inflation, 2006 to 2022) per year [6]
- Visualized in the sacroiliac joints (SIJs).
- Would be useful to have a diagnostic model.

Disease Progression:
Modified from https://www.faceyourbackpain.com/ankylosing-spondylitis

From an abdominal/pelvic CT scan (left), can erosion be differentiated from a control patient?



Erosion: Early AS Symptom

Young Control Patient

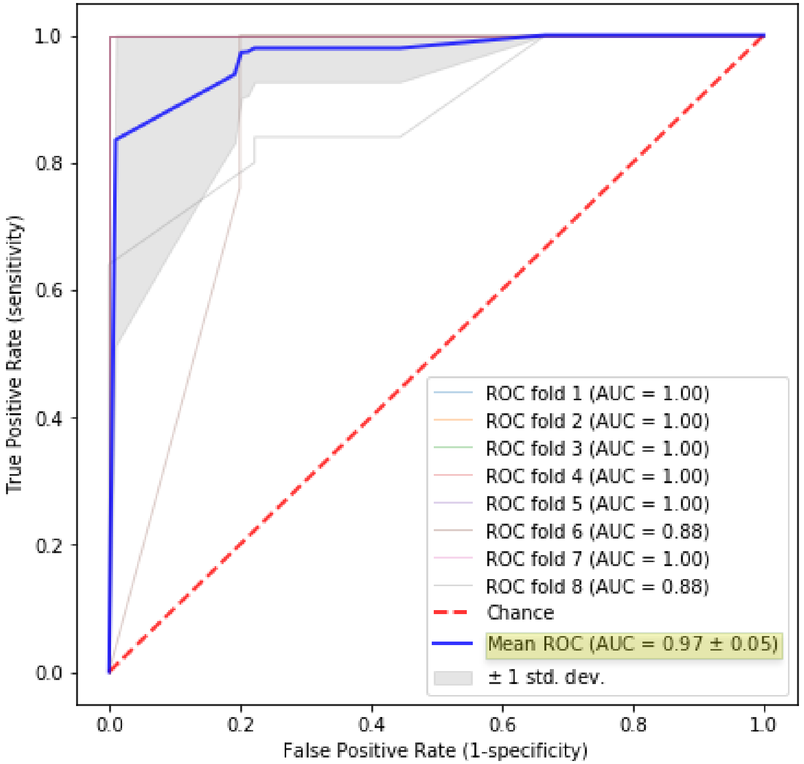
Old Control Patient

Machine and Deep Learning for Early Detection of Ankylosing Spondylitis (AS) using CT

UNIVERSITY OF SASKATCHEWAN

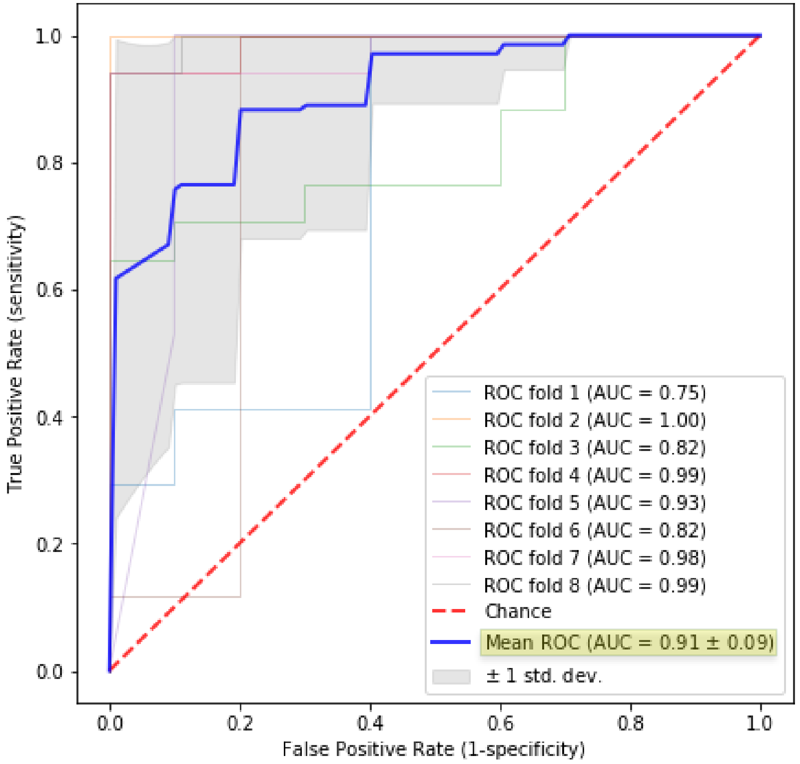
Differentiation and therefore diagnosis is possible with machine learning!

The ROC of Random Forest (10 decision trees) machine learning classifiers trained on GLCM and LBP texture features (8 features in total, 4 from each), 8-fold cross validation:



Erosion vs. Young Control

96.0% accuracy, 92.9% sensitivity

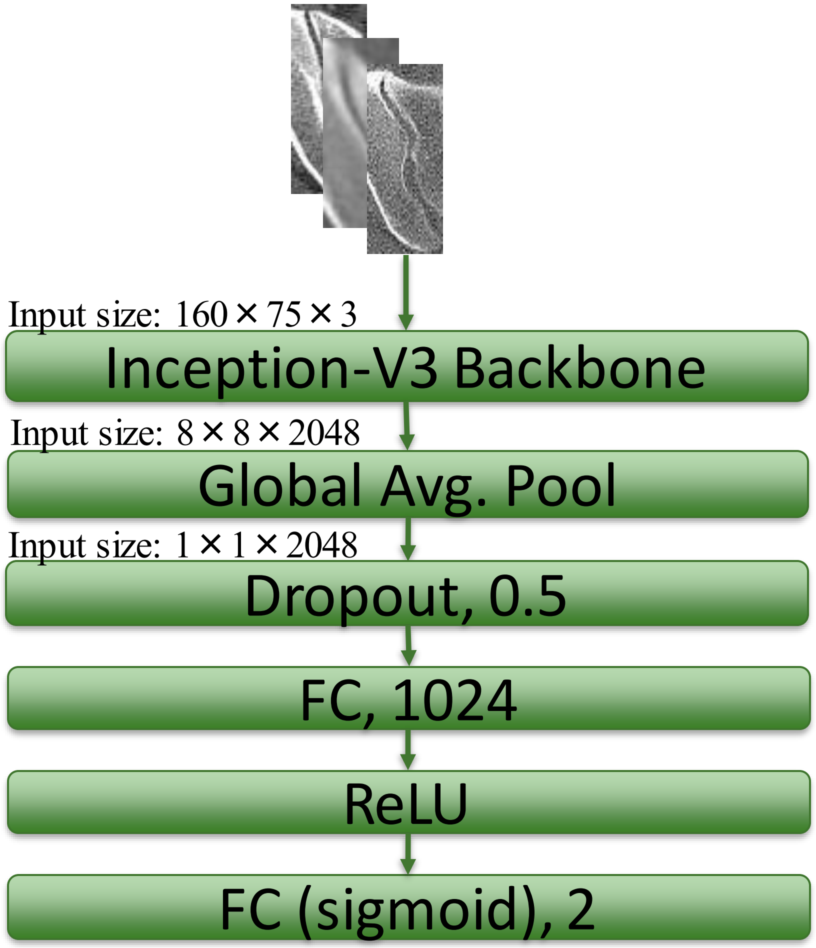


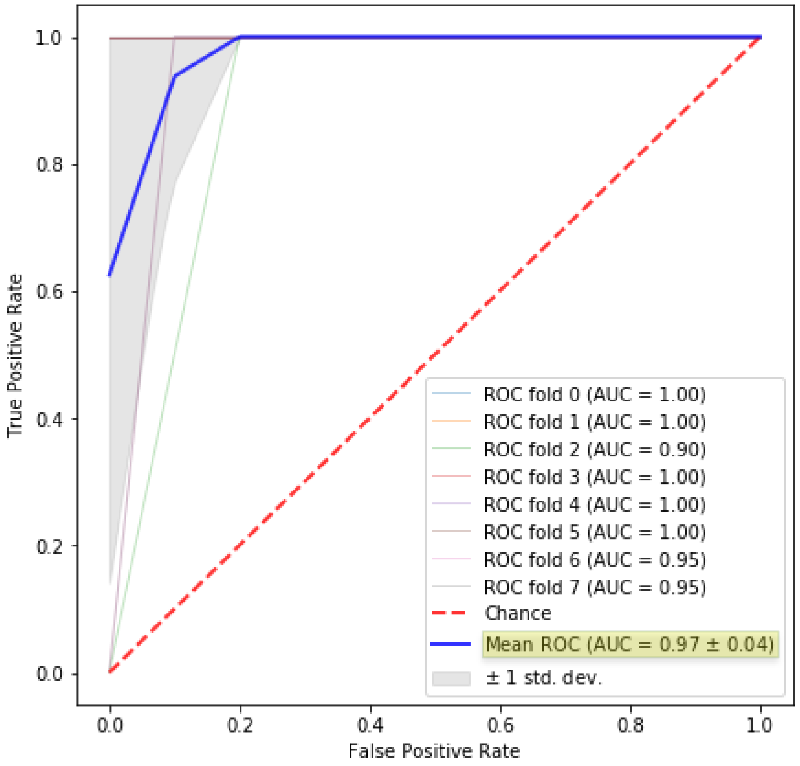
Erosion vs. Old Control

82.4% accuracy, 80.6% sensitivity

Differentiation and therefore diagnosis is even better with deep learning!

The ROC of a deep learning classifier trained via transfer learning on Inception-V3, 8-fold cross validation:





Erosion vs. ALL Control

99.0% accuracy, 97.5% sensitivity

Comparison of AS detection against similar work

([1,2] use a dataset comprised of MR imagery)

COMPARISON OF OUR STATISTICAL MACHINE LEARNING WORK WITH SIMILAR WORK.

Algorithm	ROC AUC	Sensitivity	Features Per SI Joint	SI Joint Regions
Our Work: Texture Features				
k-NN	0.90	93.9%	8	281
Random Forest	0.97	96.0%	8	281
From [2]: Texture and Spectral Features				
k-NN 1	0.91	91%	5	612
k-NN 2	0.96	77%	17	612

Castro-Zunti, R., Park, E., Younhee, C., Jin, G., and Ko, S. "Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, considering patient age." *Computerized Medical Imaging and Graphics.* 82. (2020). p. 101718

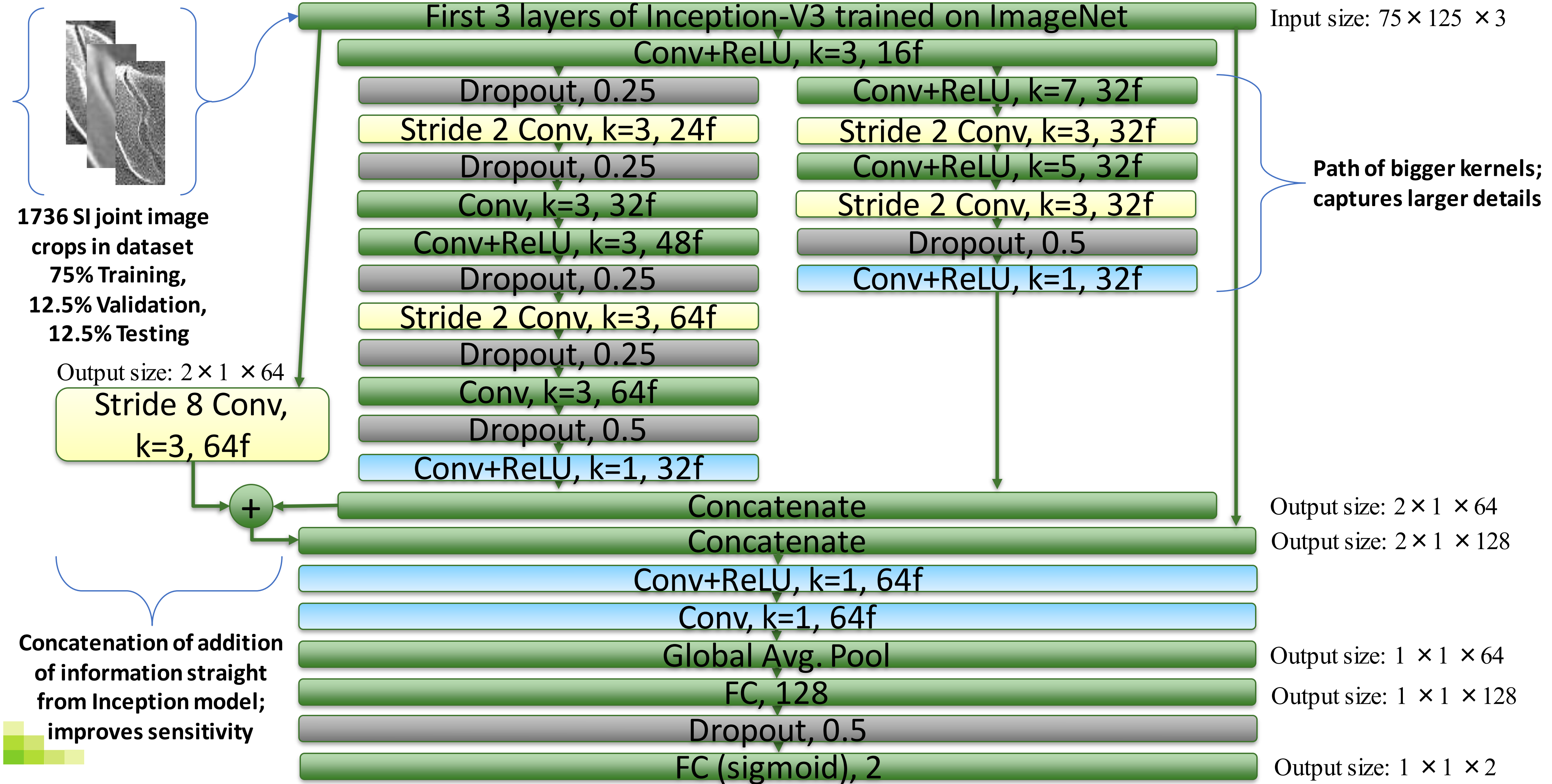
COMPARISON OF OUR NEURAL NETWORK WORK WITH SIMILAR WORK.

Algorithm	ROC AUC	Sensitivity	Parameters	Features Per SI Joint	SI Joint Regions
Our Work: CNN-acquired Features					
Modified InceptionV3	0.97	97.5%	23.9×10^6	N/A	681
From [1]: Texture Features					
MLP ANN	0.93	73%	Not mentioned	39	612
From [2]: Texture and Spectral Features					
MLP ANN	0.95	73%	Not mentioned	13	612
From [3]: CNN-acquired Features					
5-layer CNN + Random Forest Ensemble	0.97	95%	$> 1.9\times10^6$	N/A	34894

Machine and Deep Learning for Early Detection of Ankylosing Spondylitis (AS) using CT

University of Saskatchewan

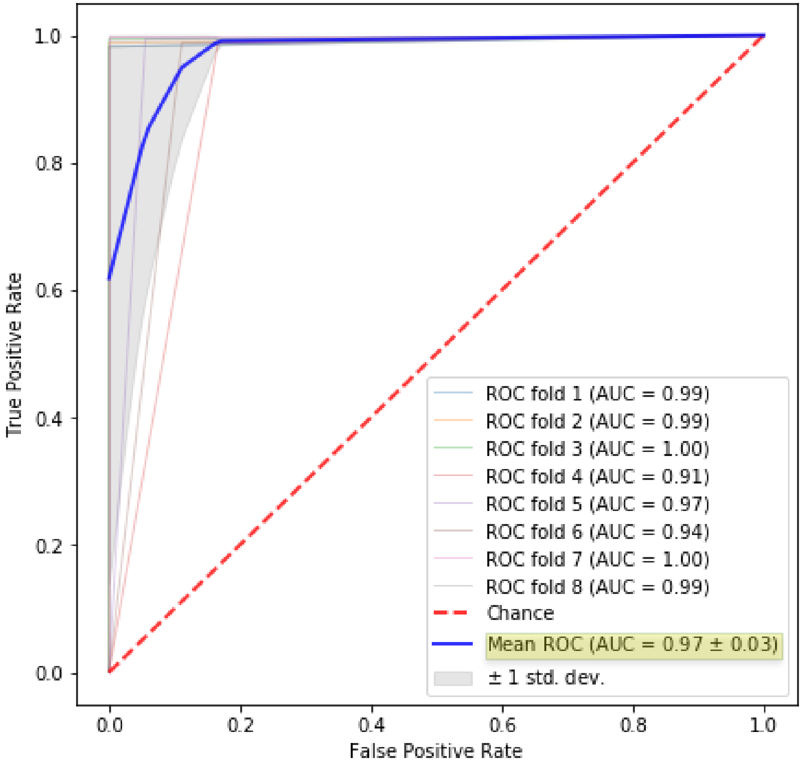
What about a custom deep learning architecture to lower computational cost?



UNIVERSITY OF SASKATCHEWAN

What about a custom deep learning architecture to lower computational cost? It doesn't do that bad against fine-tuning on a vanilla InceptionV3 model despite using 93.7% less parameters.

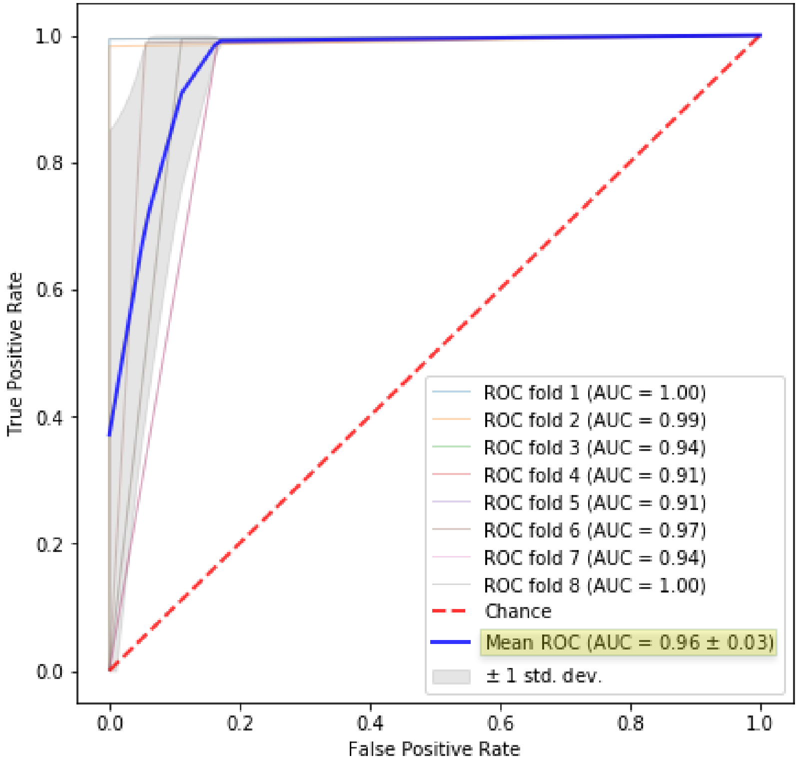
Our Custom Model



Erosion vs. ALL Control

98.7% accuracy, 97.4% sensitivity

Fine-Tuned InceptionV3



Erosion vs. ALL Control

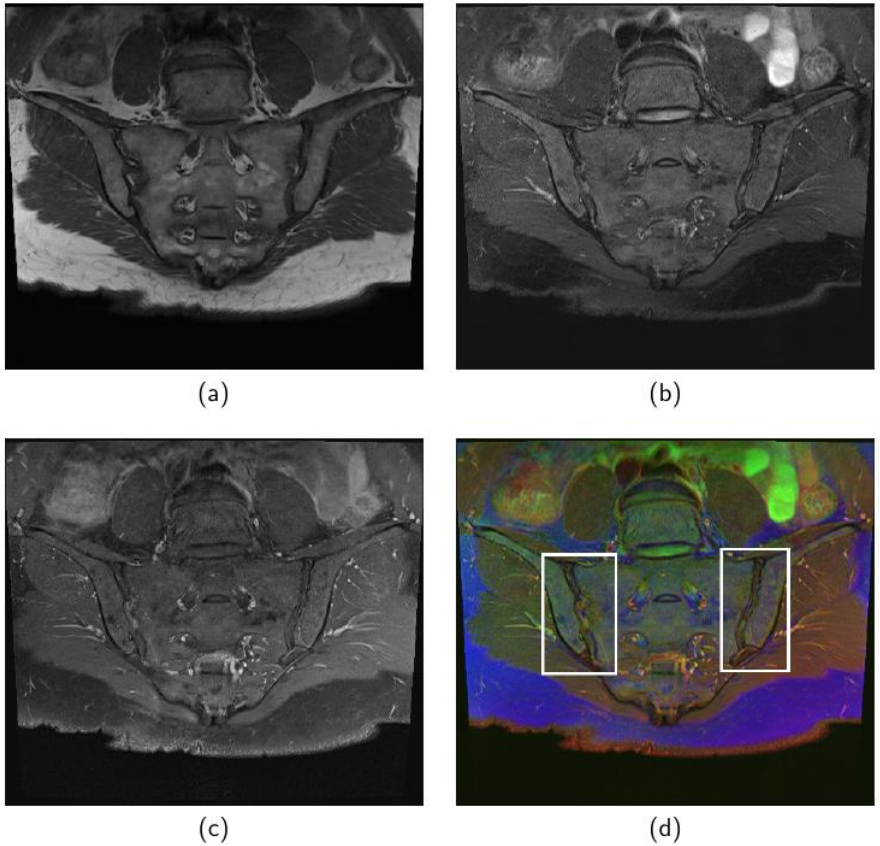
98.4% accuracy, 95.7% sensitivity

Ankylosing Spondylitis Classification from MR

UNIVERSITY OF SASKATCHEWAN

AS Diagnostic System

- AS diagnostics commonly done using MR because CT has radiation, which is unsafe for patients [7].
- MR can also image the very earliest symptoms (inflammation) before radiographic changes (i.e. erosion) [8].
- However, CT tends to outperform MR for AS diagnoses [7] due to the inflammation also being prevalent in control populations [9].
- Sought to develop a general AS diagnostics system using conventional MR as an input. Showed that combining 3 types (T1wTSE, T2wFS, and T2wPCFS) as RGB is better than any single type alone for classification.



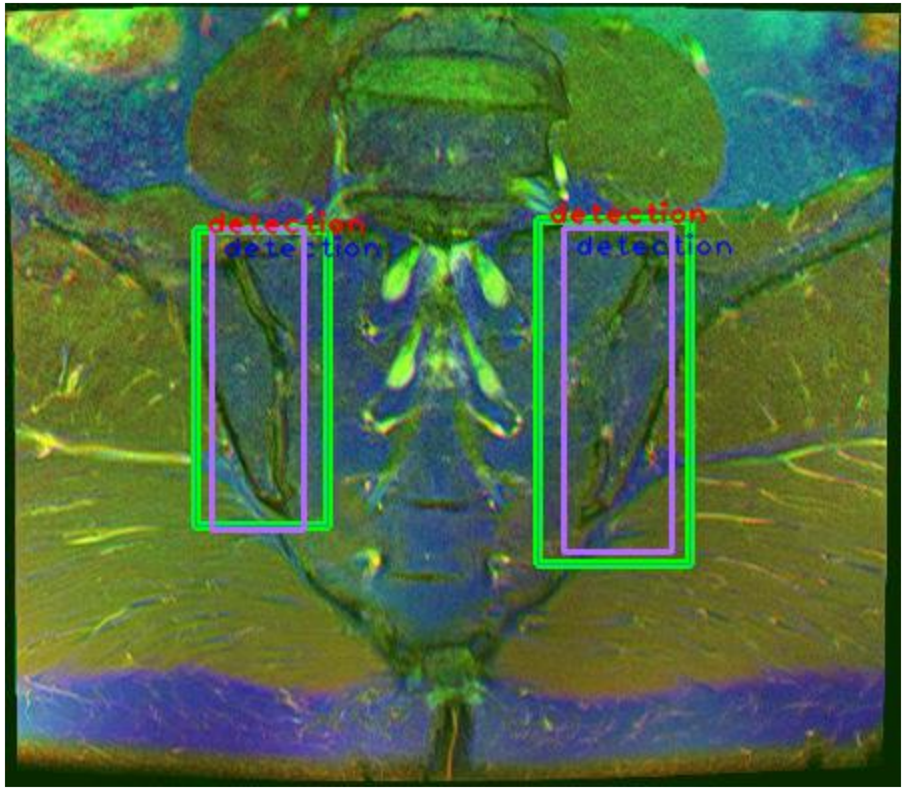
Different MR imaging sequences (a-c) and their combined (RGB) image (d). SIJs' bounding boxes are shown in white.

Ankylosing Spondylitis Classification from MR

UNIVERSITY OF SASKATCHEWAN

SIJ Detection

- We fine-tuned a YOLOv5 Medium object detection network using the combined imagery.
- 835 radiologist-annotated frames from 38 patients (18 AS and 20 control).
- Per-patient-based (i.e. the patients in the train, validate, and test folds were disjoint) 6-fold cross validation F1-score of 99.9% and mAP@0.5 of 99.5%.

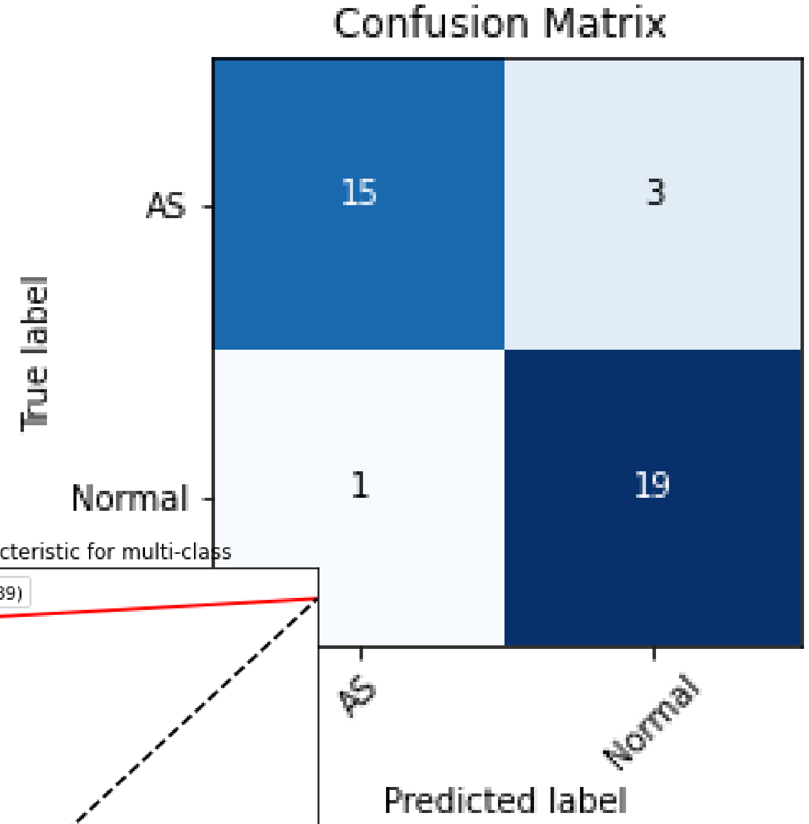


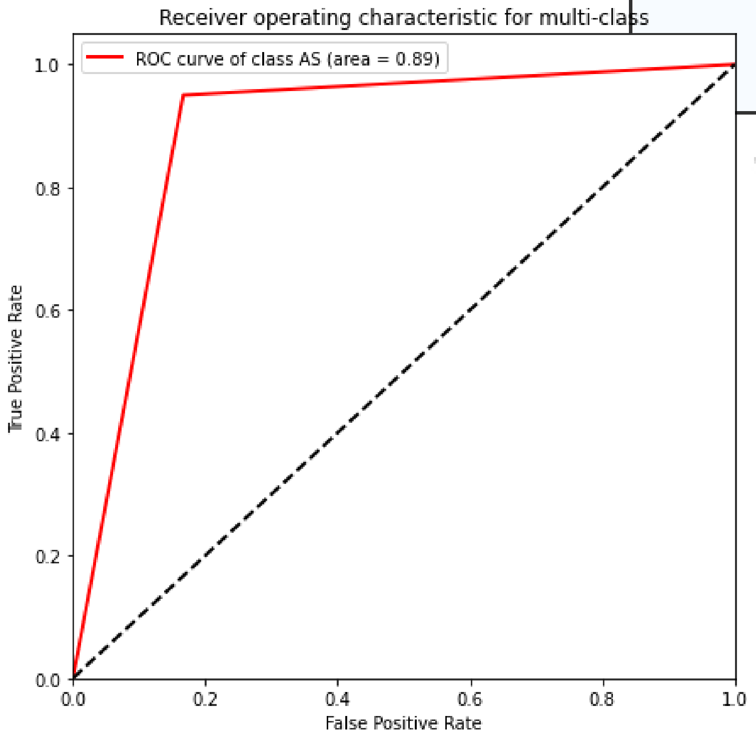
Ankylosing Spondylitis Classification from MR

UNIVERSITY OF SASKATCHEWAN

SIJ AS vs. Normal Classification

- We fine-tuned a variety of InceptionV3, ResNet50, and VGG16-based convolutional neural networks (CNNs).
- Same dataset as for detection. 531 AS and 606 control SI joints.
- Best was an ensemble of VGG16 networks trained with global average pooling.
- Per-patient Accuracy of 89.5%, Specificity of 95%, Sensitivity of 83.3%, ROC AUC of 89.2%.
  - Specificity especially interesting considering issue with inflammation in control patients.





UNIVERSITY OF SASKATCHEWAN

SIJ AS vs. Normal Classification: Comparisons

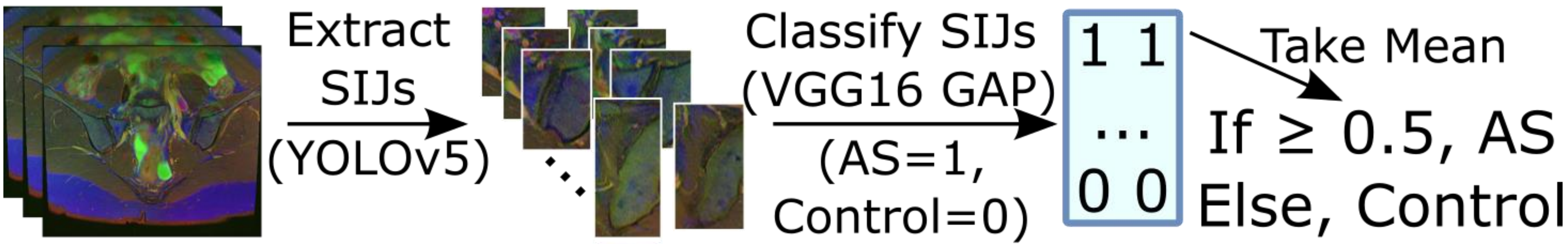
- Outperformed a 10-yr experienced radiologist by 13.13% accuracy, 11.11% sensitivity, 15.00% specificity, and 13.06% ROC AUC. (Not statistically significant though.)
- Additionally, reimplemented related work using the dataset and showed ours to be superior:

	Acc. (%)	Sens. (%)	Spec. (%)	AUC (%)	Params.
VGG16GAP Ensemble (Proposed)	**89.47**	**83.33**	**95.00**	**89.17**	14.7M× 2
MLP with Hand-crafted Features [5]	73.68	72.22	75.00	73.61	**37.1k**
5-layer CNN [3]	71.05	61.11	80.00	70.56	3.00M
Our retrained CT AS classifier model	78.95	**83.33**	75.00	79.17	23.9M× 2

  – Meets or exceeds truth-related metrics.
  – Likely faster than the MLP if a GPU is used for the CNN, considering the MLP requires generating 158 features, which would take non-negligible time.

UNIVERSITY OF SASKATCHEWAN

Full System Architecture



Ankylosing Spondylitis Diagnostic Systems

UNIVERSITY OF SASKATCHEWAN

Future Work

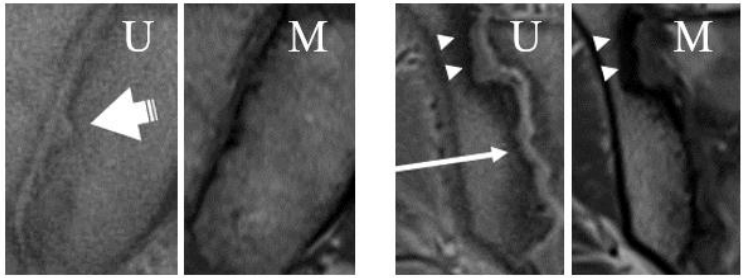


Fig. 1. *UTE (U) and conventional T1W MRI (M) SIJ crops for early (left) and advanced (right) AS. Arrows indicate AS symptoms.*

- Adapting our network and methodology to develop a system that works with Ultrashort TE (UTE) MR imagery, shown to be effective in early AS diagnosis [4] (Fig. 1.)
- Developing a system able to extract SI joint regions of interest (RoIs) from a patient CT scan without manual input to bound the locations of RoIs, similar to what was done for MR; this would help develop a fully automatic erosion or AS detection system requiring only a patient CT/UTE scan video as input.

Some parts from: Castro-Zunti, R., Park, E., Younhee, C., Jin, G., and Ko, S. "Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, considering patient age." *Computerized Medical Imaging and Graphics.* 82. (2020). p. 101718

References:

[1] Faleiros, M. C. *et al.*, "Computer-aided classification of inflammatory sacroiliitis in magnetic resonance imaging," in *31st International Congress and Exhibition on Computer-Assisted Radiology and Surgery*, vol. 12, Barcelona, Spain, 2017, pp. S154–S155.

[2] Faleiros, M. C. *et al.*, "Pattern recognition of inflammatory sacroiliitis in magnetic resonance imaging," in *European Congress on Computational Methods in Applied Sciences and Engineering*, 2018, pp. 640–644.

[3] Shenkman, Y. *et al.*, "Automatic Detection and Diagnosis of Sacroiliitis in CT Scans as Incidental Findings," in *Medical Image Analysis*, vol. 57, 2019, pp. 165–175.

[4] Hahn, S. *et al*. "Can Bone Erosion in Axial Spondyloarthropathy be Detected by Ultrashort Echo Time Imaging? A Comparison With Computed Tomography in the Sacroiliac Joint," in *Journal of Magnetic Resonance Imaging*, 2022.

[5] Faleiros, M. C. *et al.*, "Machine Learning Techniques for Computer-Aided Classification of Active Inflammatory Sacroiliitis in Magnetic Resonance Imaging," in Advances in Rheumatology (London, England), vol. 60, no. 1, 2020, p. 25.

[6] Kobelt, G. *et al*. "Costs and Quality of Life of Patients with Ankylosing Spondylitis in Canada," in *Journal of Rheumatology*, vol. 33, no. 2, 2006, pp. 289–295.

[7] Diekhoff, T. *et al*. "Choose Wisely: Imaging for Diagnosis of Axial Spondyloarthritis," in *Annals of the Rheumatic Diseases*, vol. 81, no. 2, 2021, pp. 237–242.

[8] ØStergaard, M. *et. al*. "Imaging in Ankylosing Spondylitis," in *Therapeutic Advances in Musculoskeletal Disease*, vol. 4, no. 4, 2012, pp. 301–311.

[9] Baumbach, S. F. *et al*. "How We Manage Bone Marrow Edema—An Interdisciplinary Approach," in *Journal of Clinical Medicine*, vol. 9, no. 2, 2020, p. 551.

Rib Tracking and Fracture Classification

UNIVERSITY OF SASKATCHEWAN

Why is this important?

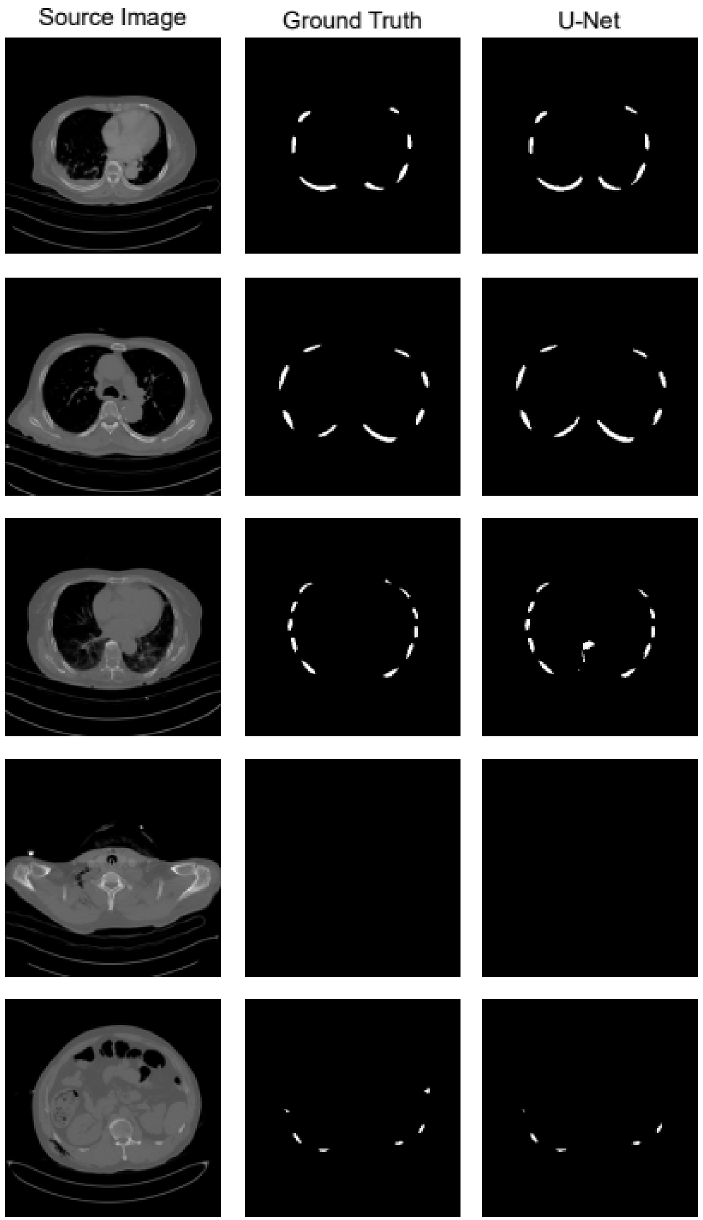
- Ribs are the most common bones to be fractured.
- >66% of patients admitted to trauma centres for chest trauma have a rib fracture
- ~12 of patients admitted with a rib fracture will die as a result of their injuries
  - 10% for young adults
  - 22% for the elderly
- Although it is relatively easy for a radiologist to track and determine which of a human's 12 ribs have fractures, looking through 200+ CT scan frames can be tedious for a radiologist.
  - Would be useful to streamline this.

Modified from https://shopeverydaymedical.com/blogs/home/broken-ribs-braces-to-treat-injured-or-broken-ribs

Rib Tracking and Fracture Classification

Rib Segmentation

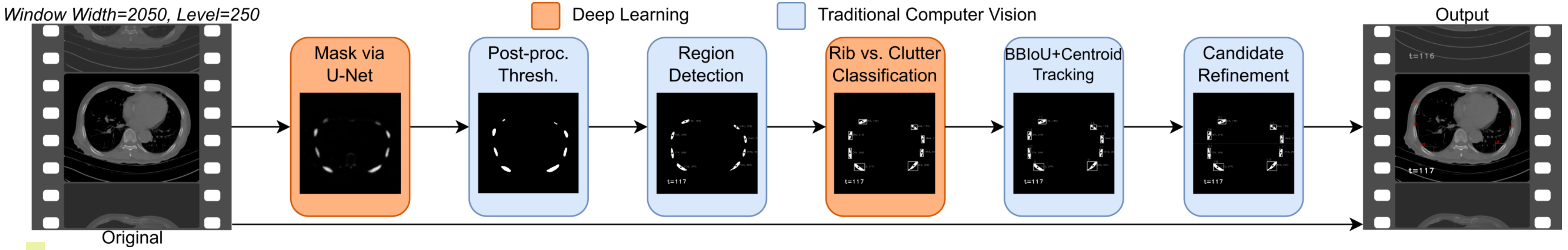
- Via U-Net. Experimented with both single-class (rib only) and multiclass (rib and clutter).
- Experimentation with various thresholding levels and morphological operations when finalizing output probability masks.



Rib Tracking and Fracture Classification

Rib Tracking

- Inputs are RoIs found from the segmentation step.
- Custom computer vision software and multi-object tracking system.
  - Close to 2000 lines of Python code!
- IoU- and centroid-based tracking.
- Many applied heuristics pertinent to the way ribs move through a scan frame.
- Current best tracking model (with multiclass segmentation) achieves 45% correct accuracy and 69% mostly correct (ribs 3-8 correctly tracked) over 98 patients.

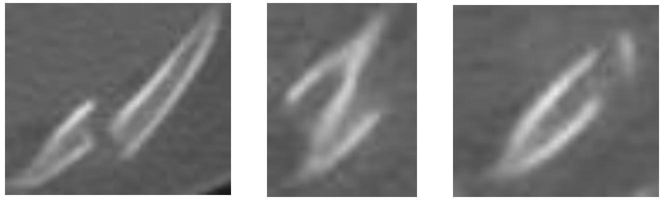


Rib Tracking and Fracture Classification

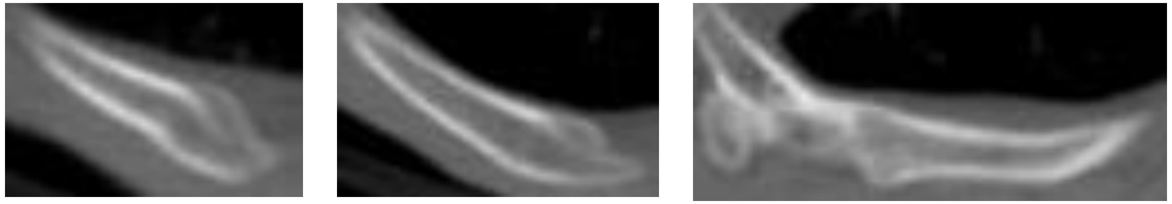
UNIVERSITY OF SASKATCHEWAN

Dataset

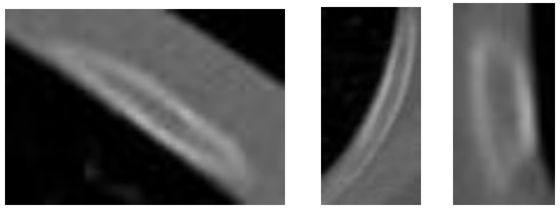
Acute (Recent) Fracture: 2,546 RoIs



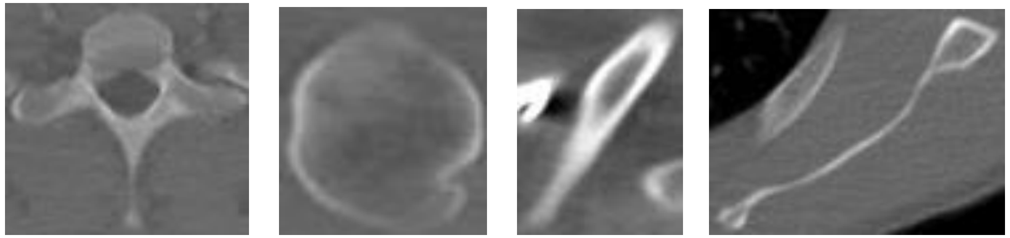
Old (Healed) Fracture: 5,983 RoIs



Regular (No Fracture): 18,630 RoIs



Various Anatomical Clutter: 7,122 RoIs



Rib Tracking and Fracture Classification

Rib Fracture Classification

- We fine-tuned a variety of InceptionV3, ResNet50, and VGG16-based CNNs with 5-fold cross validation. Best was first 7 blocks of InceptionV3.

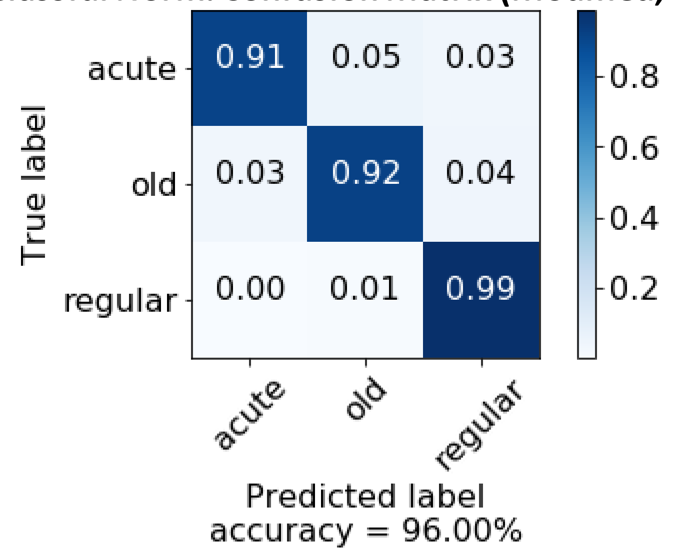
Classful Results

	Acc. (%)	Mac. Sens. (%)	Params. $\times 10^6$	Per-Crop Time (ms) CPU	Per-Crop Time (ms) GPU
InceptionV3					
Modified	96.00	94.0	6.83	13.6	12.2
Full	95.07	92.6	21.81	22.7	21.0

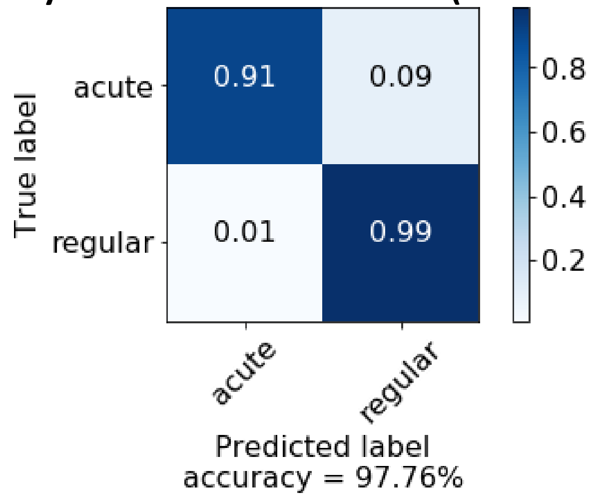
Binary Results

	Acc. (%)	Mac. Sens. (%)	AUC (%)	Params. $\times 10^6$	Per-Crop Time (ms) CPU	Per-Crop Time (ms) GPU
InceptionV3						
Modified	97.76	94.6	94.7	6.83	13.4	12.4
Full	97.21	92.2	92.1	21.81	22.4	20.5

Classful Norm. Confusion Matrix (Modified, Ours)



Binary Norm. Confusion Matrix (Modified, Ours)

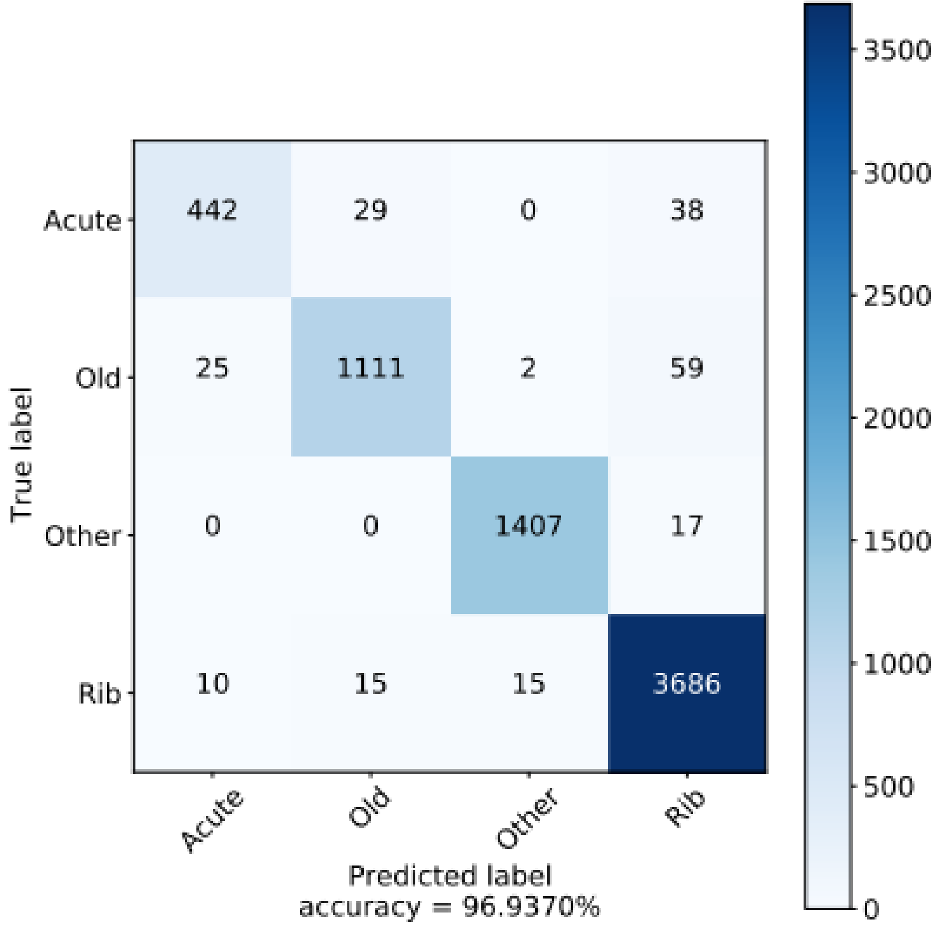


Rib Tracking and Fracture Classification

UNIVERSITY OF SASKATCHEWAN

Combined Rib vs. Anatomical Clutter and Rib Fracture Classification

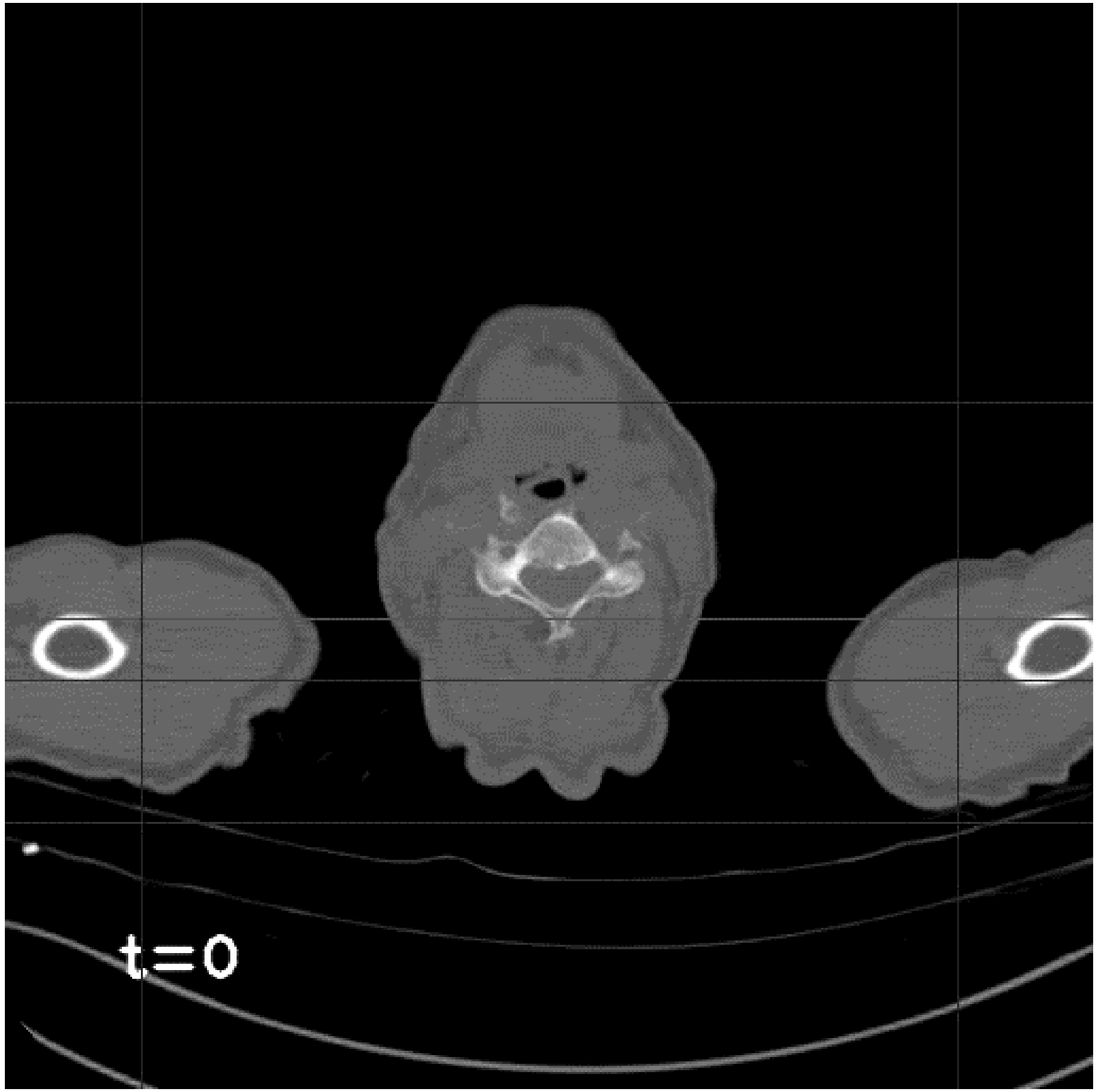
- We fine-tuned a variety of InceptionV3, ResNet50, and VGG16-based CNNs using 6-fold cross validation.
- Best model was first 15 blocks ResNet50.
  - Accuracy of 96.6% and macro sensitivity of 94.8%.
  - When ran as binary (rib vs. other) Accuracy of 99.5%, sensitivity of 99.7%, specificity of 98.8%, and ROC AUC of 99.2% over 5-fold cross-validation test set.
- 31% completely correct fracture classifications and 76% partially correct fracture classifications over 98 patients.



(a)

Rib Tracking and Fracture Classification

Example: (Red bounding box is predicted acute, blue is old, green is regular, grey is non-rib)



Rib Tracking and Fracture Classification

Current Status of Research

- **Rib Tracking/Labelling:**
  - A brief paper about the clinical aspects of the work to date is being prepared by partnering radiologists
- **Rib Fracture Classification:**
  - A paper about the computer vision / deep learning rib fracture classification system has been published.
    - R. Castro-Zunti, K. Chae, Y. Choi, G. Jin and S. Ko, “Assessing the Speed-Accuracy Trade-Offs of Popular Convolutional Neural Networks for Single-Crop Rib Fracture Classification,” *Elsevier Computerized Medical Imaging and Graphics*, vol. 91, 2021, p. 101937.

Rib Tracking and Fracture Classification

UNIVERSITY OF SASKATCHEWAN

Future Projects

- Improve segmentation component.
  - Tendency for false negative predictions especially in the smaller, latter-stage ribs.
- Improve tracking component.
  - Could develop additional heuristics or try to make some parts of the program simpler / more streamlined.
  - Could also possibly try training a deep learning- (rather than computer vision-) based tracker? Previously thought to have lack of labelled data, but one could take the fully correct labelled patient scans, extract bounding boxes, and train.
    - Would have ~50 patients.

Rib Tracking and Fracture Classification

Future Projects

- Improve fracture classification component.
  - Could employ more sophisticated methods/models, have 2 models (one for fracture classification and one for rib vs. clutter) rather than 1.
  - Could design a heuristic/threshold for deciding fractures.
    - E.g., rib X detected as acute 1 time vs. as 4 times?
  - Additionally, tendency for false positive old samples in areas of rib region creation (bottom) and destruction (top), and false positives on early- and later-sequence ribs.
    - Probably best served using some selection algorithm for what rib RoIs get predicted by classifier.

Gout Classification from DECT Imagery

University of Saskatchewan

What is the problem?

- Gout (tophi) is a uric acid accumulation that can lead to massive swelling and inflammation.
  - Can deteriorate joints, cartilage, and bone.
- Dual-energy CT (DECT), introduced ~2010, was shown to easily image tophi as green patched.
- Though at first believed to be revolutionary, later found that some green regions of interest (RoIs) can appear even in normal/healthy patient DECT scans.
- Would be useful to have a diagnostic model to overcome challenges associated with these false positive RoIs.

From https://www.bestpodiatristnyc.com/what-is-gout-and-foot-surgery-for-gout/

Gout Classification from DECT Imagery

UNIVERSITY OF SASKATCHEWAN

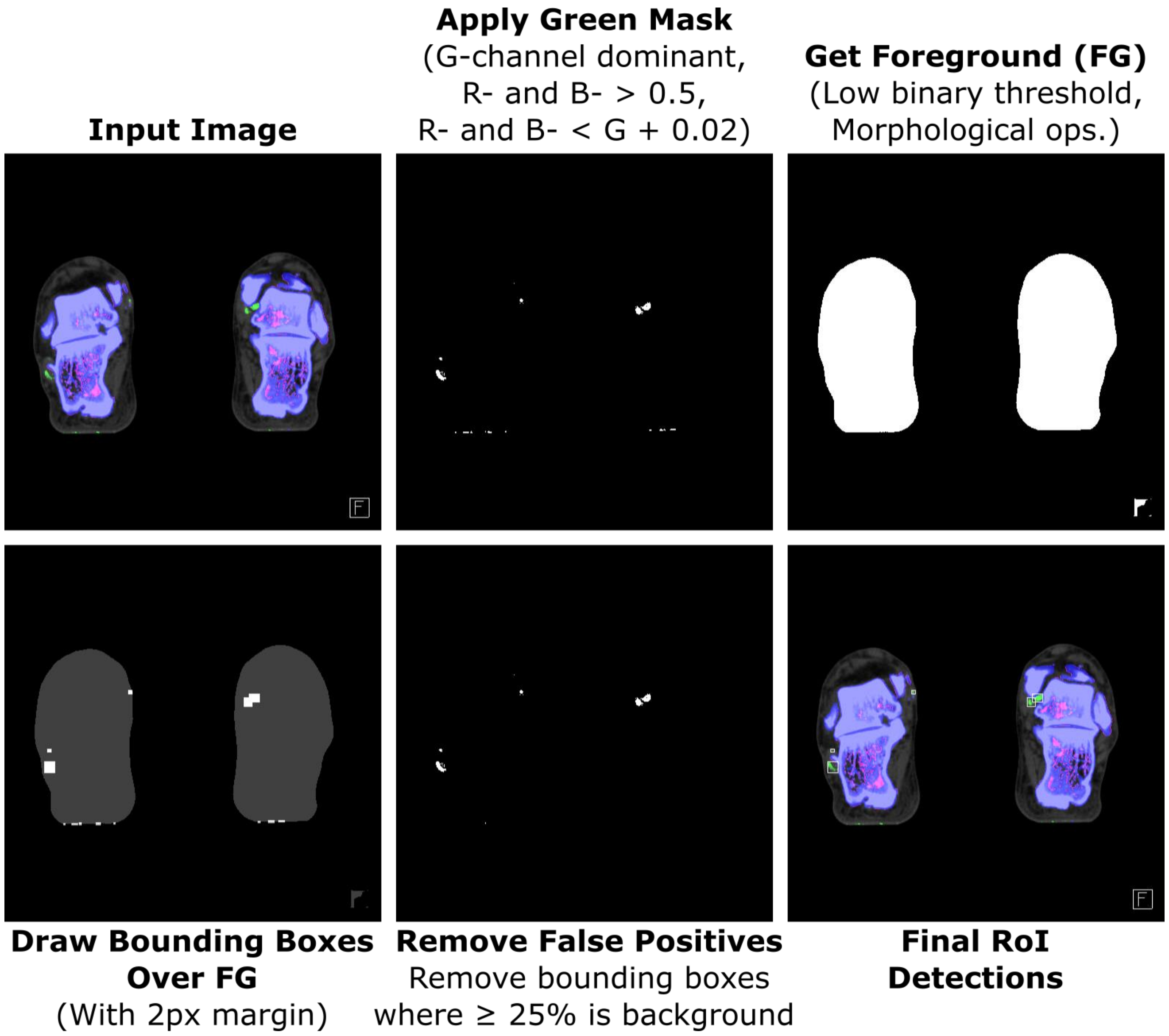
Dataset

- 47 gout patients and 27 control patients.
- Extracted 10,912 gout RoIs and 7,792 control RoIs.
- First DECT gout image dataset to be used for machine learning purposes.
  - Earlier work used structured patient records, not images [1].
    - [1] Bahra, G., Wiese, L. (2018). Classifying Leukemia and Gout Patients with Neural Networks. In: Database and Expert Systems Applications. DEXA 2018. *Communications in Computer and Information Science*, vol 903. Springer, Cham.

Gout Classification from DECT Imagery

Tophi/Artifact Detection

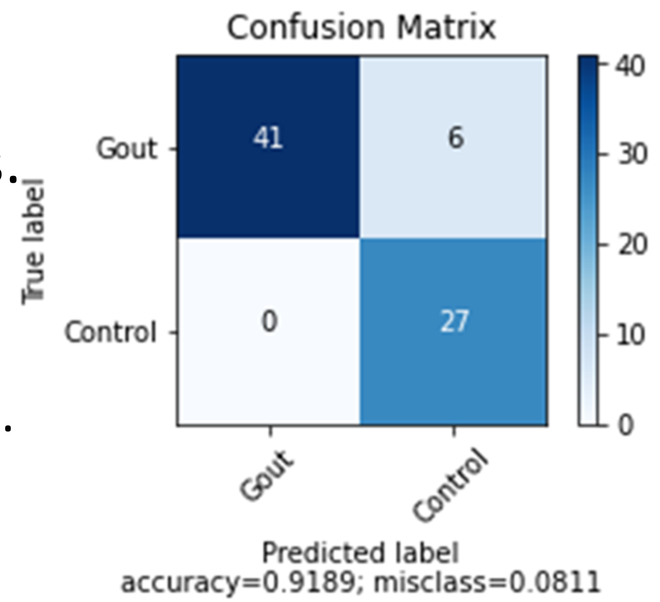
- Because tophi/artifacts appear green in DECT, designed a computer vision algorithm to crop (with a small padding) regions of green.
- Ignored false positives at the edge of the appendage via a technique that looked at the % background in a RoI's bounding box.
- No clear method to remove sidebar if present.
  – Presumed that a radiologist could do this manually. Possible computer vision future work project.

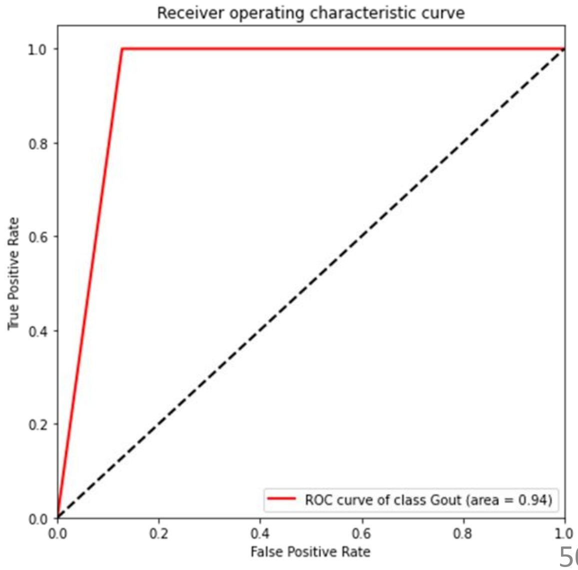


Gout Classification from DECT Imagery

Tophi vs. Artifact Classification

- We fine-tuned InceptionV3, ResNet50, and VGG16-based CNNs.
- Because size of RoIs is variable, trained 3 models per architecture:
  - Small, medium, and large, based off area-size quartiles over the dataset.
- Per-patient dataset split based on large area-size quartile (because large had fewest cropped samples). Highest 6-fold cross-validation validation accuracy/AUC for each area-size dataset was VGG16 trained with global average pooling.
- Trained traditional ML classifiers using general patient features.
  - 7 features: # boxes found, % boxes in each area-size dataset, % RoIs predicted as gout in each area-size dataset.
- Best was linear SVM. Per-patient cross-validation test results:
  - Accuracy 91.89%, sensitivity 87.23%, specificity 100%, ROC AUC 93.62%.

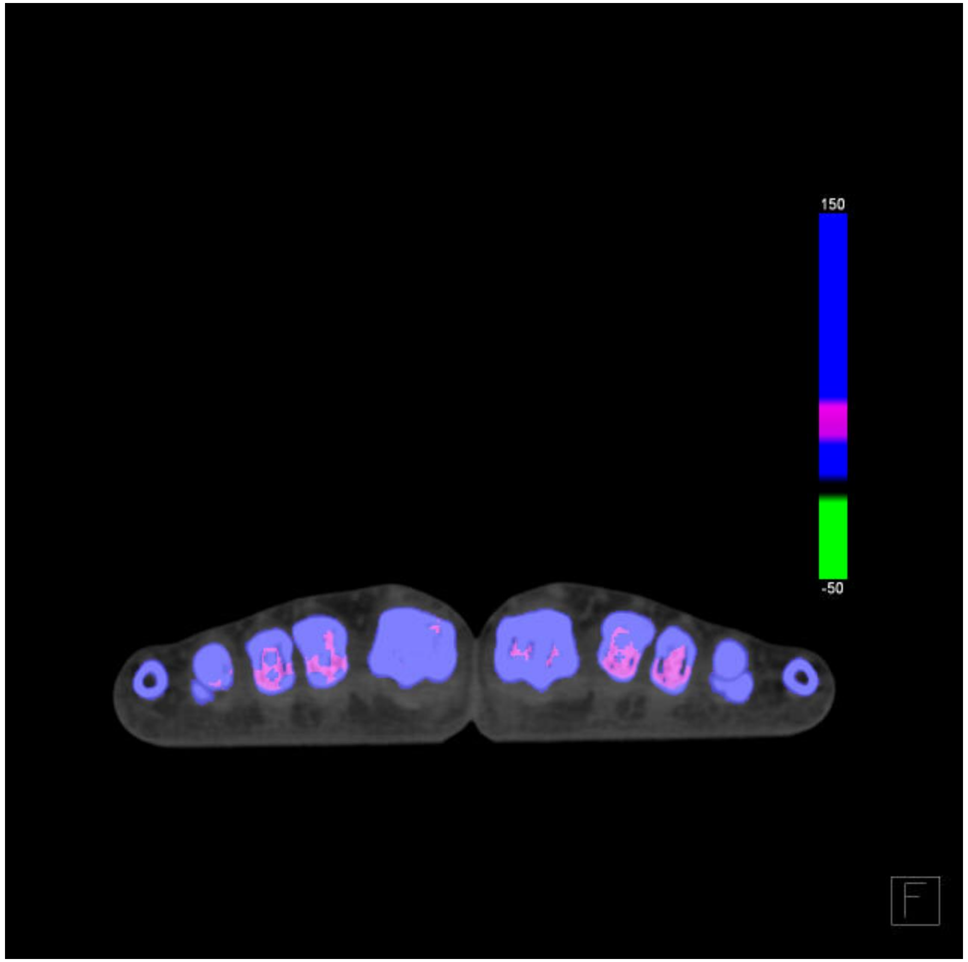




Gout Classification from DECT Imagery

Future Work

- Improve RoI detection.
  - Seems to already work decently, but one could try a deep learning approach, or to develop a component to automatically remove sidebars when present (right image).
- Improve classification CNNs or final per-patient prediction algorithm.
  - Could employ more sophisticated methods/models, investigate an ensemble of methods/models, etc.
- W. Yoo, E. Park, D. Lee, R. Casto-Zunti, S. Ko and Y. Choi, “Solving the Final Puzzle of Gout Detection in DECT via Machine Learning-Based Mitigation of Pseudolesion-Related Challenges: Enhancing Diagnostic Accuracy” accepted to *ACR Convergence 2023*

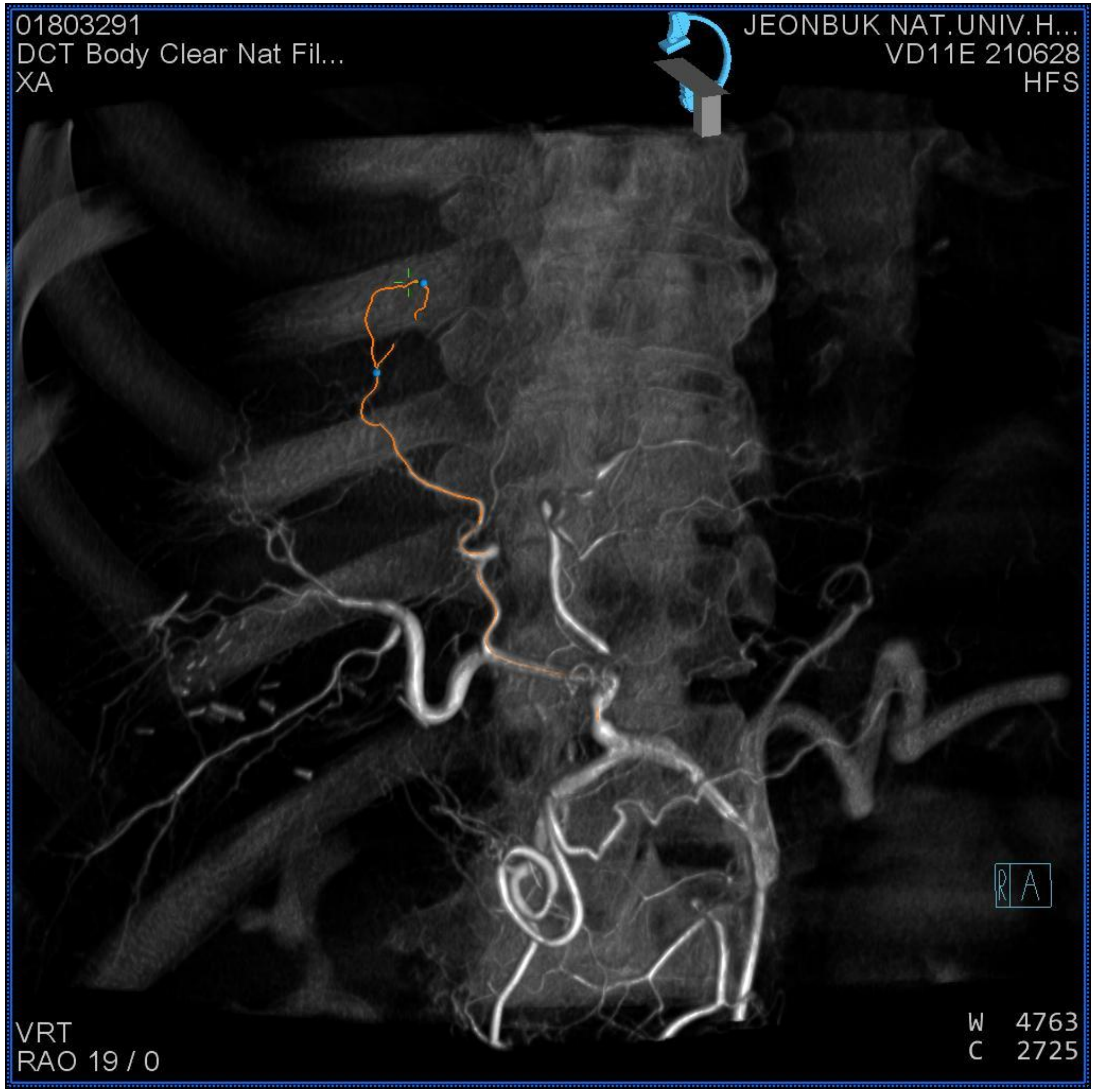


Optimal Viewpoint for Angiography

UNIVERSITY OF SASKATCHEWAN

Project Overview

- Angiography refers to blood (etc.) vessels. There are techniques to capture, segment (in orange), and develop 3D reconstructions of vessels (right).
- However, when conducting surgery, a surgeon must locate an optimal viewpoint wherein the visibility of the vessel of interest is maximized.
  - Time consuming, especially where there is potentially critical surgery to be performed!
- A program that could process all images across all the system's rotations and angles and automatically find the maximized visibility would be invaluable.

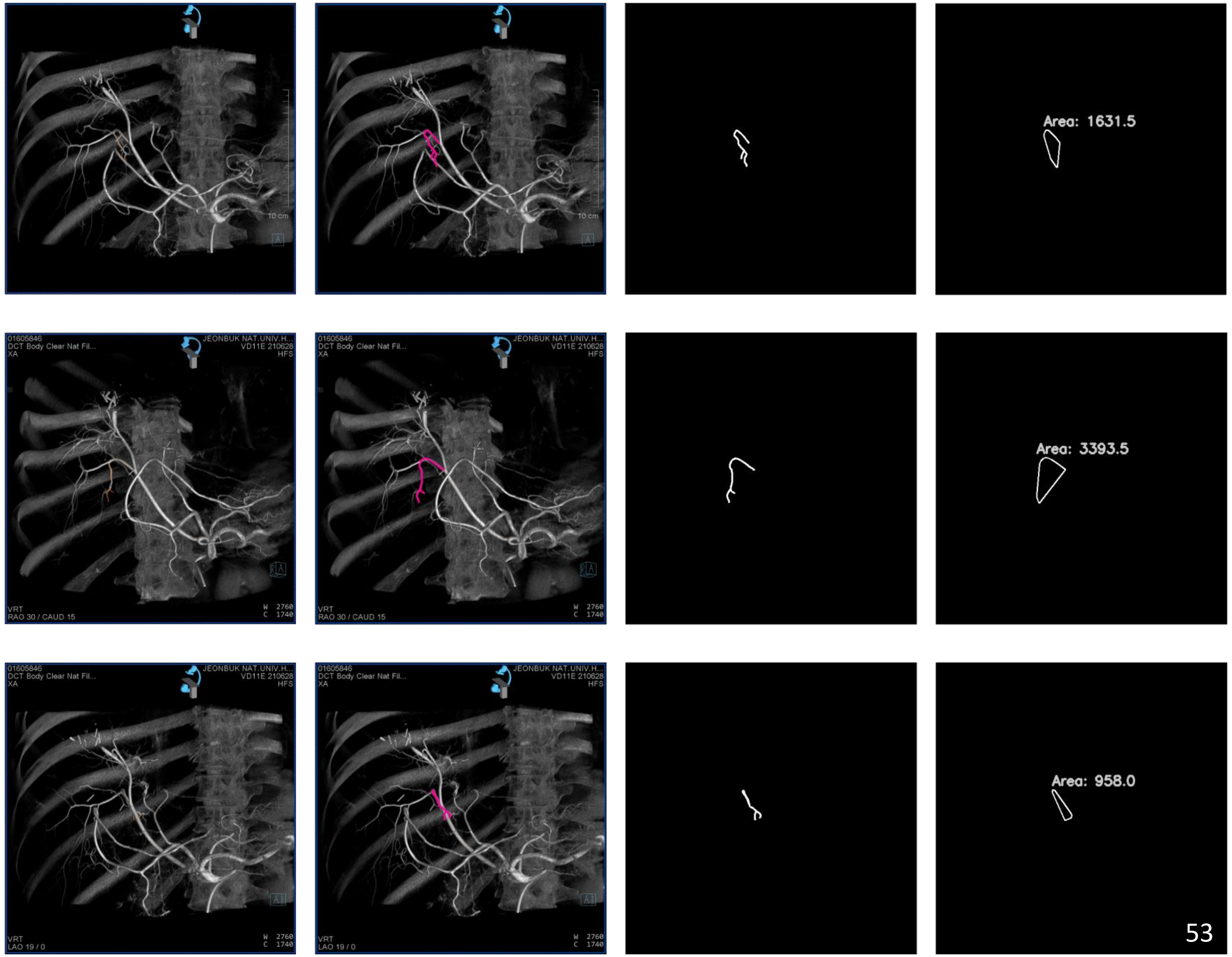


Optimal Viewpoint for Angiography

Methodology

- We are presently experimenting with functions based on the convex hulls of the segmented lines.
  - Traditional computer vision.
- Research is still relatively preliminary.

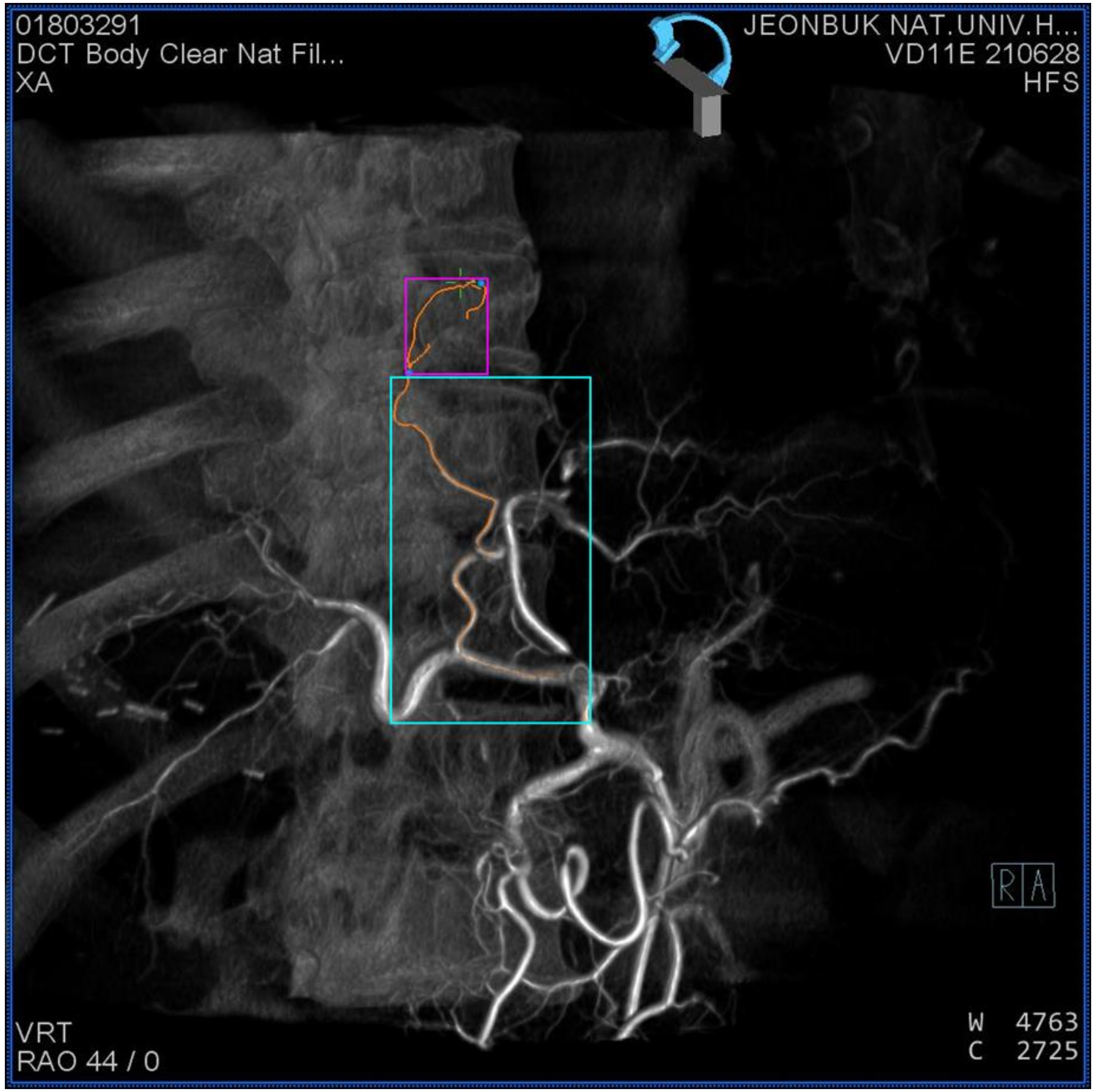
From PPT	Orange Line Redrawn (Pink)	Get Line Mask	Get Convex Hull Area (Max = Best?)



Optimal Viewpoint for Angiography

Future Work

- Continue developing system, integrating more components, refining, etc.
  - Would like to incorporate aspects like “tip” (magenta) vs. “branch” (cyan) maximization, and possibly use information about “holes” in the line segmentation to determine undesirable vessel cross-overs.
- Work with radiologists to develop a metric.
  - E.g., weight of “best” viewing rotation+angle vs. a nearby rotation+angle.
- Possibly (if we can get a sufficiently large annotated dataset) integrate some form of machine/deep learning to the project.

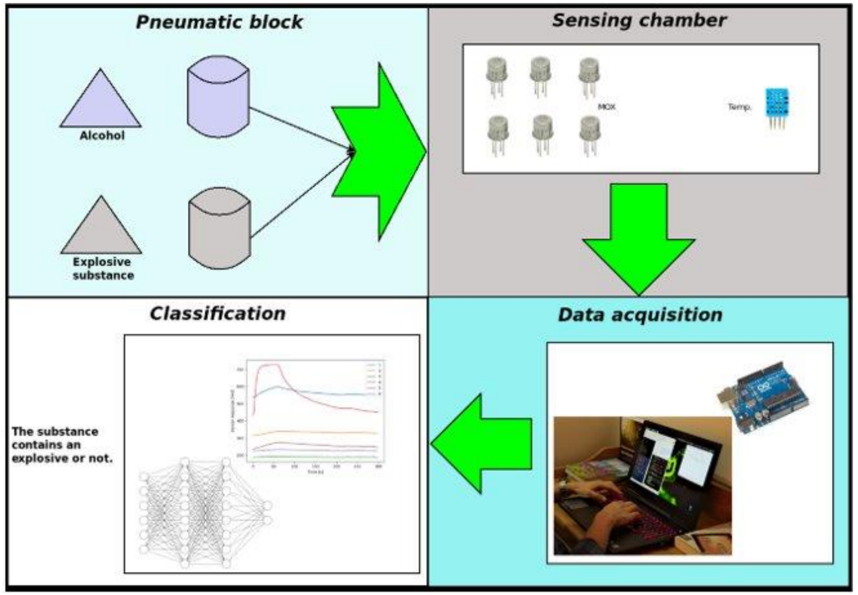


Using a Low-Cost Electronic Nose

- Entities throughout the world face the problem of detecting hidden explosive devices where human and canine inspection might not be a viable solution.
- The goal is to develop a fast and light-weight classification model to be used in an electronic nose to identify very small concentrations of explosive substances (Gunpowder and trinitrotoluene), by means of deep learning.
- 149 samples were taken, combining TNT or gunpowder with either soap or toothpaste, or acquiring raw samples of those substances in amounts ranging from as low as 0.1 g to 2 g.
- For the classification problem, five models were evaluated: k Nearest Neighbor, Support Vector Machine, Random Forest, Convolutional Neural Network and Long Short Term Memory.

Using a Low-Cost Electronic Nose

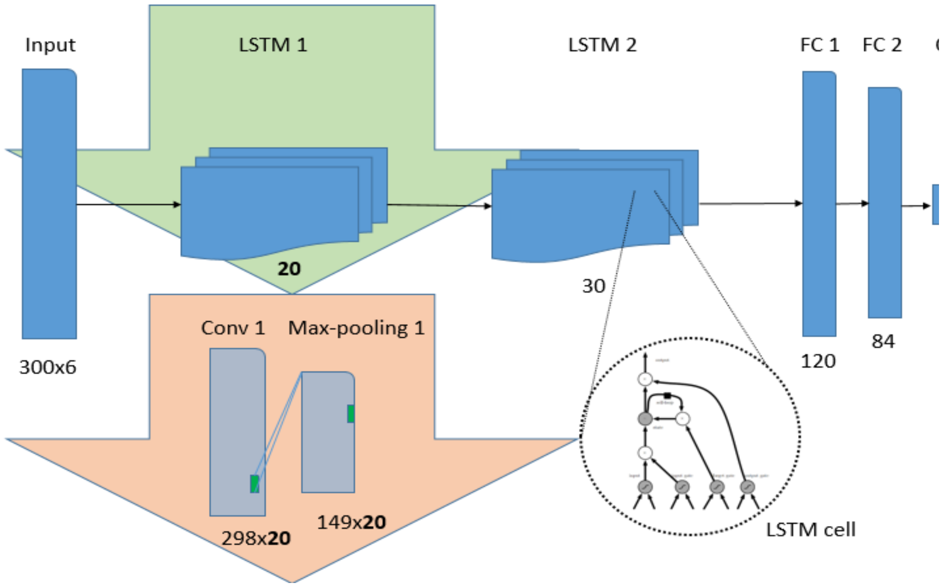
- The goal is to develop a fast and light-weight classification model to be used in an electronic nose to identify very small concentrations of explosive substances (Gunpowder and trinitrotoluene), by means of deep learning.
- 149 samples were taken, combining TNT or gunpowder with either soap or toothpaste, or acquiring raw samples of those substances in amounts ranging from as low as 0.1 g to 2 g.
- For the classification problem, five models were evaluated: k Nearest Neighbor, Support Vector Machine, Random Forest, Convolutional Neural Network and Long Short Term Memory.

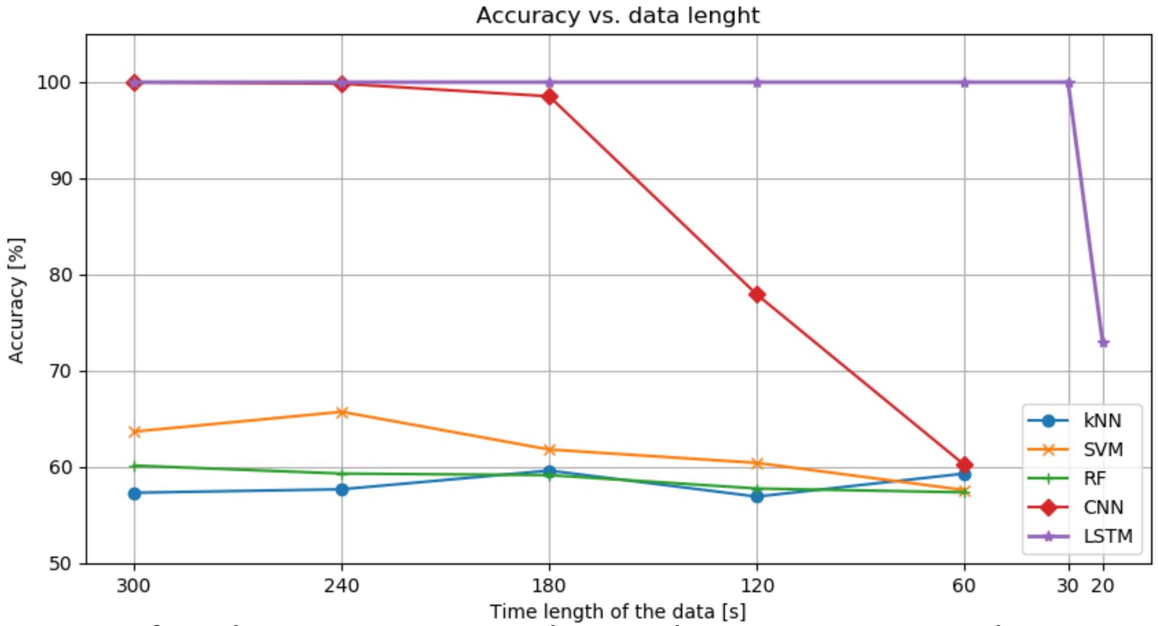


Deep Learning for Fast Detection of Explosives

The Final Model

- It implements an LSTM modification of a LeNet model, already proven to work in gas classification.
- It maintains good accuracy with only 30 seconds. The samples captured data during 6 minutes. This speeds up detection and reduces model size.





J. Torres-Tello, A. V. Guaman and S. Ko, 2020. “Improving the Detection of Explosives in a MOX Chemical Sensors Array with LSTM Networks,” *IEEE Sensors Journal*, Vol. 20, Iss. 23, pp. 14302-14309.

Yield Prediction in Aeroponics

UNIVERSITY OF SASKATCHEWAN

Ensemble Learning for Improving Generalization in Aeroponics Yield Prediction

- Aeroponic process can grow 30% more food up to 3 times faster than traditional methods, using a fraction of the water and land.
- Three ML models have been implemented and evaluated: SVM, RF, and NN.
- The last two gave good results, and were used to create an Ensemble predictor for yield, that increased the accuracy of the predictions.
- Six crops were evaluated, with an average R^2 value of 0.81 over the complete dataset.
- Aeropod (Farmboys) is a system that allows total control of variables, and records data of crop production.

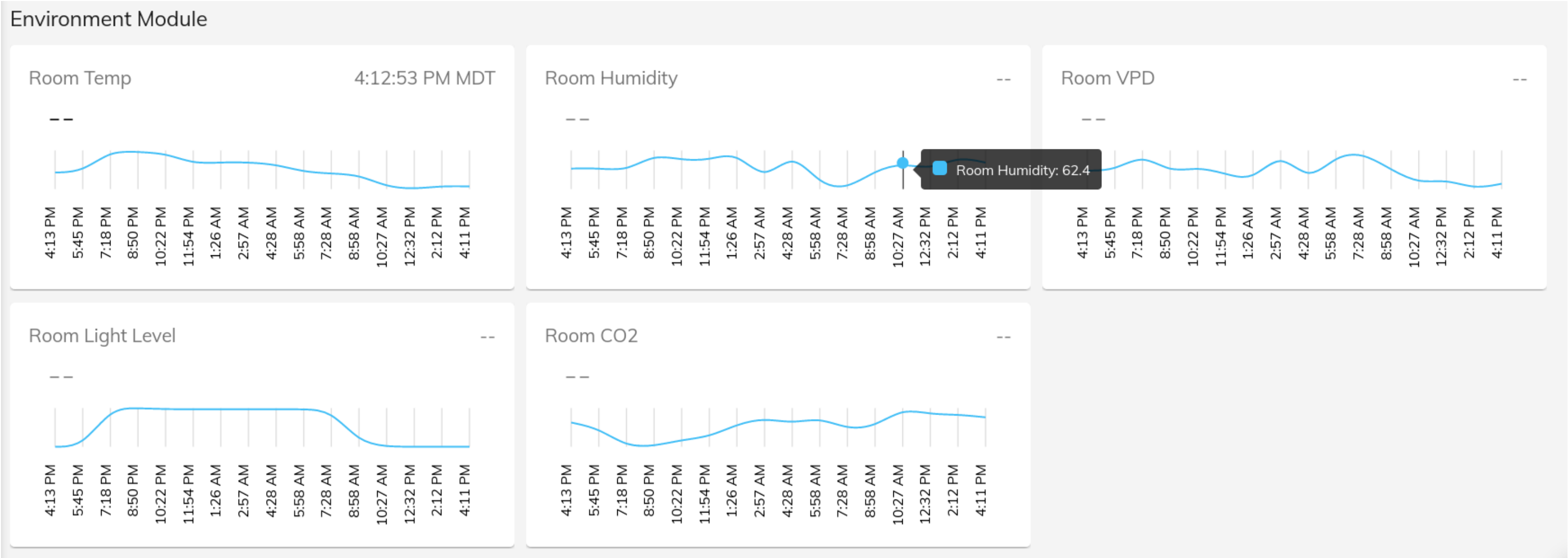
J. Torres-Tello, S. Venkatachalam, L. Moreno and S. -B. Ko, "Ensemble Learning for Improving Generalization in Aeroponics Yield Prediction," 2020 *IEEE International Symposium on Circuits and Systems*, Sevilla, 2020, pp. 1-5, doi: 10.1109/ISCAS45731.2020.9181283.

Yield Prediction in Aeroponics (ongoing)

UNIVERSITY OF SASKATCHEWAN

Interpretability of AI Models that use Data Fusion to Predict Yield in Aeroponics

- This work has two main goals: (i) use data fusion to improve yield prediction in aeroponics, and (ii) find which features are more relevant for yield prediction of six different crops.
- To reach these goals, a number of artificial intelligence models and an interpretability analysis based on Shapely Additive exPlanations (SHAP) have been implemented.
- The models were trained using 200 samples that were collected for almost a year, including information from different air and water quality sensors besides manually recorded data.

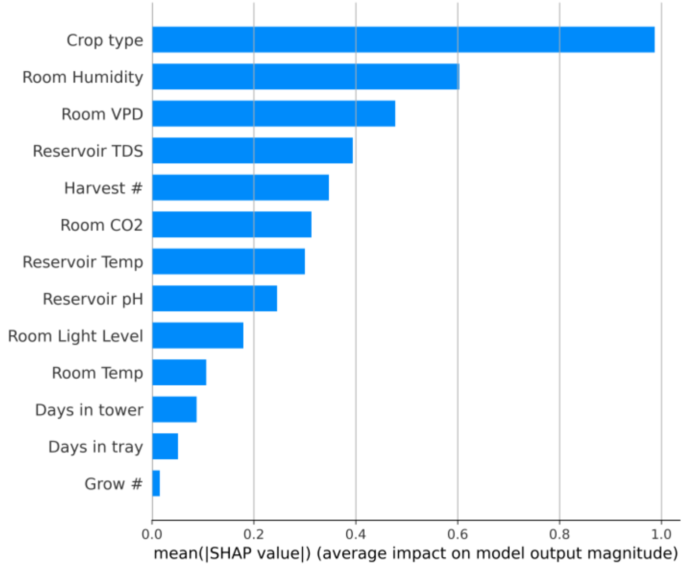


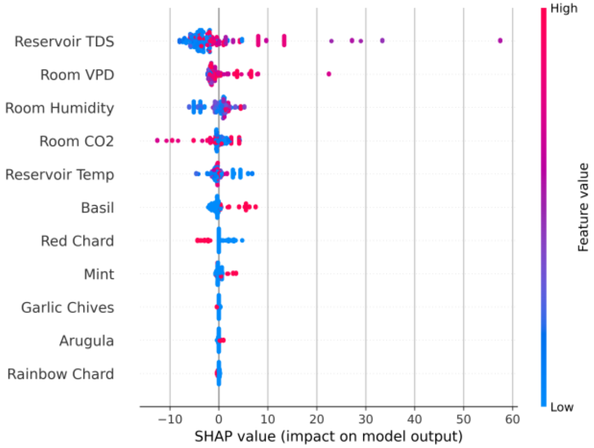
Yield Prediction in Aeroponics (ongoing)

UNIVERSITY OF SASKATCHEWAN

Interpretability of AI Models that use Data Fusion to Predict Yield in Aeroponics

- Our models reached a coefficient of determination value $R^2 = 0.752$ for the validation dataset in the best case (CNN-based model).
- As a result, two main features were identified in the dataset: Room CO_2 and Reservoir Temperature.
- SHAP values also provided important information for feature selection. These results could be the first steps towards the full automation of an aeroponics crop production system.



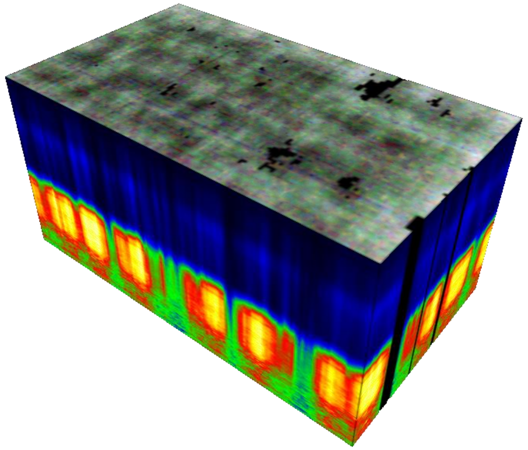


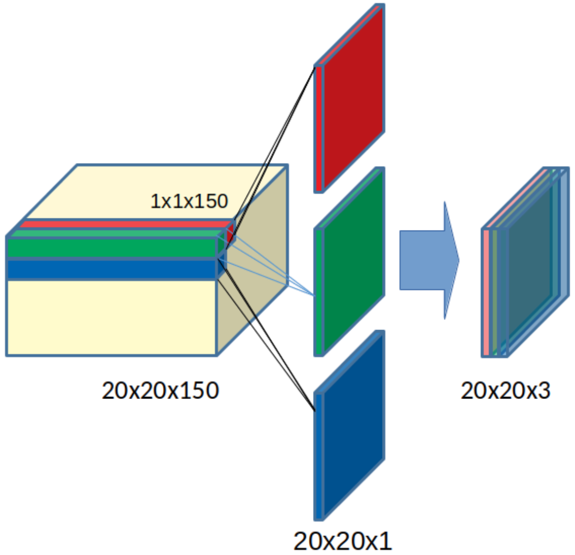
Plant phenotyping from Hyperspectral Images

UNIVERSITY OF SASKATCHEWAN

Transfer learning from RGB to HSI by means of pointwise convolutions

- A common technique for small datasets is the use of pre-trained models.
- However, the availability of such models is mainly in the RGB image domain. Therefore, this work explores the potential of pointwise convolutions to adapt the higher number of channels in HSI (150) to match the dimensions of a pre-trained VGG-16 model (3 channels).
- By using this technique, we could train a model that predicts the moisture content of canola plots.
- The developed DL algorithm over the test dataset resulted in a R^2 of 0.77.





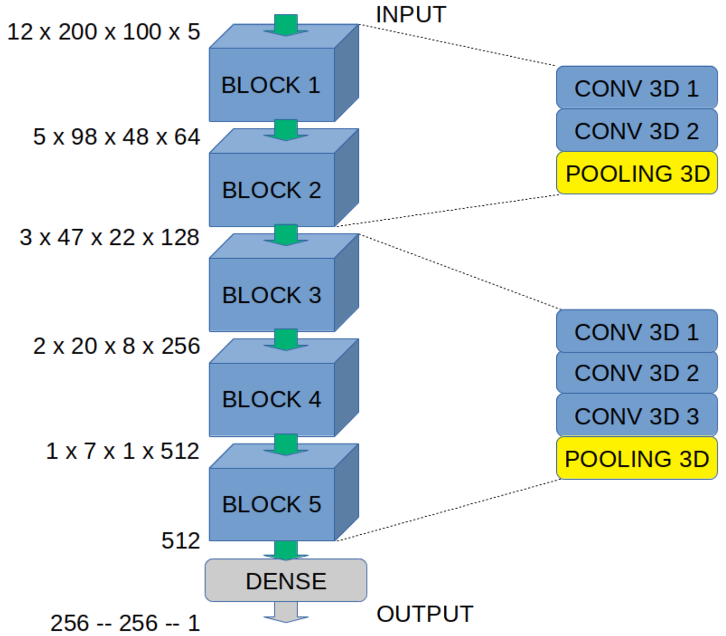
Julio Torres-Tello, Keshav D Singh, Seok-Bum Ko, Steve Shirtliffe, "Transfer Learning from RGB to Hyperspectral Images by means of Pointwise Convolutions", *5th Annual P²IRC Symposium*, 2020, Winner of Best Poster Competition.

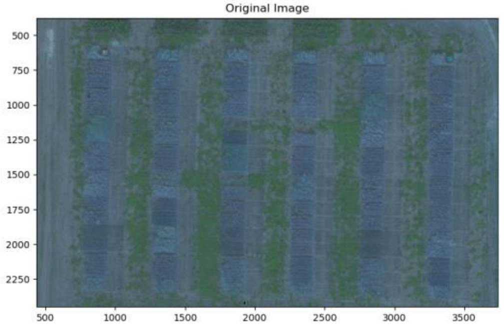
Multi/Hyper Spectral Images and Deep Learning

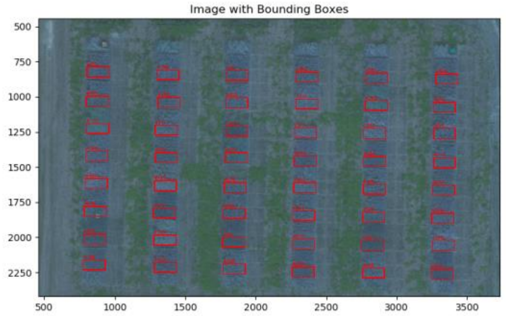
UNIVERSITY OF SASKATCHEWAN

Identifying Useful Features in MSI with DL for Optimizing Wheat Yield Prediction

- A DL based approach for optimizing the yield prediction process of spring wheat, using multispectral images.
- We assessed both the temporal features to find the most valuable time to take images, as well as the contribution of spectral bands.
- The most crucial flying times for acquiring images were at late-heading, late-flowering, dough-development, and harvesting stages.
- The two most useful colour-bands for yield prediction were red and red-edge.





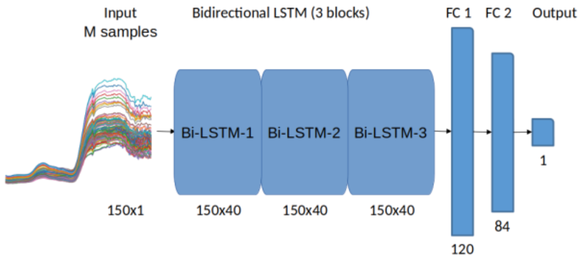


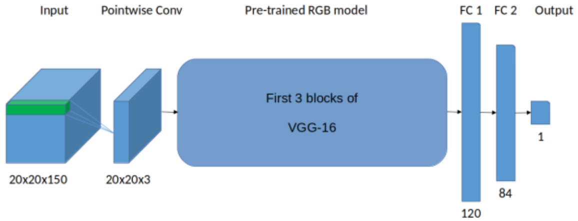
J. Torres-Tello and S. -B. Ko, "Identifying Useful Features in Multispectral Images with Deep Learning for Optimizing Wheat Yield Prediction," 2021 *IEEE International Symposium on Circuits and Systems*, 2021, pp. 1-5, doi: 10.1109/ISCAS51556.2021.9401360.

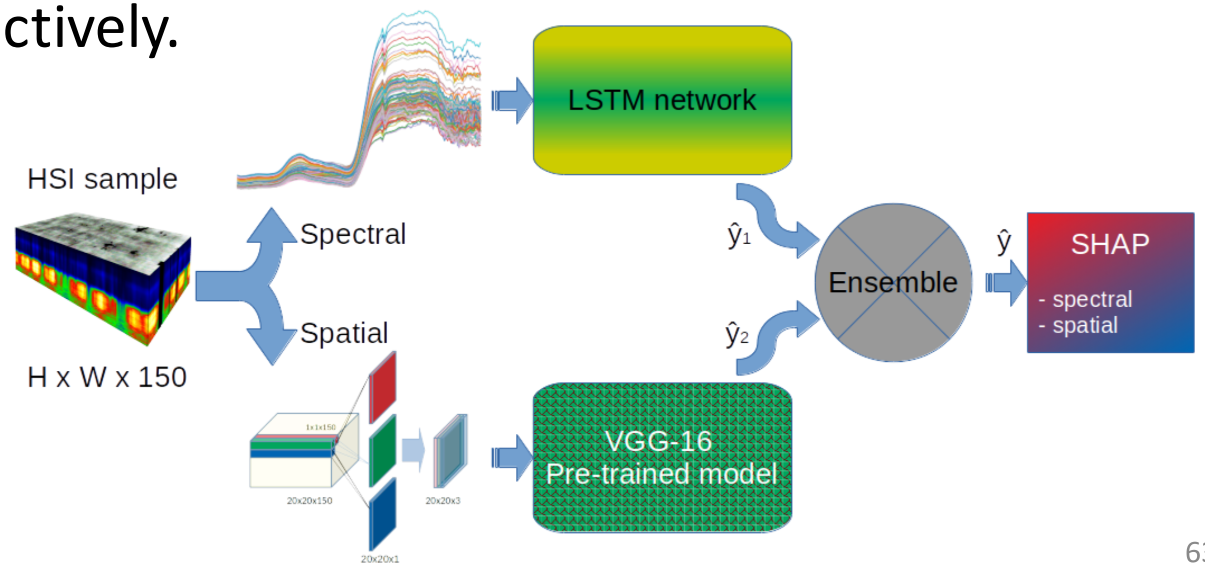
Multi/Hyper Spectral Images and Deep Learning

A Novel Approach to Identify the Spectral Bands that Predict Moisture Content in Canola and Wheat (ongoing)

- Moisture content prediction is relevant for assessing the degree of maturity of a crop, which relates to efficient harvesting and quality control.
- An accurate DL model for the prediction of the moisture content of canola and wheat crops, based on hyperspectral images taken by several drone flights.
- The model includes a final ensemble of two branches for analysis of spatial and spectral features, and it reached a coefficient of determination of 0.916 and 0.818 for the canola and wheat test datasets, respectively.





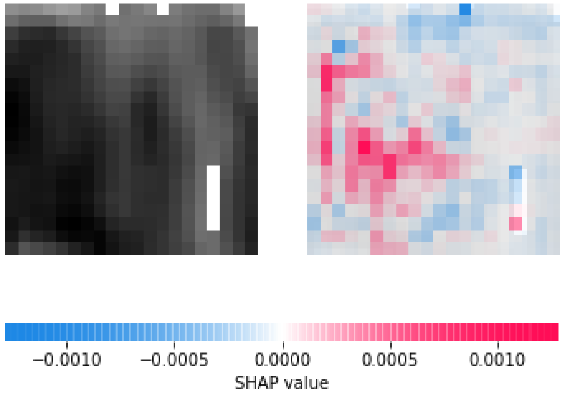


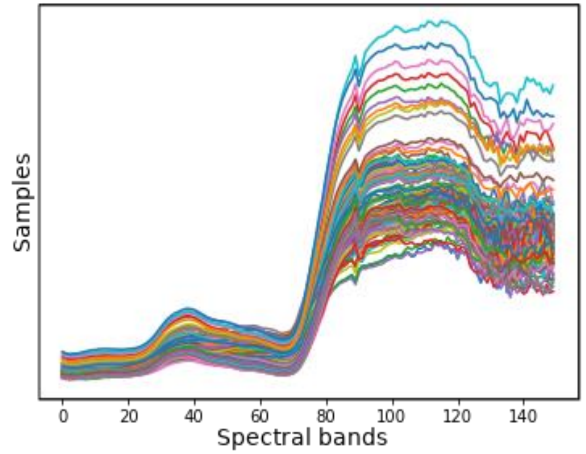
Multi/Hyper Spectral Images and Deep Learning

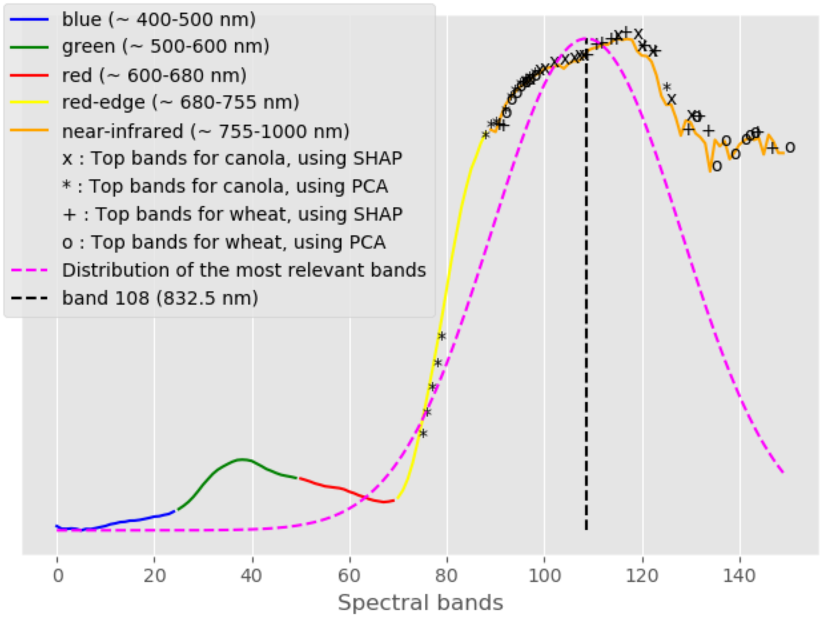
A Novel Approach to Identify the Spectral Bands that Predict Moisture Content in Canola and Wheat

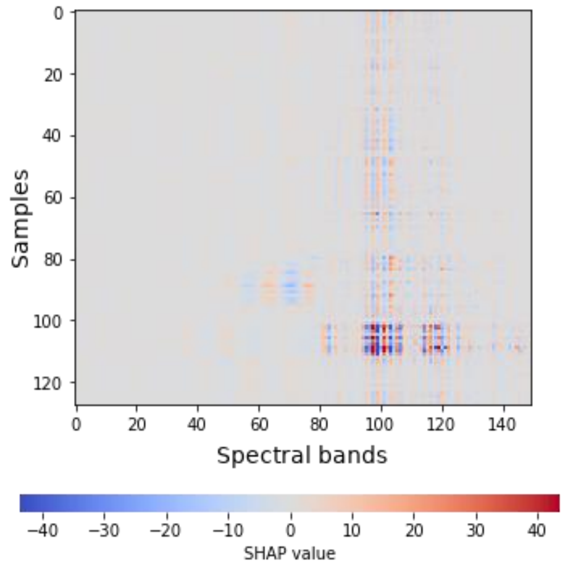
- SHapley Additive exPlanations analysis allowed us to study the individual predictions of the models.
- Using this approach actually obtains the spectral bands that are important for this task, since they are similar to PCA results, and they fall on the NIR part of the spectrum, which is widely used in moisture measurement of agricultural products and vegetation analysis.

J. Torres-Tello and S. Ko, 2021. "A novel approach to identify the spectral bands that predict moisture content in canola and wheat," *Biosystems Engineering*, Vol. 210, pp. 91-103.



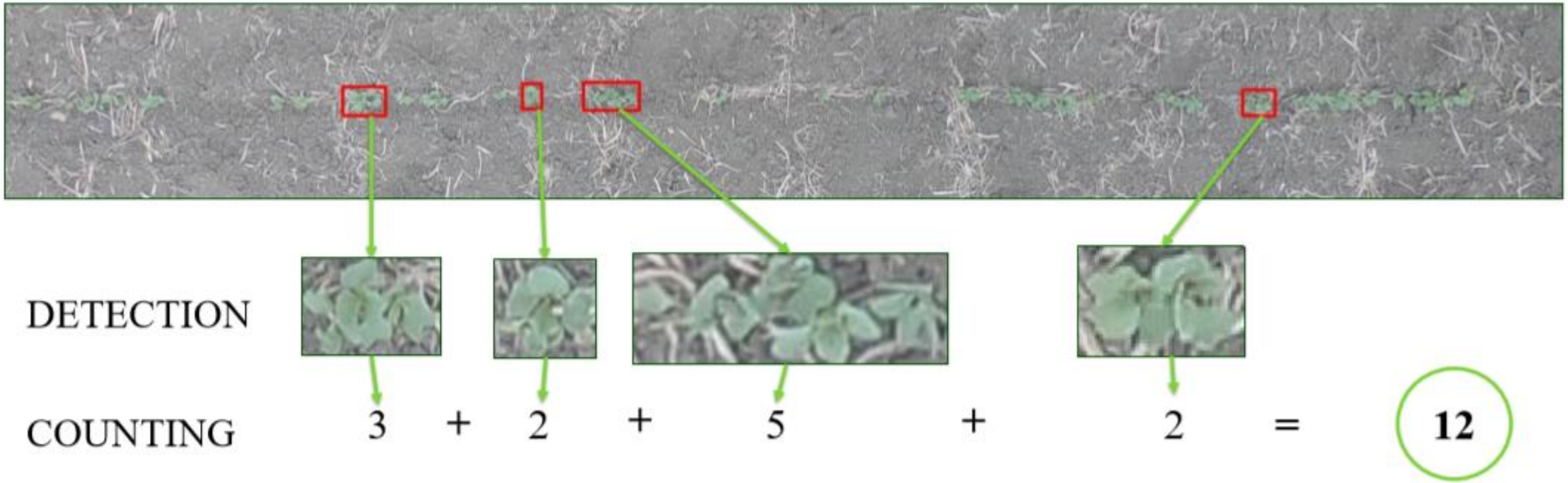






A pipeline for *Brassica Carinata* Emergence counting developed in partnership with Computer Science and Plant Science for P²IRC.

It is slow, tedious, and prone to error to count manually. Counting is valuable for yield, and thus is important to farmers and crop phenotypers.

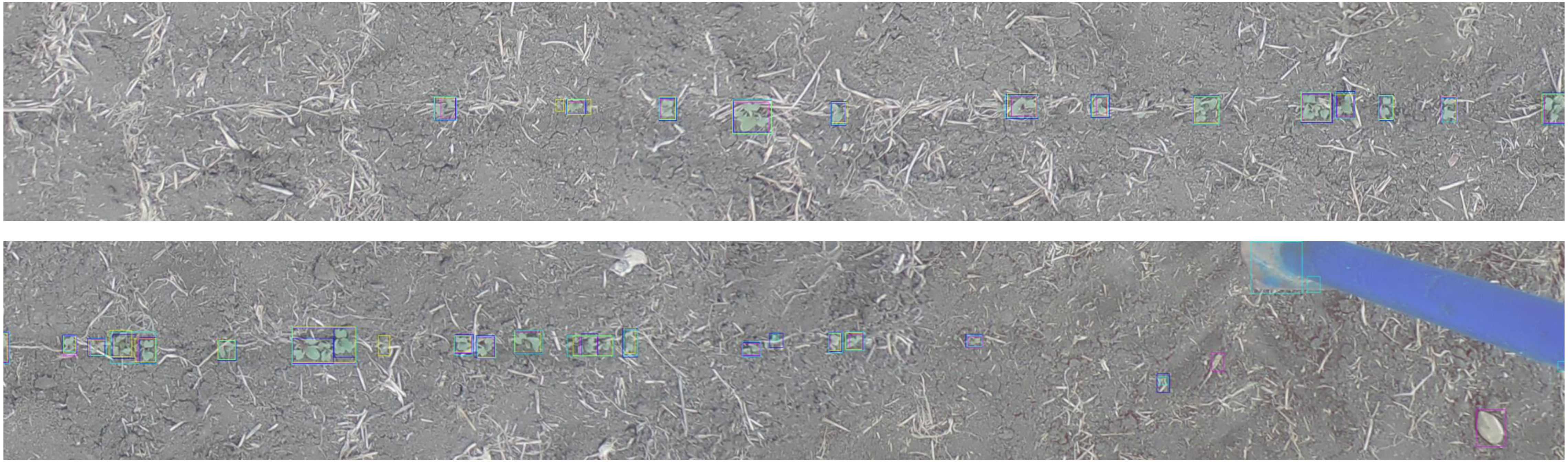


Leyeza, B., Seidenthal, K., Aziz, S., Castro-Zunti, R., Molahasani, M., Stavness, I., Ko, S., Vail, S., and Eramian, M. (October 2019) “Clump it up!: clump-based emergence counting of brassica carinata.” Poster presented at the (International) *4th Annual P²IRC Symposium*, Saskatoon, SK. Winner of Best Poster Competition.

Used combinations of 6 object detection methods paired with 12 counting methods.

Object Detection part example.

Example of 1 row divided into 2 parts.



SSDlite+MobilenetV2 = blue

SSDlite+ResNet = magenta

YOLOv3 = yellow

U-Net = cyan

Brassica Carinata Emergence Object Detection and Counting

UNIVERSITY OF SASKATCHEWAN

Results for object detection.

Best Scores Per Model	F1-Score	Train. Augs.	Dice Similarity	Train. Augs.
SSDlite + MobileNetV2	0.761	F, R, Z, B, C, H, S	0.820	F, R, Z, B, C, H, S
SSDlite + ResNet	0.788	F, R, Z	0.849	F, R, Z, B, C, H, S
YOLOv3	0.791	F, R, Z, B, C, S	0.038	F, R, Z, B, C
U-Net	0.723	F, R, Z, B, C, S	0.847	F, R, Z, B, C, H, S

Results for best models of the 72 models tried.

Detection Model	Counting Model	Abs. CountDiff	CountDiff
YOLOv3	Simple CNN (Synthetic)	**5.4729**	0.6997
SSDlite + MobileNetV2	Simple CNN (Synthetic)	5.6103	**0.1225**
Faster R-CNN	Encoder CNN (Real)	5.7497	1.8893
Faster R-CNN	Simple CNN (Synthetic)	6.0872	3.8450

- AbsCountDiff - absolute difference between the predicted and true total number of plants in a row
- CountDiff – The difference between the predicted and true total number of plants in a row

Automatic license plate recognition (ALPR) is fairly easy for humans, but difficult for computers—especially when plates are blurry, skewed, partially occluded, over or underexposed, etc.—because standard image processing techniques to locate a license plate amongst a complex background, and recognize the characters on the plate, fail.

The Problem

Finding the plate in complex background

Different obstructions on license plates in images from the NTUA Medialab dataset. (a) Rotated, Underexposed Plate. (b) Blurry, Overexposed Plate. (c) Skewed Plate. (d) Partially Shadow-covered Plate.

The Goal

UNIVERSITY OF SASKATCHEWAN

Deep learning can help for both localization and recognition! We use a MobileNet-like architecture: a modified SSD object detection network with linear bottlenecks and depth-separable convolutions (most important) for speeding convolution with negligible accuracy loss.

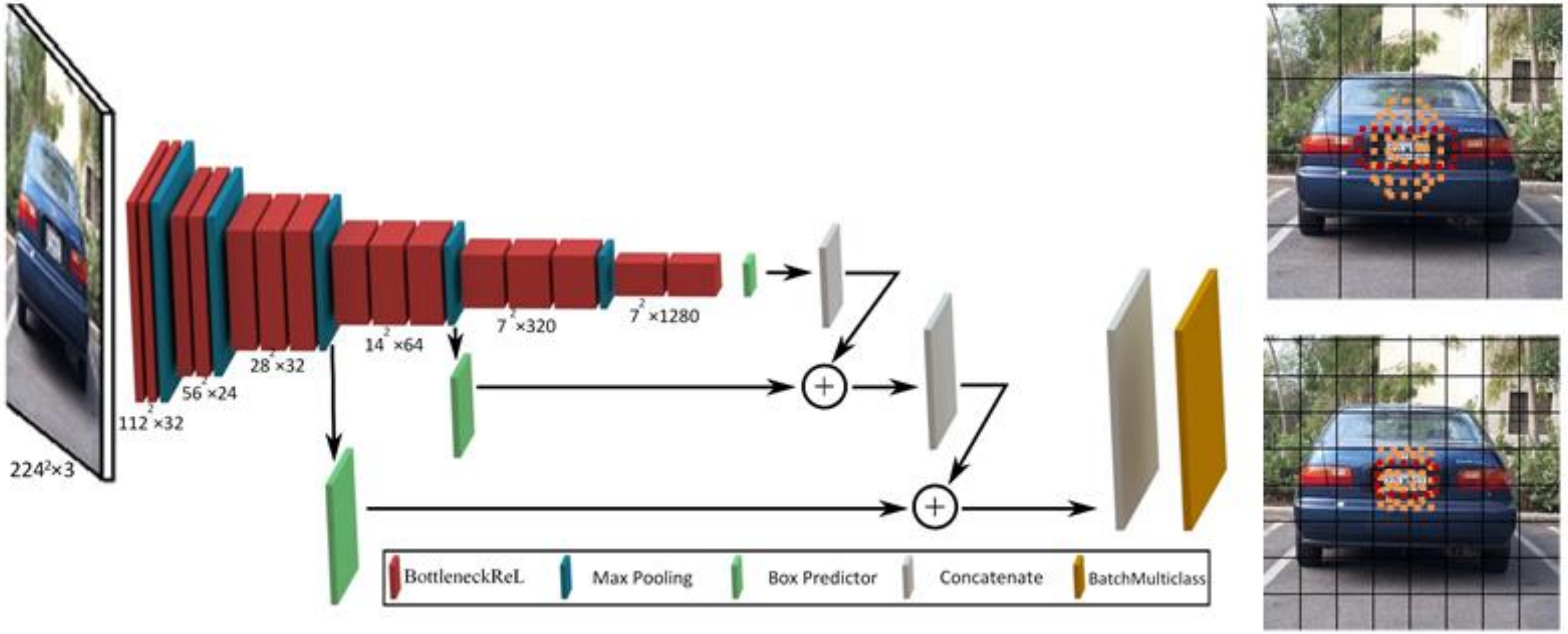
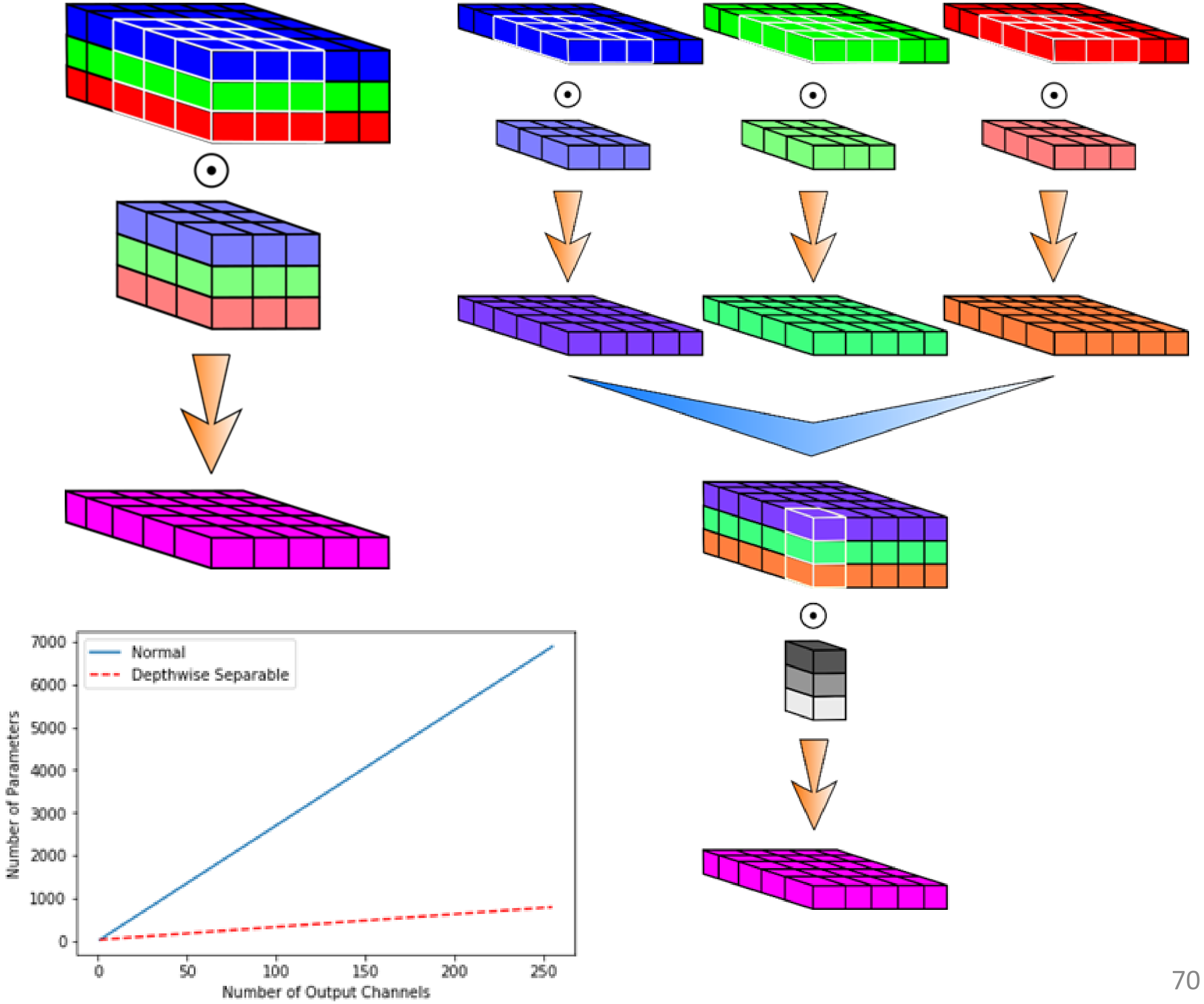


Fig. 2. Architecture of our license plate localization method. This network is based on Single Shot MultiBox Detector, but differs in that our architecture uses depthwise separable convolution, as opposed to standard convolutions, and layers of linear bottlenecks with inverted residuals. Note that the bottleneck layers contain the depthwise separable convolutions, which are comprised of a depthwise convolution followed by a pointwise convolution. Car images from the Caltech dataset.

Yépez, Juan; Castro-Zunti, Riel D.; Ko, Seok-Bum: 'Deep learning-based embedded license plate localisation system', *IET Intelligent Transport Systems*, 2019, 13, (10), p. 1569-1578, DOI: 10.1049/iet-its.2019.0082IET Digital Library, https://digital-library.theiet.org/content/journals/10.1049/iet-its.2019.0082

Depthwise Separable Convolution (DSC)

- A DSC is a two-step process involving a depthwise convolution followed by a pointwise convolution [1], [2]. A depthwise convolution performs filtering over multiple channels while allowing the channels to remain separate [2]. DSCs (right) reduce computation compared to standard convolution (top left) by a factor of $k^2d_j/(k^2 + d_j)$, where k is the kernel size and d_j is the dimension of the output channel layers (computational savings graph bottom left).



We trained/tested the localization system on license plates from 3 public datasets (Caltech Cars [American], NTUA [Greek], and University of Zagreb [Croatian]). Accuracy:

Comparison of license plate localization algorithms on the Caltech Cars 1999 (rear) 2 dataset

Description of System/Algorithm	Correct Detection %	Processing Time Per Image (s)
Ours	**98.4**	**0.02**
Faster R-CNN for vehicle detection + CNN Classifier [3]	98.39	None given
Faster R-CNN + RPN [4]	98.04	0.279 (estimated)
Line Density Filter + SVM-based Classifier [5]	91.27	< 0.042 (estimated)
Feature Extraction + Principal Visual Word [6]	84.4	7.19

Comparison of license plate localization algorithms on the University of Zagreb plate detection, recognition, and automated storage dataset

Description of System/Algorithm	Correct Detection %	Processing Time Per Image (s)
Ours	**97.83**	**0.02**
Corner-point Detection + Linear Discriminant Analysis-based Classifier [7]	92.8	0.12 (estimated)

Comparison of License Plate Localization Algorithms on the NTUA Medialab LPR Database

Description of System/Algorithm	Correct Detection Percentage By Set # (%)								Correct Detection %	Processing Time Per Image (s)
	1	2	3	4	5	6	7	8		
Ours	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**99.37**	**99.8**	**0.02**
Morphological Operations [8]	100	100	100	97	100	100	100	95.65	98.45	0.02
Connected Component Analysis [9]	92.02	82.48	88.73	87.24	74	90.84	N/A	N/A	89.45 (sets 1 through 6 only)	0.035

We tested the localization system using Saskatchewan (Canadian) License Plates.

- We used a low-cost embedded system
  - Raspberry Pi 3 with a quad core 1.2GHz processing power chip and 1 GB of RAM.
  - Intel Neural Compute Stick 2 (NCS2)
    - Enables a CNN to be deployed on a low-power chip for real-time inference
    - No connection to the cloud or large processing server
    - Allows 1 trillion operations per seconds (TOPS)
  - Samsung S5K2L1 camera
    - 12 MP resolution and a sensor of 1.4 µm
    - 1/2.6" and 10x optical zoom was used
    - 60 FPS video streamed to the Raspberry Pi 3.
- 99.77% localization accuracy over 898 vehicles.
- The system can run at an average of 13 FPS.
- Multiple plates can be localized in the same frame.
- https://www.youtube.com/watch?v=7eyfGCW_UwQ

We trained the recognition system on Californian license plates from the Caltech Cars and UCSD-Stills, with dataset augmentations via the following image transformations:

Normal

Equalizing the histogram of the Y channel in YUV color space

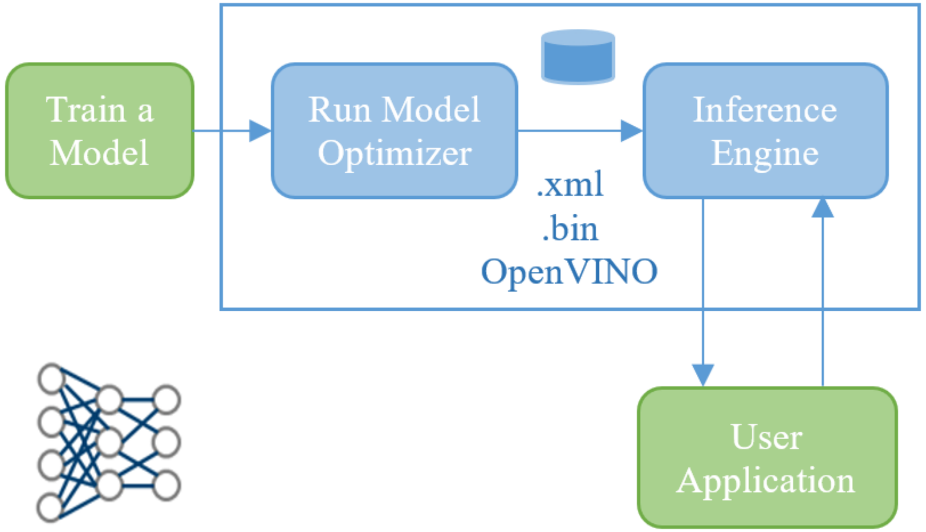
Horizontal motion blurring

Greyscale representation (V channel in HSV color space)

Greyscale representation (OpenCV BGR2GRAY)

Binarizing (Otsu's Thresholding), and its inverse binarization

We used OpenVINO to accelerate the recognition model onto a Raspberry Pi 3 + NNA (specifically, Neural Compute Stick 2) hardware platform.



OpenVINO performs static model analysis and redesigns a deep learning model for optimal execution on a target device. It consists of two parts: the Model Optimizer (MO), and the Inference Engine (IE). OpenVINO creates files for an Intermediate Representation (IR) using the MO, and the input to the MO is the network model trained using Tensorflow. The output of the MO is a model optimized for execution on a specific Intel CPU, GPU, VPU, FPGA, or a combination thereof.

The MO optimizes the model via the following mechanisms:

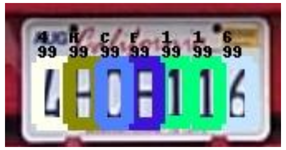
- Pruning extraneous model components that are required at the time of training, but not at the time of inference.
- Fusing operations. Some multiple operations can be combined into a single operation, and the MO detects such operations and fuses them.
- Bit-width reduction of weights (from 32- to 16-bit floating point) for compatibility with NCS2.

Deep Learning for License Plate Localization and Recognition

Results of our work, and against others' works:

Comparison to other academic works by dataset

Work	Per-Character OCR Accuracy (%)		Processing Time Per-Image (ms)	Size of Dataset Used in Training Process	Parameters
	Caltech Cars	UCSD-Stills			
Ours	96.23	99.79	14 ms (CPU+OpenVINO)	2,845 plates, or > 19,000 chars.	Approx. 3.4 M
Modified YOLO [10]	96.1 (without heuristics)	97.3 (without heuristics)	Approx. 14 ms (high-end GPU)	6,205 plates	Approx. 3.1 M
7-layer CNN [11]	94.28	N/A	None given	> 90,000 plates	Unclear From Paper
Curve Joining + 6-layer CNN [12]	94.8	N/A	None given	Approx. 368,000 chars.	Approx. 690,000






Castro-Zunti, Riel D.; Yépez, Juan; Ko, Seok-Bum: "License plate segmentation and recognition system using deep learning and OpenVINO." *IET Intelligent Transport Systems*, 2020, 14, (2), p. 119-126, DOI: 10.1049/iet-its.2019.0481 IET Digital Library, https://digital-library.theiet.org/content/journals/10.1049/iet-its.2019.0481

Current Work

- Use OpenVINO (or another acceleration method) with localization research

Future Work

- Expand plate recognition research to more use cases / jurisdictions
- Develop end-to-end system by combining both localization and recognition and retraining
- Deployment of system(s)

References

[1] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.: 'MobileNetV2: Inverted Residuals and Linear Bottlenecks'. *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, United States of America, 2018, pp. 4510-4520

[2] Bendersky, E.: 'Depthwise separable convolutions for machine learning', https://eli.thegreenplace.net/2018/depthwise-separable-convolutions-for-machine-learning/, accessed July 2018

[3] Kim, S., Jeon, H., Koo, J.: 'Deep-learning-based license plate detection method using vehicle region extraction', *Electronics Letters*, 2017, 53, (15), pp. 1034-1036

[4] Li, H., Wang, P., Shen, C.: 'Toward End-to-End Car License Plate Detection and Recognition With Deep Neural Networks', *IEEE Transactions On Intelligent Transportation Systems*, 2019, 20, (3), pp. 1126-1136

[5] Yuan, Y., Zou, W., Zhao, Y., Wang, X., Hu, X., Komodakis, N.: 'A Robust and Efficient Approach to License Plate Detection', *IEEE Transactions On Image Processing*, 2017, 26, (3), pp. 1102-1114

[6] Zhou, W., Li, H., Lu, Y., Tian, Q.: 'Principal Visual Word Discovery for Automatic License Plate Detection', *IEEE Transactions On Image Processing*, 2012, 21, (9), pp. 4269-4279[23] Wafy, M., Madbouly, A.: 'Efficient method for vehicle license plate identification based on learning a morphological feature', *IET Intelligent Transport Systems*, 2016, 10, (6), pp. 389-395

[7] Wafy, M., Madbouly, A.: 'Efficient method for vehicle license plate identification based on learning a morphological feature', *IET Intelligent Transport Systems*, 2016, 10, (6), pp. 389-395

[8] Yépez, J., Ko, S.: 'Improved License Plate Localization Algorithm Based on Morphological Operations', *IET Intelligent Transport Systems*, 2018, 12, (6), pp. 542-549

[9] Zhu, S., Dianat S., Mestha, L.: 'End-to-end system of license plate localization and recognition', *Journal of Electronic Imaging*, 2015, 24, (2), p. 23020

[10] Laroca, R., Zanlorensi, L., Gonçalves, G., *et al.*: 'An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO detector', arXiv.org: 1909.01754v2, accessed October 2018

[11] Polishetty, R., Roopaei M., Rad, P.: 'A Next-Generation Secure Cloud-Based Deep Learning License Plate Recognition for Smart Cities', in *IEEE Int. Conf. on Machine Learning and Applications (ICMLA)*, Anaheim, United States of America, 2016, pp. 286-294

[12] Selmi, Z., Halima, M., Alimi, A.: 'Deep Learning System for Automatic License Plate Detection and Recognition', in *IAPR Int. Conf. on Document Analysis and Recognition (ICDAR)*, Kyoto, Japan, 2017, pp. 1132-1138

University of Saskatchewan

INTERNATIONAL ROAD DYNAMICS INC.

General Objective

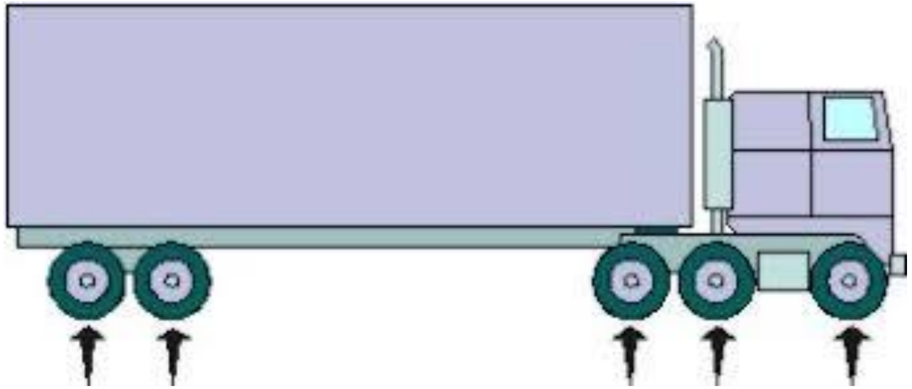
- A 3-year project to produce a low-cost fast edge microcontroller-based system that receives camera input and processes it to extract relevant information, including the following:
  - License plate number (see previous slide section)
  - HAZMAT decal[1] (current work)
  - CVSA inspection sticker[2] (current work)
  - USDOT code[3] (future work)
  - Vehicle Axle Spacing[4] (future work)
- In partnership with International Road Dynamics through a MITACS Accelerate Grant.

Mitacs

1


2


3


4


Deep Learning for Commercial Vehicle Characteristics

University of Saskatchewan

INTERNATIONAL ROAD DYNAMICS INC.

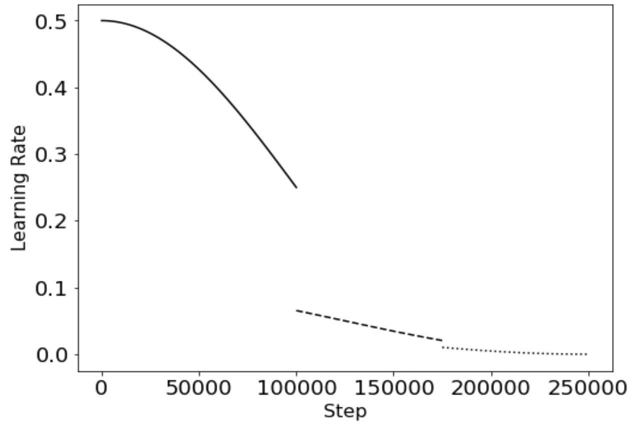
HAZMAT Detection and Classification

- 2229 HAZMAT placards over 1864 images mostly captured from a U.S. vehicle check stop.
- From MobileDet EdgeTPU [1] we developed a novel architecture trained quantize-aware (INT8) and with a custom piecewise cosine decay learning rate function.
- We implemented the model on a low cost ($115 USD) Raspberry Pi 4 and Google Coral USB Accelerator edge system.

TABLE III
NEURAL NETWORK ARCHITECTURE FOR HAZMAT PLACARD LOCALIZATION AND CLASSIFICATION: QUANTIZED FEATURE EXTRACTOR

Operator	Scaling Factor	Output Channels	Kernel	Stride
Conv2d+RELU6	N/A	32	3×3	2
Tucker	S_{IR}=0.25, S_{OR}=0.75	16	3×3	1
Fused	E=8	16	3×3	2
Fused*	E=4	16	3×3	1
Fused*	E=8	16	3×3	1
Fused*	E=4	16	3×3	1
Fused	E=8	40	5×5	2
Fused*	E=4	40	3×3	1
Fused*	E=4	40	5×5	1
IBN*	E=4	40	3×3	1
IBN	E=8	72	3×3	2
IBN*	E=8	72	3×3	1
Fused*	E=4	72	3×3	1
Fused*	E=4	72	3×3	1
IBN	E=8	96	5×5	1
IBN*	E=8	96	5×5	1
IBN†	E=8	120	5×5	1
IBN*	E=8	120	5×5	1
IBN	E=8	128	3×3	2
IBN*	E=4	128	5×5	1
IBN*	E=8	128	3×3	1
IBN†	E=8	384	5×5	1

* denotes the block uses a skip connection between its input and output
† denotes the block's output is an endpoint into the SSDlite architecture

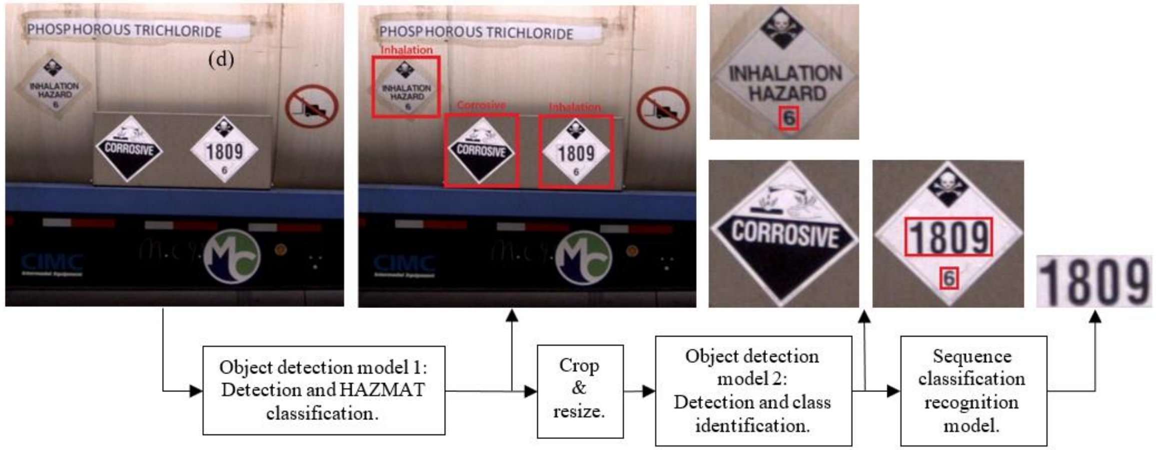


Castro-Zunti, R., Yépez, J., Choi, Y., Johnson, T., and Ko, S. "Real-Time Deep Learning-based HAZMAT Detection and Classification Edge System." submitted to *Springer Multimedia Tools and Applications*: 16 pages

Mitacs

HAZMAT Detection and Classification

- Over 15 HAZMAT classes, achieved 85.0% mAP@0.5.
- System achieved **104.49 FPS** on a Jetson Xavier and 30 FPS on Jetson Nano.
- Compared to a previous work [2] on an 8-class public dataset (on which we transfer learned our model) ours is 21-91% faster and generally more accurate (average of 98.8% F1-score vs. 82.5% accuracy).
- Our model works despite unideal lighting, plate skews, and complex backgrounds.



INTERNATIONAL ROAD DYNAMICS INC.

Mitacs

Deep Learning for Commercial Vehicle Characteristics

CVSA Detection and Recognition

- 2 stage object detection system developed by modifying MobileDet EdgeTPU [1] (baseline).
  - Stage 1 detects windshield, decal, and classifies decal colour. For better small object detection, 14% more parameters than baseline.
  - Stage 2 detects and classifies year digit and corner cut. Because easier task, 17% less parameters than baseline. If 3-7 decals found <30 frames from each other, use them all with majority vote.
- Models implemented in parallel on variety of edge computing / accelerator hardware.

TABLE I FIRST STAGE BACKBONE.

Input	Layer	*k*	*e*	*n*	*r*	*s*
$320^2 \times 3$	Conv	3×3	N/A	1	0	2
$160^2 \times 32$	Tucker	3×3	0.25-0.75	1	0	1
$160^2 \times 16$	Fused	3×3	8	1	0	2
$80^2 \times 16$	Fused	3×3	4	1	1	1
$80^2 \times 16$	Fused	3×3	8	1	1	1
$80^2 \times 16$	Fused	3×3	4	1	1	1
$80^2 \times 16$	Fused	5×5	8	1	0	2
$40^2 \times 40$	Fused	3×3	4	3	1	1
$40^2 \times 40$	IBN	3×3	8	1	0	2
$20^2 \times 72$	IBN	3×3	8	1	1	1
$20^2 \times 72$	Fused	3×3	4	2	1	1
$20^2 \times 72$	IBN	5×5	8	1	0	1
$20^2 \times 96$	IBN	5×5	8	1	1	1
$20^2 \times 96$ (C4)	IBN	3×3	8	2	1	1
$20^2 \times 96$	IBN	5×5	8	1	0	2
$10^2 \times 120$	IBN	3×3	8	1	1	1
$10^2 \times 120$	IBN	5×5	4	1	1	1
$10^2 \times 120$	IBN	3×3	8	1	1	1
$10^2 \times 120$ (C5)	IBN	5×5	8	1	1	1

INTERNATIONAL ROAD DYNAMICS INC.

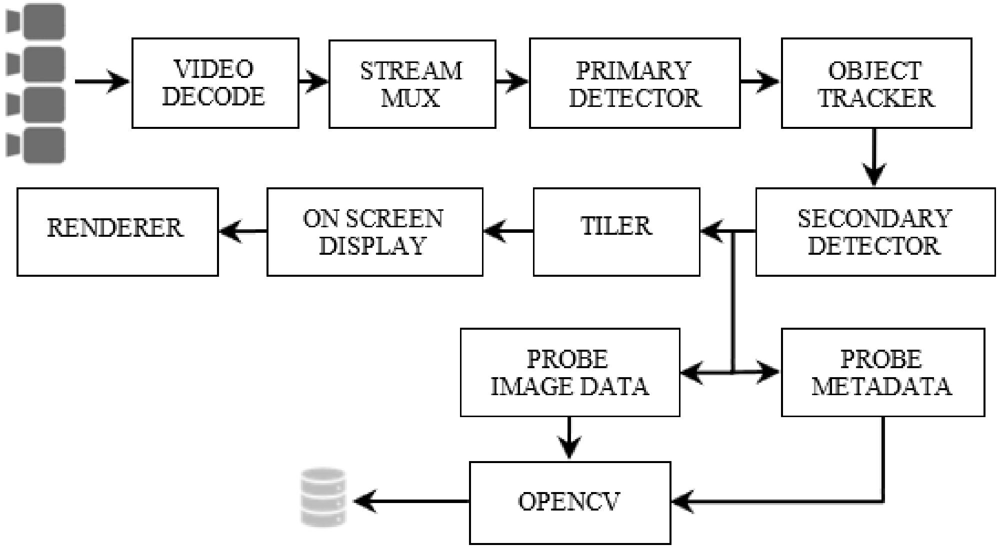
TABLE I SECOND STAGE BACKBONE.

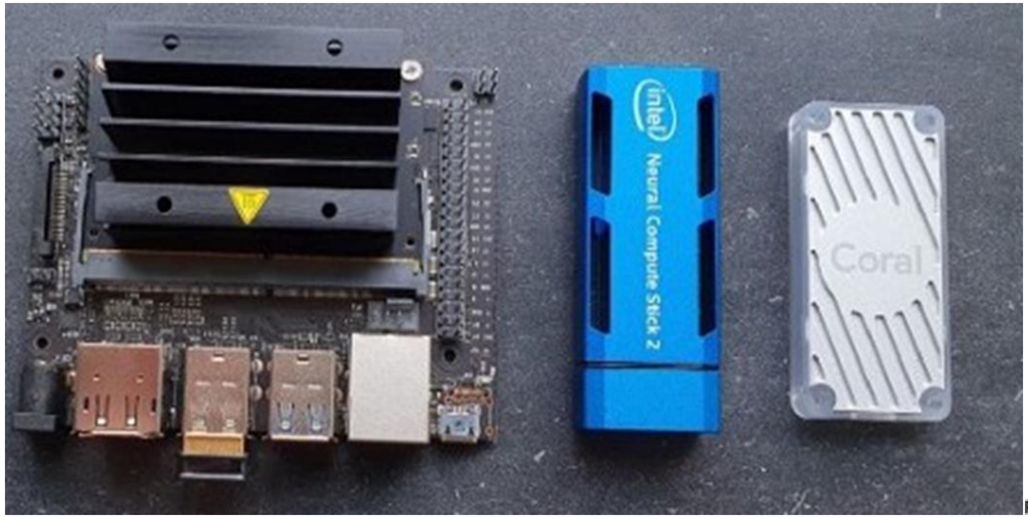
Input	Layer	*k*	*E*	*S*
$320^2 \times 3$	Conv	3×3	N/A	2
$160^2 \times 32$	Tucker	3×3	0.25-0.75	1
$160^2 \times 16$	Fused	3×3	8	2
$80^2 \times 16$	Fused	5×5	8	2
$40^2 \times 40$	IBN	3×3	8	2
$20^2 \times 72$ (C4)	IBN	5×5	8	1
$20^2 \times 96$ (C5)	IBN	5×5	8	2

Yépez, J., Castro-Zunti, R., Choi, Y., Johnson, T., and Ko, S. 2021. "Real-time CVSA Decal Recognition System Using Deep Convolutional Neural Network Architectures." *IET Intelligent Transport Systems*, Vol. 15, Iss. 11, pp. 1359-1371.

AI Hardware accelerators

- To assess the real-time suitability of our proposed system, we tested the following implementations of each custom model:
  - On the RPi4 alone, using TFLite.
  - On the RPi4 + NCS2 using OpenVINO's IR files.
  - On the RPi4 + Coral USB Accelerator using TFLite.
  - On the Jetsons (Nano and Xavier) using the converted TensorRT files.
- The figure shows the GStreamer pipeline used for the system. GStreamer can reads frames in different formats and even from different sources in parallel, process them, and export them to a file or stream them over a network.





Deep Learning for Commercial Vehicle Characteristics

University of Saskatchewan

International Road Dynamics Inc.

CVSA Detection and Recognition

TABLE III
SUMMARY OF MODEL TRAINING AND RESULTS

Model	Input Image (W×H×3)	MAdds (B)	Params (M)	File Size (MB)	mAP@0.5 Stage One (%)	mAP@0.5 Stage Two (%)	
						Single	7-spots
SSDLite+MobileDet EdgeTPU *[1]*	320×320	1.53	4.20	5.04	96.6	100	100
SSD+MobileNetV1 *[3]*	300×300	1.20	6.80	5.52	89.8	94.3	93.7
SSDLite+MobileNetV2 *[4]*	300×300	0.80	4.30	4.71	92.1	96.9	96.5
Proposed Stage One Model	320×320	1.77	4.80	5.20	**97.5**	—	—
Proposed Stage Two Model	320×320	0.38	3.47	1.44	—	**100**	**100**

TABLE IV
HARDWARE ACCELERATOR BENCHMARK

Device	Price USD	Bits	Stage one Pre-process (ms)	Stage one Inference (ms)	Stage one Post-process (ms)	Stage two Pre-process (ms)	Stage two Inference (ms)	Stage two Post-process (ms)	FPS*
NVIDIA Jetson Xavier (TRT)	$699	FP16	2.30	7.94	2.23	**2.11**	6.16	**2.15**	**81.90**
Coral USB Accelerator	$60	INT8	4.35	7.64	4.45	4.88	**5.91**	3.89	60.83
NVIDIA Jetson Nano (TRT)	$99	FP16	3.50	29.6	3.10	3.38	23.9	2.96	27.62
Intel NCS2	$69	FP16	5.85	29.19	5.45	5.34	28.98	4.72	24.70
NVIDIA Jetson Xavier	$699	FP32	4.75	48.75	3.51	4.31	39.91	3.18	17.54
Intel NCS	$40	FP16	6.19	107.06	5.73	5.56	88.48	4.86	8.43
NVIDIA Jetson Nano	$69	FP32	6.87	105.21	6.54	6.25	84.85	5.78	8.37
Raspberry Pi 4 (TF Lite)	$55	FP16	6.51	144.57	5.91	5.87	120.48	5.10	6.37
Raspberry Pi 4	$55	FP32	9.07	260.26	8.45	8.34	218.70	7.19	3.60

* The stages run in parallel, the speed is limited by stage one which is slower

Mitacs

Future Work

- Develop remaining deep learning application modules
  - USDOT
  - Vehicle Axle Spacing
- Package modules together in an edge hardware-optimized system

References

[1] Yunyang Xiong et al., "MobileDets: Searching for Object Detection Architectures for Mobile Accelerators," arXiv:2004.14525v1 [cs.CV], 2020.

[2] M. A. Mohamed, J. Tünnermann and B. Mertsching, "Seeing Signs of Danger: Attention-Accelerated Hazmat Label Detection," 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Philadelphia, PA, 2018, pp. 1-6.

[3] A. G. Howard et al., “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” in arXiv preprint arXiv:1704.04861, 2017.

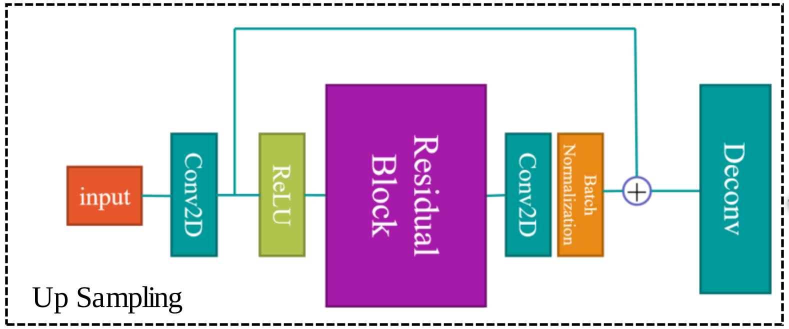
[4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520.

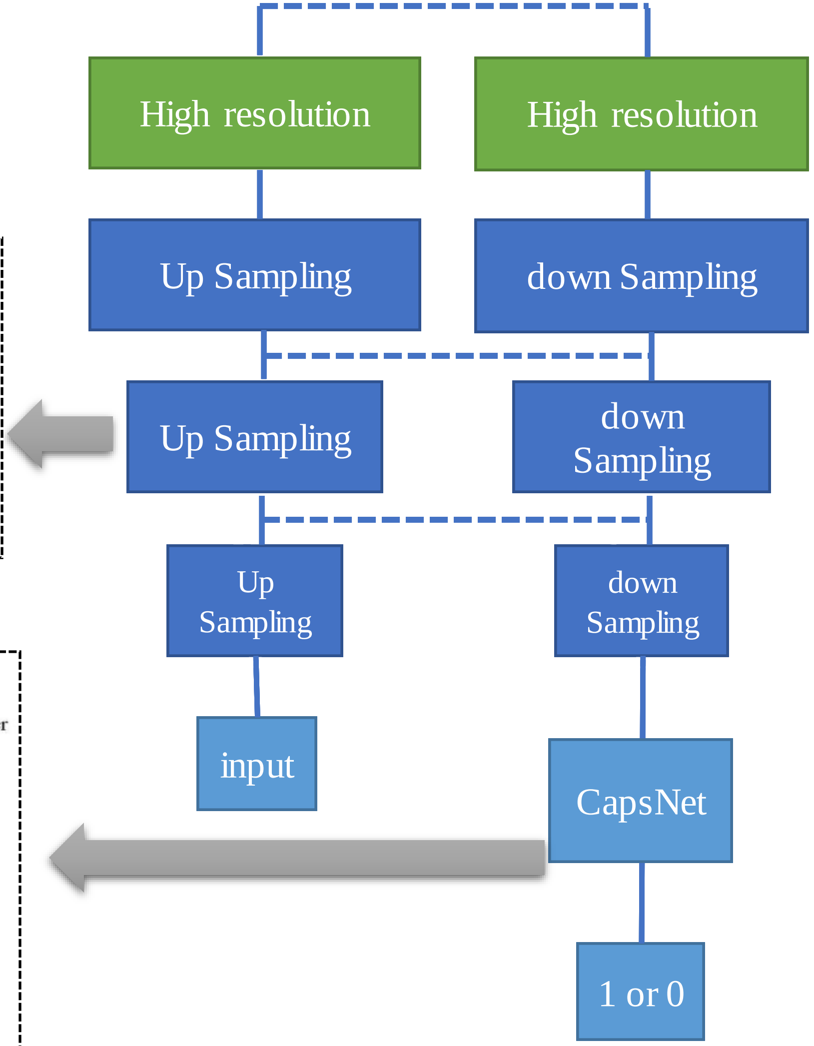
Face Super-Resolution with Deep Learning

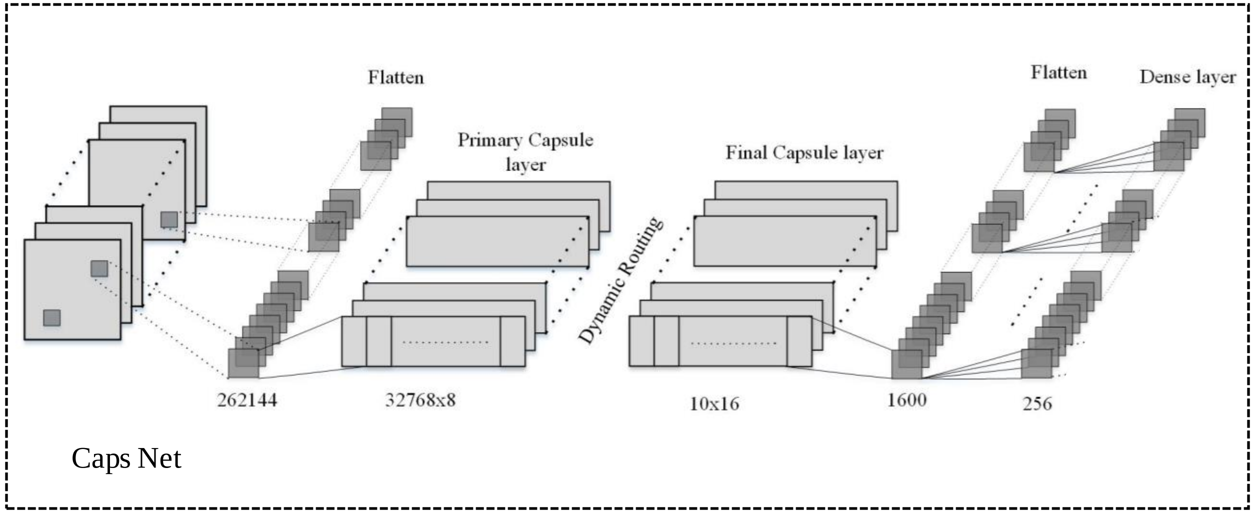
UNIVERSITY OF SASKATCHEWAN

MSG-CapsGAN: Multi-Scale Gradient Capsule GAN for Face Super Resolution

- First CapsGAN for Super-Resolution
- Using MSG-GAN for Super-Resolution for the first time

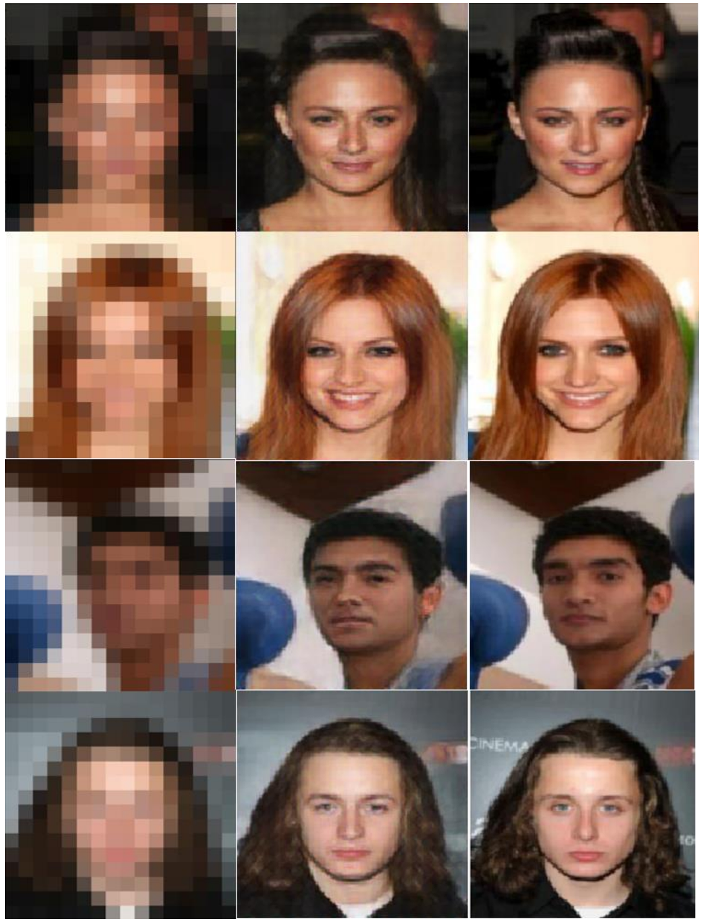






MSG-CapsGAN*: Multi-Scale Gradient Capsule GAN for Face Super Resolution

- Better Peak Signal to Noise Ratio (PSNR) than state-of-the-art
- Without using any attribute domain information

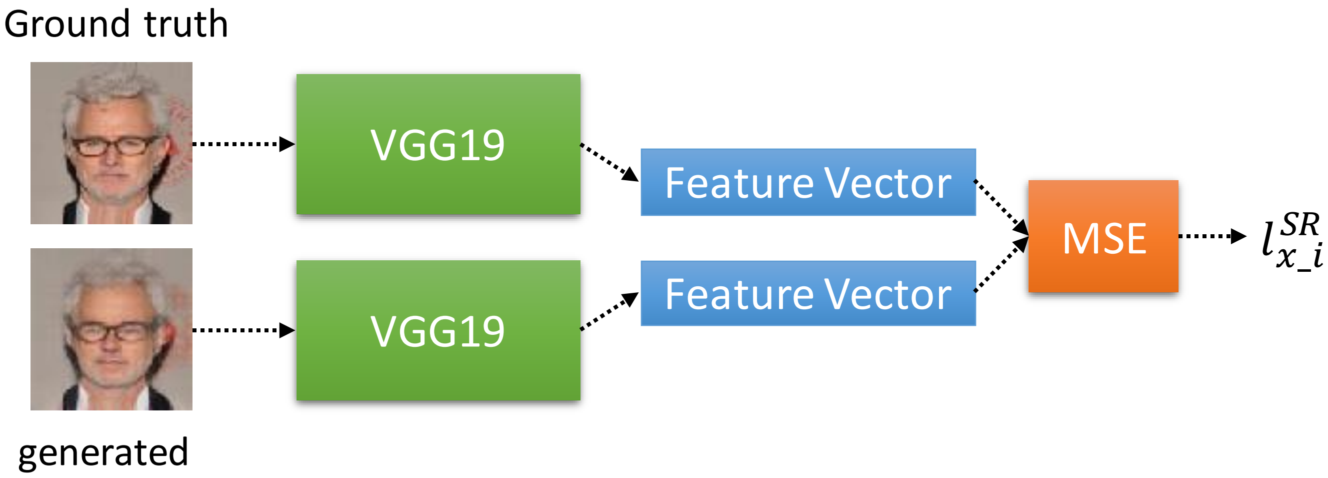


Method	PSNR	SSIM
Bilinear	20.85	0.574
VDSR [1]	22.96	0.652
Progressive [2]	22.66	**0.685**
Proposed MSG-CapsGAN	**23.35**	0.673

* Majdabadi, Mahdiyar Molahasani, and Seok-Bum Ko, "MSG-CapsGAN: Multi-Scale Gradient Capsule GAN for Face Super Resolution," *2020 International Conference on Electronics, Information, and Communication (ICEIC),* IEEE, 2020.

MSG-CapsGAN for multi-scale SR

- Use the error of all scales (x3,x4,x8) for tainting
- $l^{SR} = (l^{SR}_{x\_128} + l^{SR}_{x\_64} + l^{SR}_{x\_32}) + 3 \times 10^{-3} l^{SR}_{Gen}$
  - l^{SR}_{Gen} = adversarial loss
  - $l^{SR}_{x\_i}$ = content loss of $i \times i$ generated image



psnr = 25.355462 - ssim = 0.828645

Two examples of the generator's outputs

UNIVERSITY OF SASKATCHEWAN

References

[1] J. Kim, J. Kwon Lee, and K. Mu Lee, “Accurate image super-resolution using very deep convolutional networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1646–1654

[2] D. Kim, M. Kim, G. Kwon, and D.-S. Kim, “Progressive face super-resolution via attention to facial landmark,” arXiv preprint arXiv:1908.08239, 2019

[3] Rajpurkar, Pranav, Jeremy Irvin, Aarti Bagul, Daisy Ding, Tony Duan, Hershel Mehta, Brandon Yang et al. "Mura: Large dataset for abnormality detection in musculoskeletal radiographs." arXiv preprint arXiv:1712.06957 (2017).

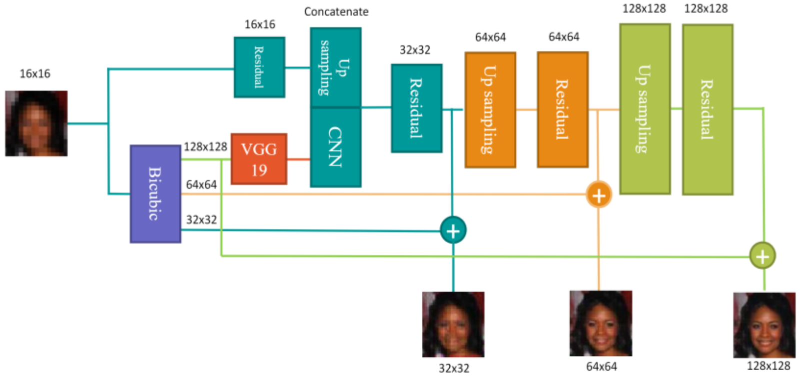
Robust Face Super-Resolution*

- **Surpassing the state-of-the-art systems in terms of PSNR, Structural SIMilarity (SSIM), Multi-Scale Structural SIMilarity(MS-SSIM), and Feature SIMilarity (FSIM)**
- **Outperforms the state-of-the-art face SR system in robustness**
- **Multiscale super-resolution**

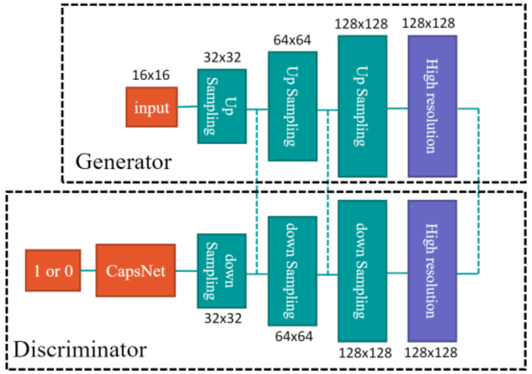
Introducing Feature SIMilarity (FSIM):

$$FSIM = e^{-C \times \frac{\sqrt{\frac{(f_1 - f_2)^2}{N}}}{f_2}}$$

Where $C = 0.3$, f is the feature vector extracted form VGG16 and N is the length of this vector.



Proposed Residual Generator



Proposed MSG-CapsGAN

Low resolution | bilinear | state-of-the-art | proposed model | ground truth

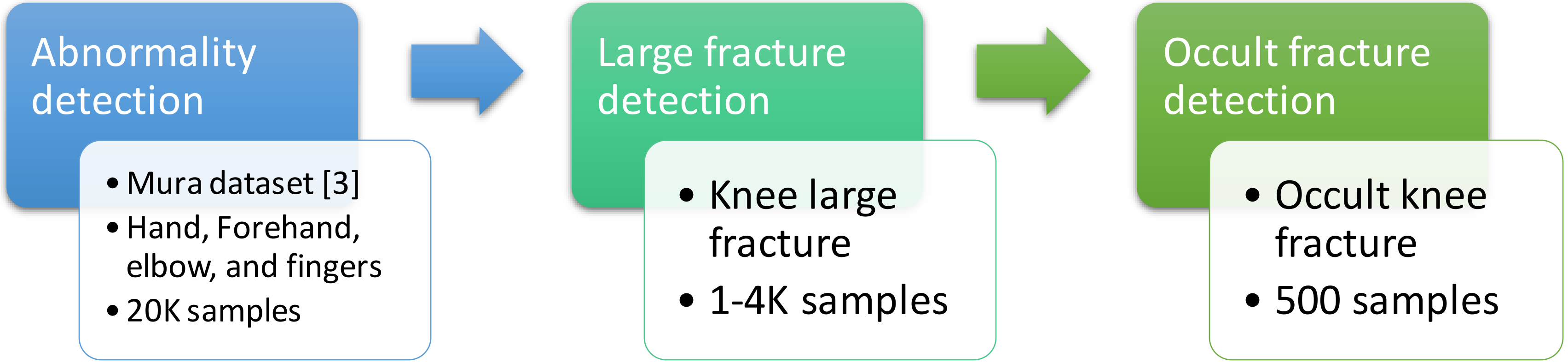
Method	PSNR	SSIM	MS-SSIM	FSIM
Bilinear	20.75	0.574	0.782	0.5320
Progressive Face SR [1]	22.67	0.687	0.908	0.6374
VDSR [2]	22.96	0.655	0.887	0.6103
MSG-CapsGAN	23.35	0.673	0.899	0.6371
Proposed Patch GAN	***23.64***	0.717	0.927	0.6788
Proposed VGG-Residual	23.53	***0.719***	***0.929***	***0.6918***

*Majdabadi, Mahdiyar Molahasani, and Seok-Bum Ko, “Capsule GAN for robust face super resolution,” *Springer Multimedia Tools and Applications*, 79, 31205-31218 (2020).

Occult Fracture Detection (in progress)

Multi-stage Transfer learning for dealing with limited available samples

- Using multi-stage transfer learning paradigm
- Start with more general problem
- Fine tuning the model on the smaller datasets



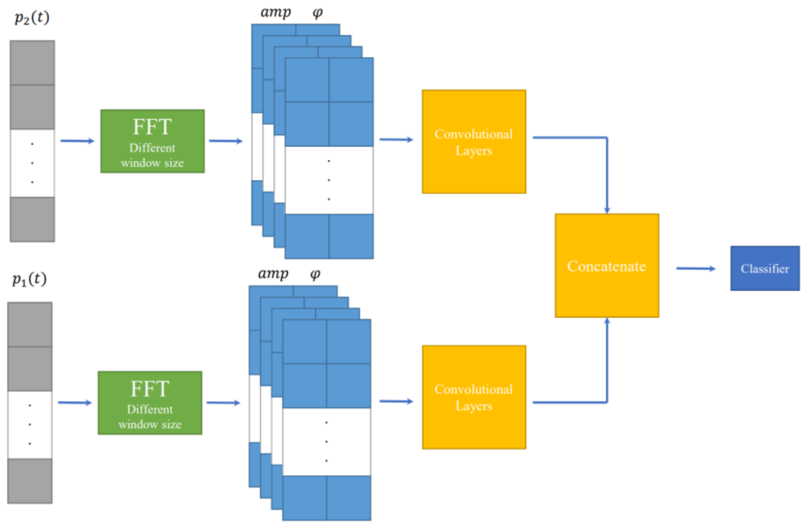
Proposed Workflow

Erosion Detection in Pipeline (in progress)

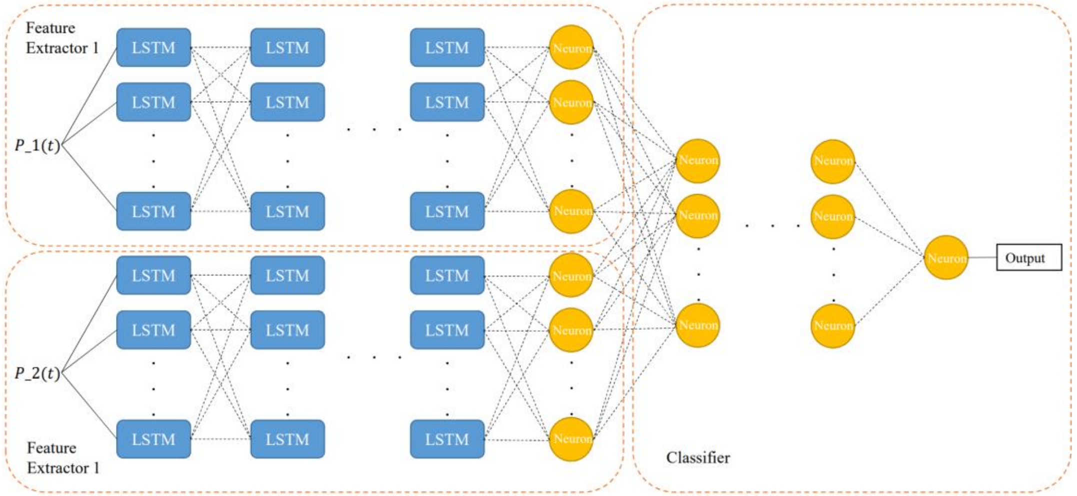
UNIVERSITY OF SASKATCHEWAN

Erosion Localization and Severity Detection Using Deep Learning in the Time and Frequency Domain

- Detect the existence of the erosion based on the peruse in the pipe $P(t)$
- Localizing the thinned area
- In both time and frequency domain



Frequency domain model

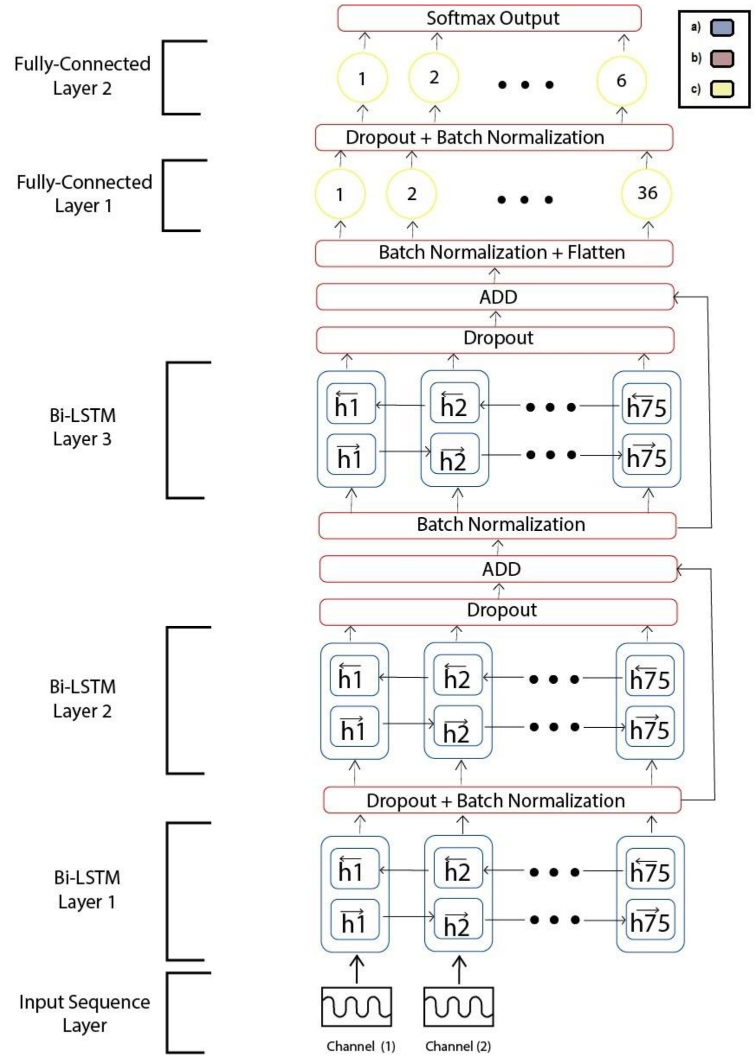


Time domain model

An LSTM-based interconnected architecture for the classification of grasp types using sEMG signals

Recently, studies show Deep Neural Networks could obtain surface electromyographic (sEMG) signal features in their internal architecture and use them directly over a classification task, avoiding all pre-processing steps and improving the obtained accuracy.

The current study proposes a deep architecture based on LSTM Networks for the classification of 6 grasp types as an end-to-end deep model approach, working with raw surface electromyographic signals.



An LSTM-based interconnected architecture for the classification of grasp types using sEMG signals

- Classification accuracy of 99.12% was obtained and compared with previous studies which use different machine learning techniques over the same dataset.

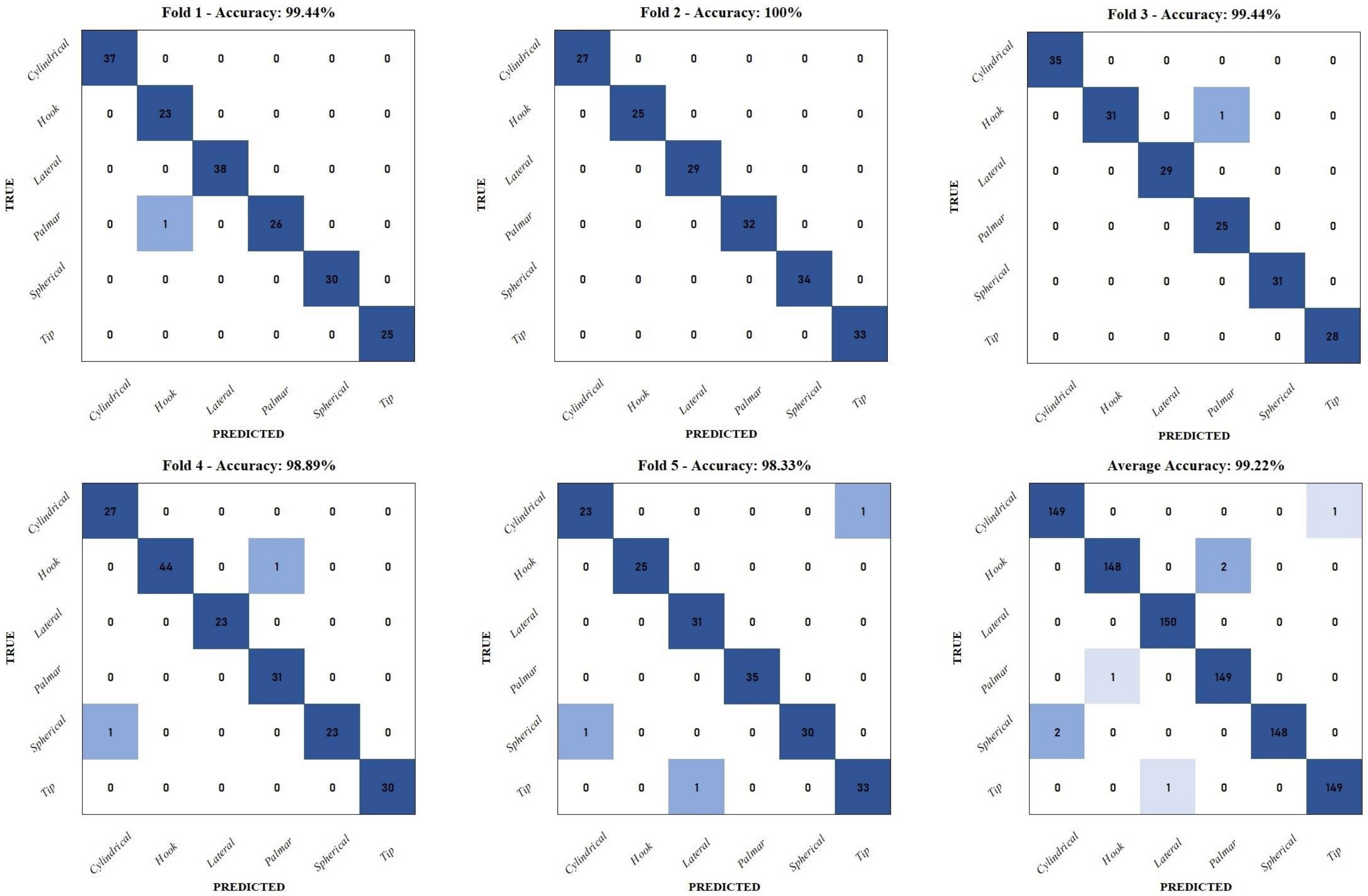
- Results obtained showed that our model's architecture improves previous results as well as provides a robust solution avoiding overfitting, with an F1-score higher than 99% for all grasp types.

Reference Study	Number of channels[1]	Feature extraction method[2]	Depth Neural Layers	Classification method	Accuracy (%)
Reference [18]	2	2D-CNN	6	Dueling Deep Q-learning	83.5
Reference [19]	1	1D-CNN	4	SoftMax classifier	94.94
Reference [25]	1	1D-CNN/LSTM	5	SoftMax classifier	98.8
The proposed	2	2D-LSTM	4	SoftMax classifier	99.13

[1] The dataset contains 2 channels information.
[2] The denomination "#D" refers to the dimension used at the indicated layers.

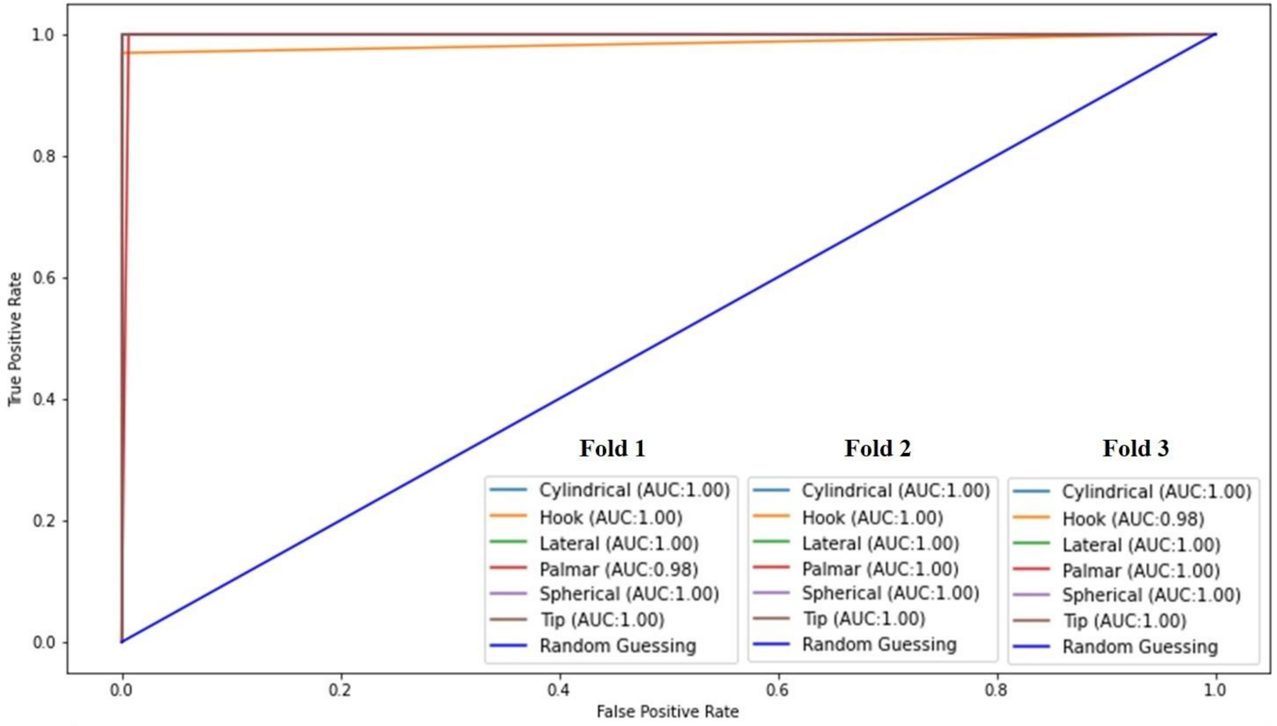
An LSTM-based interconnected architecture for the classification of grasp types using sEMG signals

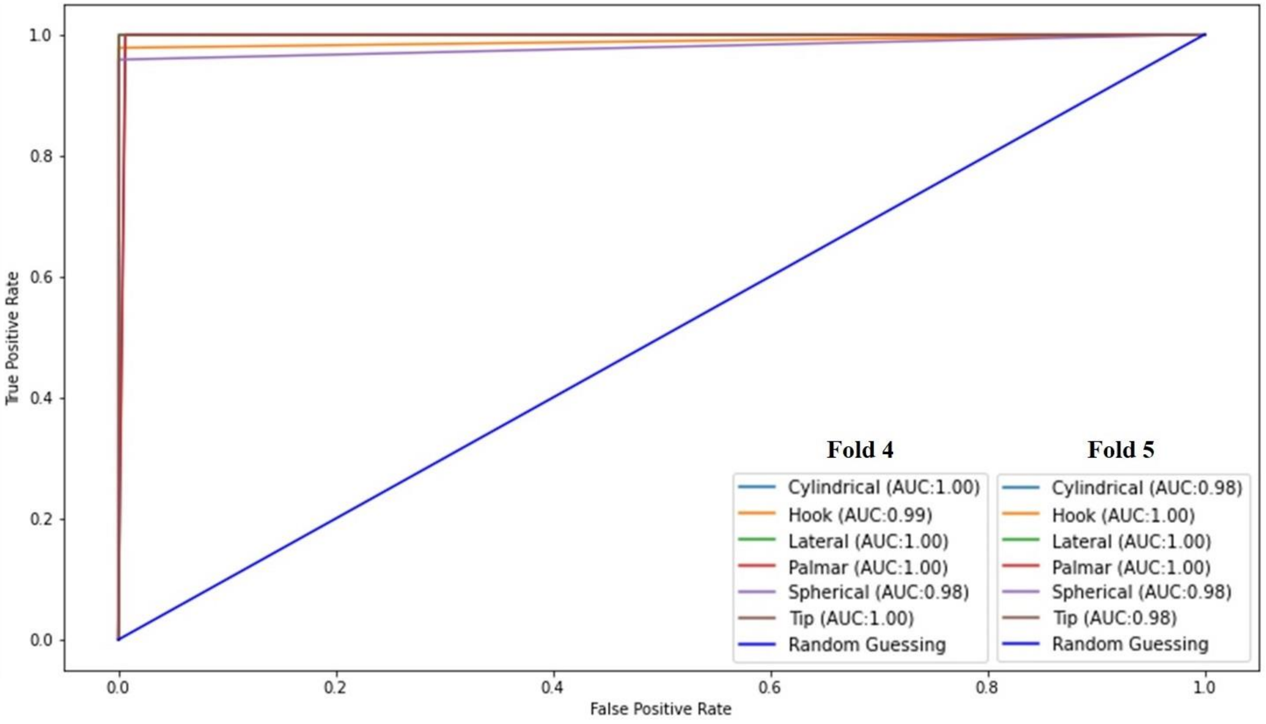
- The proposed model was cross-validated over 5 subjects and the results show an improvement of the average accuracy presented in previous studies by at least 0.33%.
- Confusion matrixes are presented to show the low percentage of errors obtained with our model.



An LSTM-based interconnected architecture for the classification of grasp types using sEMG signals

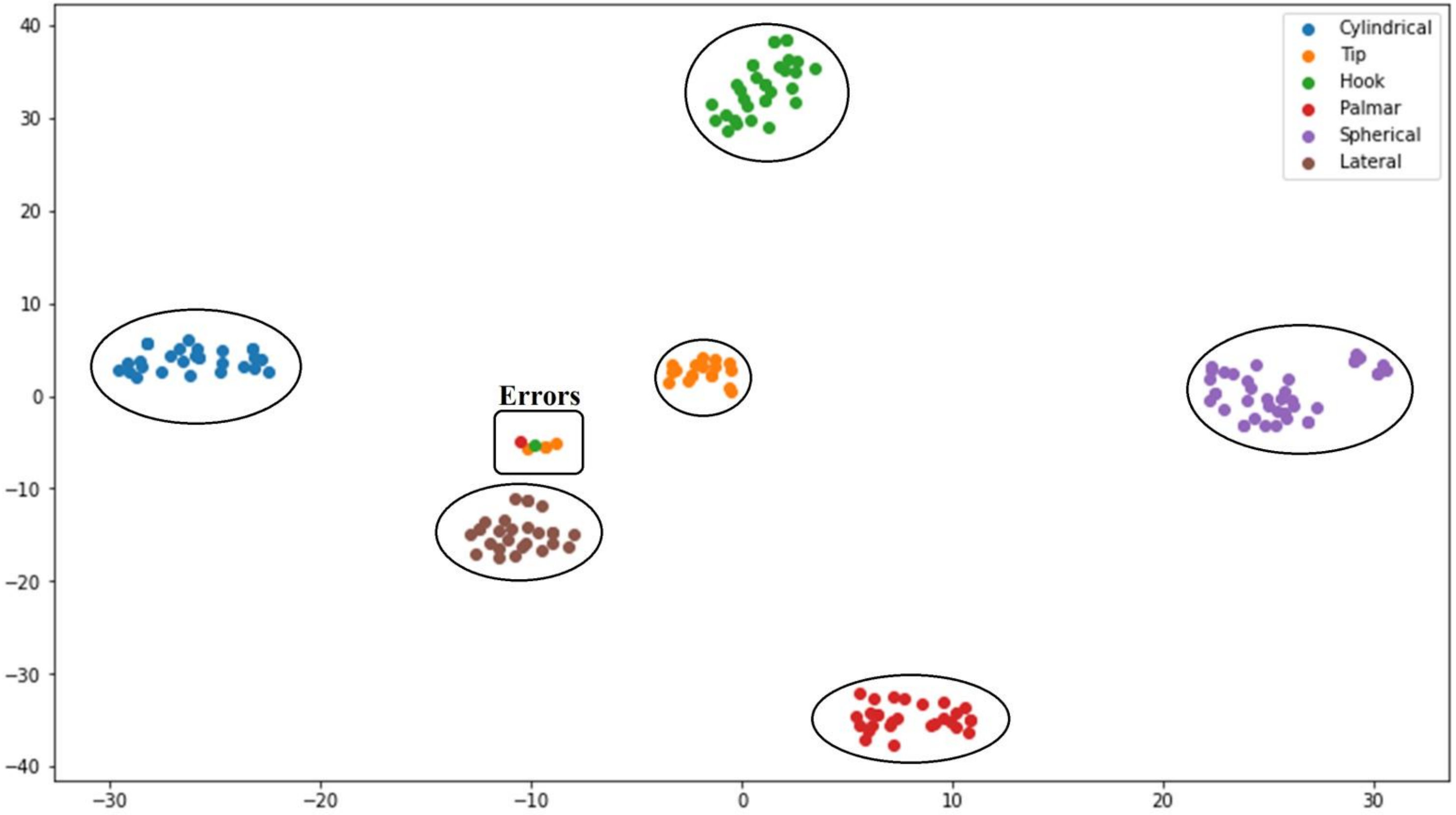
- Receiver operating characteristic (ROC) curves were obtained for each training fold to better visualize the optimization of sensitivity versus specificity for our implementation.





An LSTM-based interconnected architecture for the classification of grasp types using sEMG signals

- A graphical representation in two dimensions (t-DSNE) presents errors for grasp classification related to Hook, Tip, and Palmar grasps. Yet, most of the samples are well grouped which also conveys to the high accuracy obtained by our model.

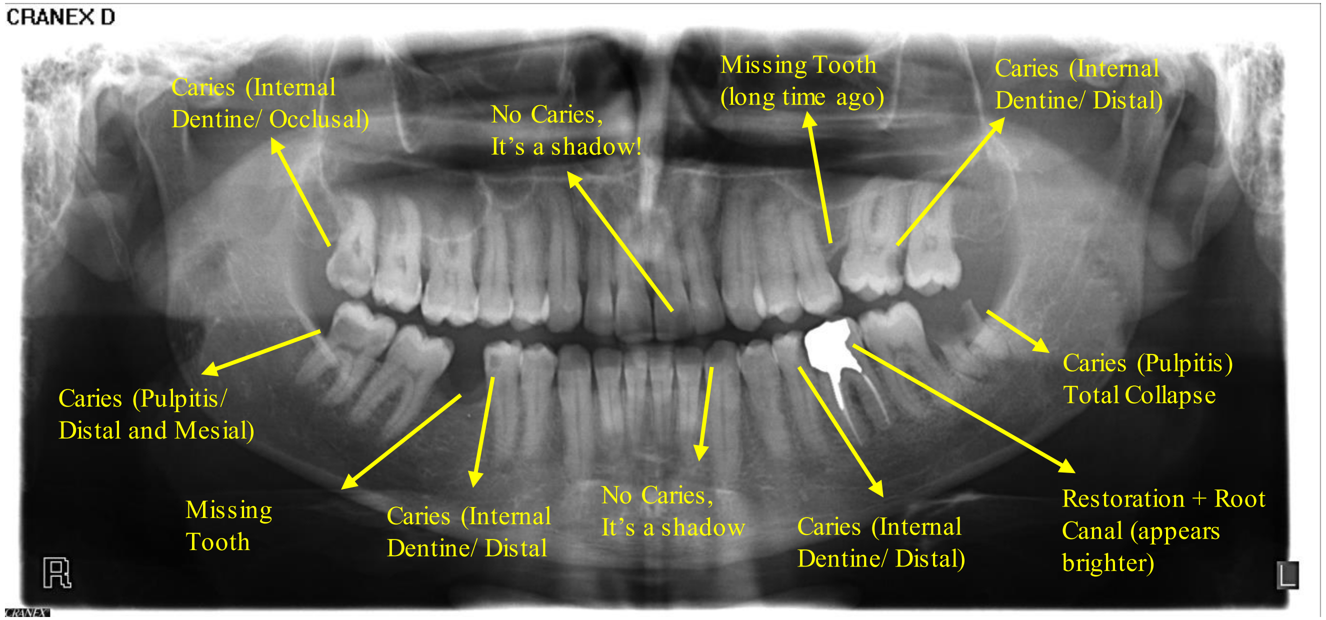


Teeth Extraction

UNIVERSITY OF SASKATCHEWAN

Material

- Dental Panoramic x-ray images
- Too noisy, has shadows, and interpretation is more sophisticated in comparison with Bitewing or Periapical images
- Covers whole mouth with relatively low radiation
- Low cost, widely used, reveals dental caries and jaw-bone fractions



Problem Description

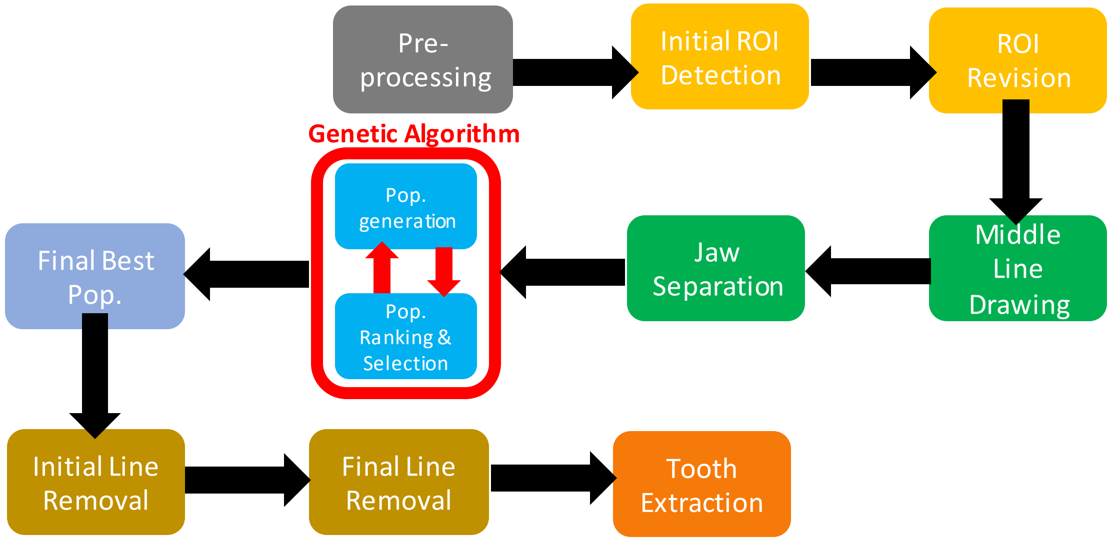
- Panoramic Images are used for various tasks; dental caries classification, lesion detection, human identification, etc.
- In this research, we focus on tooth decay detection. It means to find teeth which appeared to have dark spots on them.
- Dentist/radiologist must investigate many x-ray images everyday, hence the chance of wrong detection (false positive mainly) increases throughout the day. Also, this process causes eyestrain for them.
- True detection of caries in Panoramic images is challenging. Even radiologists working for years can not detect tooth decay from image shadows through all the samples.
- Tooth extraction/segmentation is an important part required by most of the applications, such as dental caries detection.

Teeth Extraction

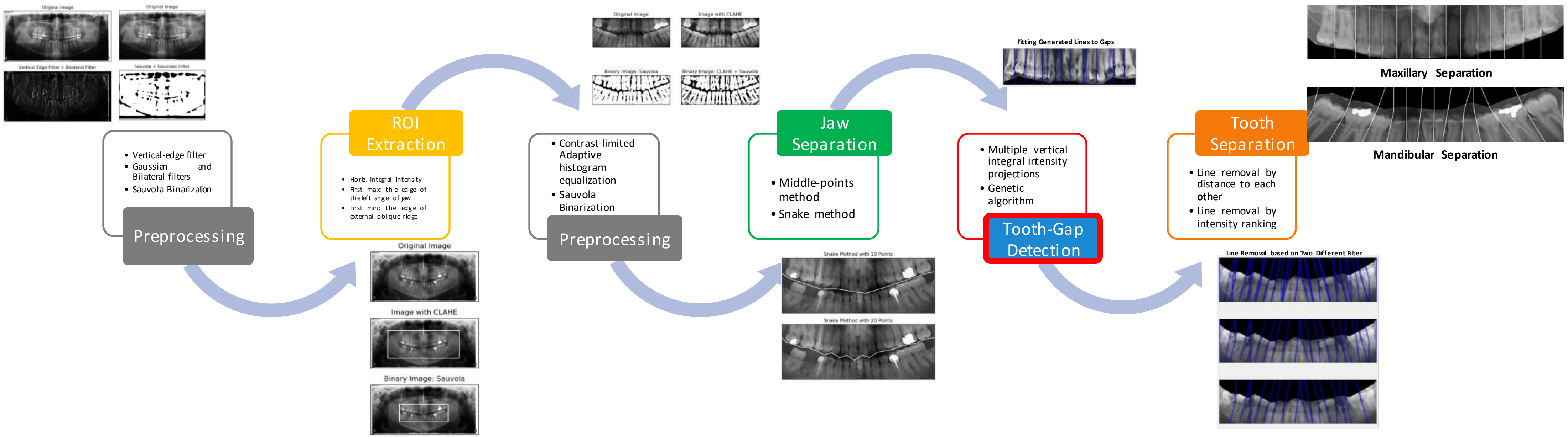
UNIVERSITY OF SASKATCHEWAN

Diagram of the Process

- **Preprocessing:** Contrast Limited Adaptive Histogram Equalization (CLAHE)/ Sauvola Binarization/ Blurring Filters
- **ROI Extraction:** Edge Filters/ Checkpoint detection
- **Jaw Separation:** CLAHE + Sauvola/ Starting point + Snake through the gap
- **Tooth Isolation:** Vertical Intensity Projection/ Genetic Alg./ Line Removal



Detailed Diagram and Algorithms



Teeth Extraction

University of Saskatchewan

Results

- Enhancements made on cost function and line removal method to improve the accuracy of mandible tooth extraction
- This study is the first to implement tooth extraction on Panoramic images and report the results. Panoramic images lack explicit boundaries between segments, include unwanted parts, have shadows; but the acquired accuracy is in the line of previous works carried out on easier image types.
- The overall accuracy is 77.56%; 81.44% for maxillary and 73.67% for mandibular teeth

Comparison between proposed method and other teeth extraction research works

	Image type	Algorithm	Correct Upper	Correct Lower
Abdel-Mottaleb et al. [1]	Bitewing	Integral Intensity Method	169/195 - 85%	149/181 - 81%
Olberg and Goodwin [2]	Bitewing	Path-based Method	300/336 - 89.3%	270/306 - 88.2%
Nomir et al. [3]	Bitewing	Integral Intensity Method	329/391 - 84%	293/361 - 81%
Al-Sherif [4]	Bitewing	Energy-based Method	1604/1833 - 87.5%	1422/1692 - 84%
Ehsani Rad et al. [5]	Periapical	Integral Intensity Method	Overall: 90.83%	
Proposed Study	Panoramic	Genetic-based Method	474/582 - 81.44%	428/581 - 73.67%

[1] M. Abdel-Mottaleb, O. Nomir, D. E. Nassar, G. Fahmy, and H. H. Ammar, “Challenges of developing an automated dental identification system,” in 2003 46th Midwest Symposium on Circuits and Systems, vol. 1. IEEE, 2003, pp. 411–414.

[2] J.-V. Ølberg and M. Goodwin, “Automated dental identification with lowest cost path-based teeth and jaw separation,” Scandinavian Journal of Forensic Science, vol. 22, no. 2, pp. 44–56, 2016.

[3] O. Nomir and M. Abdel-Mottaleb, “A system for human identification from X-ray dental radiographs,” Pattern Recognition, vol. 38, no. 8, pp. 1295–1305, 2005.

[4] N. Al-Sherif, G. Guo, and H. H. Ammar, “A new approach to teeth segmentation,” Proceedings - 2012 IEEE International Symposium on Multimedia, ISM 2012, no. 09, pp. 145 148, 2012.

[5] A. E. Rad, M. S. M. Rahim, H. Kolivand, and A. Norouzi, “Automatic computer-aided caries detection from dental x-ray images using intelligent level set,” Multimedia Tools and Applications, vol. 77, no. 21, pp. 28 843–28 862, 2018.

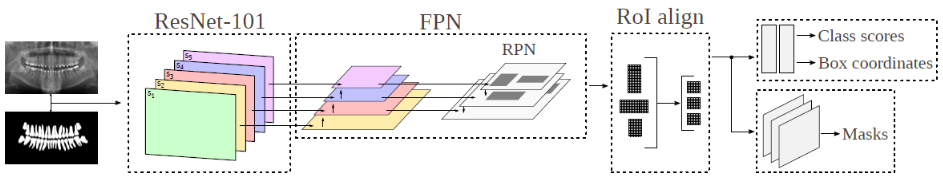
Related paper:

A. Haghanifar, M. M. Majdabadi and S. -B. Ko, “Automated Teeth Extraction from Dental Panoramic X-Ray Images using Genetic Algorithm,” *IEEE International Symposium on Circuits and Systems*, Sevilla, 2020, pp. 1-5, doi: 10.1109/ISCAS45731.2020.9180937.

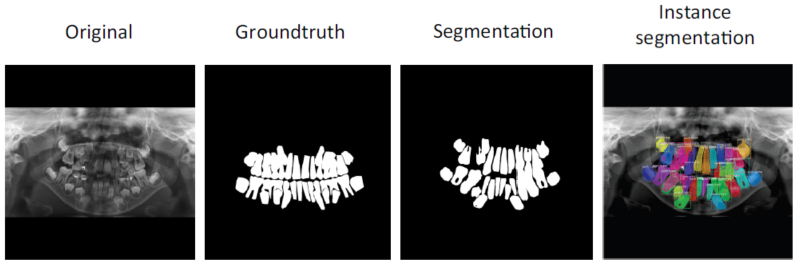
Deep Learning for Teeth Segmentation

Deep Learning for Segmentation

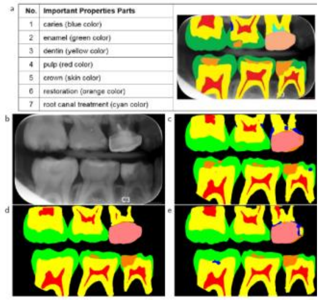
- To segment each tooth completely; finding the roots' boundary
- Mask R-CNN to obtain state-of-the-art results on Panoramic images [6]. Tooth-structure segmentation is well performed on Bitewing images using deep CNN [7].
- Annotation needed for train/validation/test; separating each tooth in each image



Architecture for teeth segmentation [1]



Sample result obtained by Mask R-CNN [1]

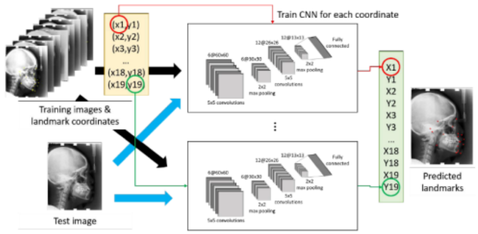


Tooth structure segmentation in Bitewing images [6]

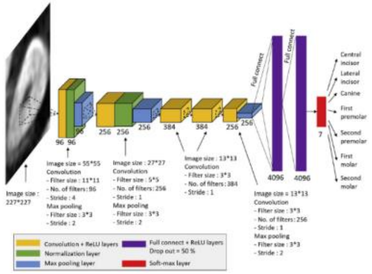
[6] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, and L. Oliveira, "Deep Instance Segmentation of Teeth in Panoramic X-Ray Images," in Proceedings - 31st Conference on Graphics, Patterns and Images, SIBGRAPI 2018, 2019, pp. 400–407.
[7] O. Ronneberger, P. Fischer, and T. Brox, "Dental X-ray Image segmentation using a U-shaped Deep convolutional network," pp. 1–13, 2015.

Deep Learning for Automating X-ray Image Analysis

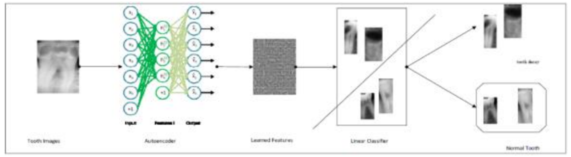
- CNNs are used for different x-ray image tasks: landmark detection in Cephalometric images [8], finding lesions [9] and osteoporosis [10] in Panoramic images, tooth classification in CBCT images [11], diagnosing periodontally compromised teeth in Periapical images [12]
- Few works have been conducted on Panoramic images for classifying tooth decays; such as [13]



Landmark detection in Cephalometric images [3]



Tooth classification in CBCT images [6]



Dental caries detection using separated image of single tooth [8]

[8] H. Lee, M. Park, and J. Kim, "Cephalometric landmark detection in dental x-ray images using convolutional neural networks," Med. Imaging 2017 Comput. Diagnosis, vol. 10134, p. 101341W, 2017.
[9] R. G. Birdal, E. Gumus, A. Sertbas, and I. S. Birdal, "Automated lesion detection in panoramic dental radiographs," *Oral Radiol.*, vol. 32, no. 2, pp. 111–118, 2016.
[10] J. S. Lee, S. Adhikari, L. Liu, H. G. Jeong, H. Kim, and S. J. Yoon, "Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: A preliminary study," *Dentomaxillofacial Radiol.*, vol. 48, no. 1, 2019.
[11] Y. Miki *et al.*, "Classification of teeth in cone-beam CT using deep convolutional neural network," *Comput. Biol. Med.*, vol. 80, no. September 2016, pp. 24–29, 2017.
[12] J. H. Lee, D. H. Kim, S. N. Jeong, and S. H. Choi, "Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm," *J. Dent.*, vol. 77, no. 2, pp. 106–111, 2018.
[13] R. Ben Ali, R. Ejbali, and M. Zaied, "Detection and Classification of Dental Caries in X-ray Images Using Deep Neural Networks," *ICSEA 2016 Elev. Int. Conf. Softw. Eng. Adv.*, no. c, pp. 223–227, 2016.

Deep Learning for Dental Caries Classification: PaXNet

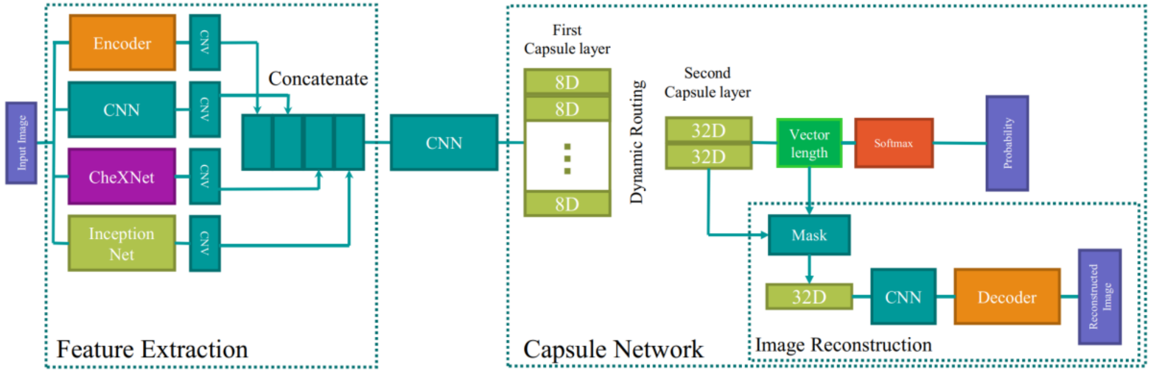
UNIVERSITY OF SASKATCHEWAN

Material

- High-resolution Panoramic x-ray images collected from two main sources: UESB dataset [14] and from a local dentistry clinic in Iran.
- UESB dataset has teeth mask annotations along with it. For our dataset, we used previously proposed genetic-based teeth extraction algorithm to isolate teeth in each image.
- Labels are provided by a radiologist. Caries are classified into two categories: mild and severe. Each tooth gets a label, and segmentation masks are not available yet, due to its time-consumption and limited availability of radiologist/dentist.

Proposed Model

- This research is the first one to perform dental caries detection on Panoramic x-rays. All previous studies have used less challenging x-ray types, like Bitewing or Periapical.
- Capsule network-based architecture is used for the first time for dental caries diagnosis.
- Feature extraction module is constructed by a voting system between different architectures. CheXNet [15] is applied for the first time in dental disease detection tasks based on x-rays.



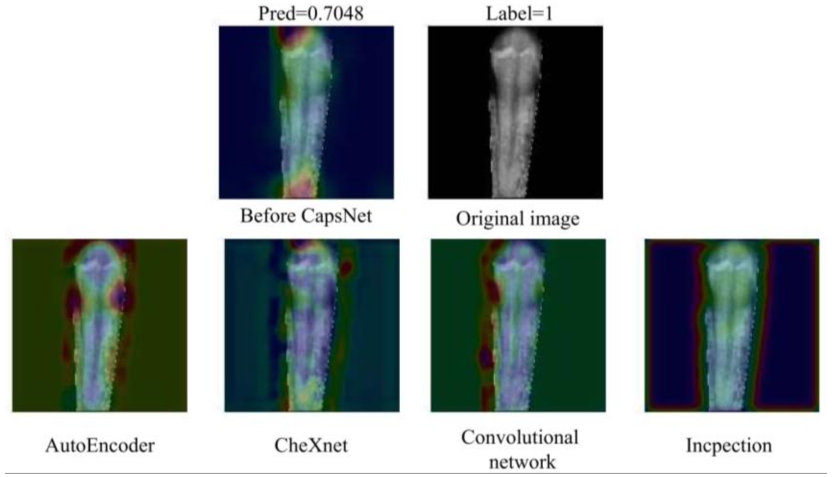
Architecture of PaXNet

[14] G. Silva, L. Oliveira, and M. Pithon, "Automatic segmenting teeth in x-ray images: Trends, a novel data set, benchmarking and future perspectives," Expert Systems with Applications, vol. 107, pp. 15–31, 2018.

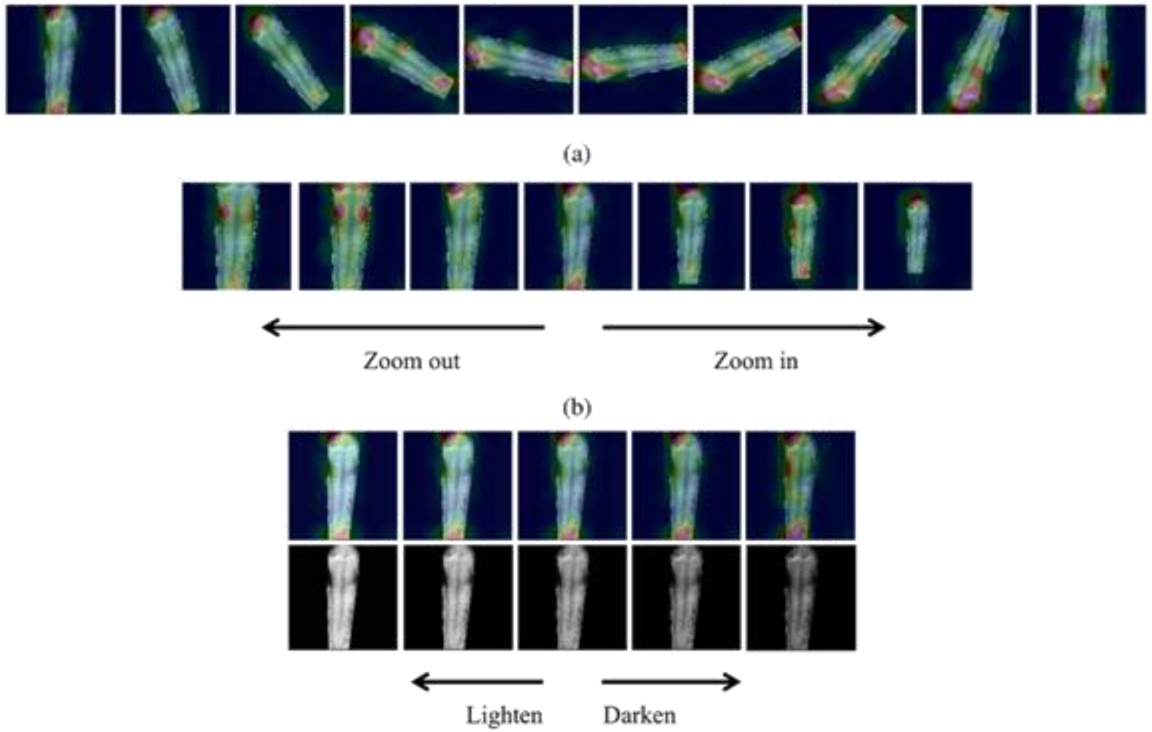
[15] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya et al., "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," arXiv preprint arXiv:1711.05225, 2017.

Results

- Gradient-weighted Class Activation Mapping (Grad-CAM) is used to highlight extracted features of each extractor models.
- Re-sampling technique is used to deal with class imbalance
- PaXNet model illustrated robustness activation map generation when applying linear transformation to any input image
- PaXNet model achieved an accuracy of 86.04% in a test-set of 368 teeth. Recall scores for classification of mild and severe caries were 69.44% and 90.52%, respectively.



Sample Result



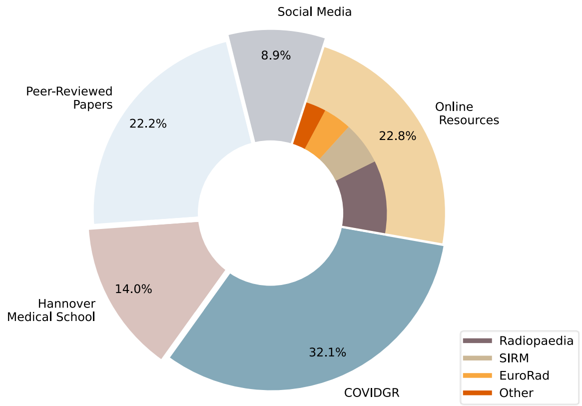
Demonstration of the robustness of the system

A. Haghanifar, M. M. Majdabadi, S. Haghnifar, Y. Choi and S. B. Ko, “PaXNet: Tooth segmentation and dental caries detection in panoramic X-ray using ensemble transfer learning and capsule classifier,” *Springer Multimedia and Applications*, Vol. 82, Iss. 18, pp. 27659-27679

Material & Preprocessing

- Frontal chest x-ray radiographs (CXRs) with PA or AP projections
- Collected from 10 different publicly available sources
- One of the largest collections of CXRs from patients with COVID-19: https://github.com/armiro/COVID-CXNet
- Including **1326** images, as of May 2021. Images have different sizes and are in different formats.
- Normal/non-COVID pneumonia images are collected from different sources; mainly NIH CXR-14 dataset [1]. 5,000 normal and 4,600 non-COVID pneumonia x-rays are used.
- Different histogram equalization (HE) preprocessing methods are used: HE, AHE, CLAHE, BEASF
- BEASF algorithm [2] is implemented in python for first time: https://github.com/armiro/COVID-CXNet/blob/master/BEASF.py

Randomly selected images from different sources



Dataset distribution

[1] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2097–2106.

[2] E. F. Arriaga-Garcia, R. E. Sanchez-Yanez, and M. Garcia-Hernandez, "Image enhancement using bi-histogram equalization with adaptive sigmoid functions," in 2014 International Conference on Electronics, Communications and Computers (CONIELECOMP). IEEE, 2014, pp. 28–34.

Problem Description

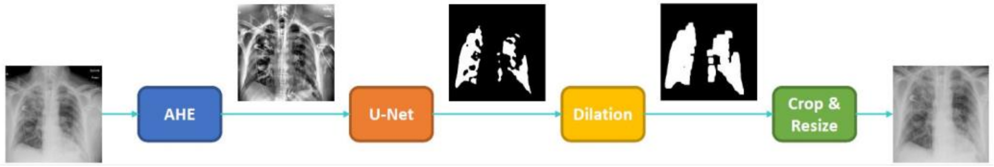
- The standard diagnosis method for detection of COVID-19 is by reverse transcription polymerase chain reaction (rRT-PCR). PCR testing has certain drawbacks; it needs special test-kits, and the results are generally available within hours to days.
- Hopefully, the disease could be assessed by detecting clinical features as well as imaging features of pneumonia [3].
- Since radiologists are visiting many patients every day and detection takes significant time, detection error rate may increase, ending up having false negatives which costs a lot to the patient and the medical staff.
- In this research study, we investigate the possibility of using deep learning-based automatic image classification system to detect COVID-19 pneumonia in CXRs.
- Although CT scans have proven to be more efficient revealing detailed features of the chest, there are less widely available and affordable than CXRs. Besides, patient's clinical situations often does not allow a CT scan.

[3] Y. H. Jin et al., “A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-ncov) infected pneumonia (standard version)," Military Medical Research, vol. 7, no. 1, p. 4, 2020.

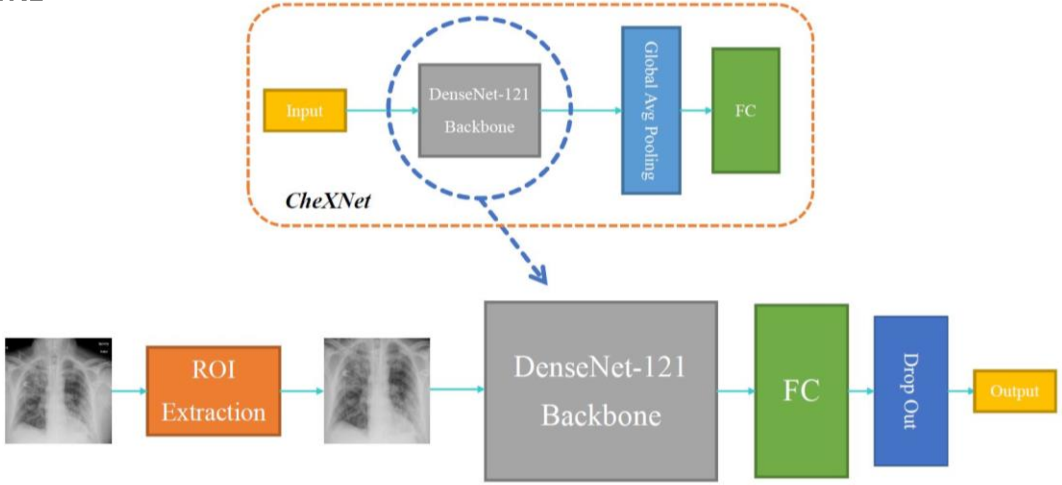
UNIVERSITY OF SASKATCHEWAN

Developed Method

- A U-Net based semantic segmentation module to extract lung from the body and background
- Trained on two datasets of manually segmented lung masks [4][5]
- Transfer learning from CheXNet [6], being trained on a la dataset of frontal CXRs to classify different lung diseases
- DenseNet-121 as the backbone



The segmentation approach based on the U-Net



COVID-CXNet model architecture based on the DenseNet-121 feature extractor as the backbone

[4] S. Stirenko, Y. Kochura, O. Alienin, O. Rokovyi, Y. Gordienko, P. Gang, and W. Zeng, “Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation,” in 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO). IEEE, 2018, pp. 422–428.
[5] S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. J. McDonald, “Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration,” IEEE Transactions on Medical Imaging, vol. 33, no. 2, pp. 577–590, 2013.
[6] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya et al., “Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning,” arXiv preprint arXiv:1711.05225, 2017.

COVID-CXNet: COVID-19 Diagnosis in CXRs using Deep Learning

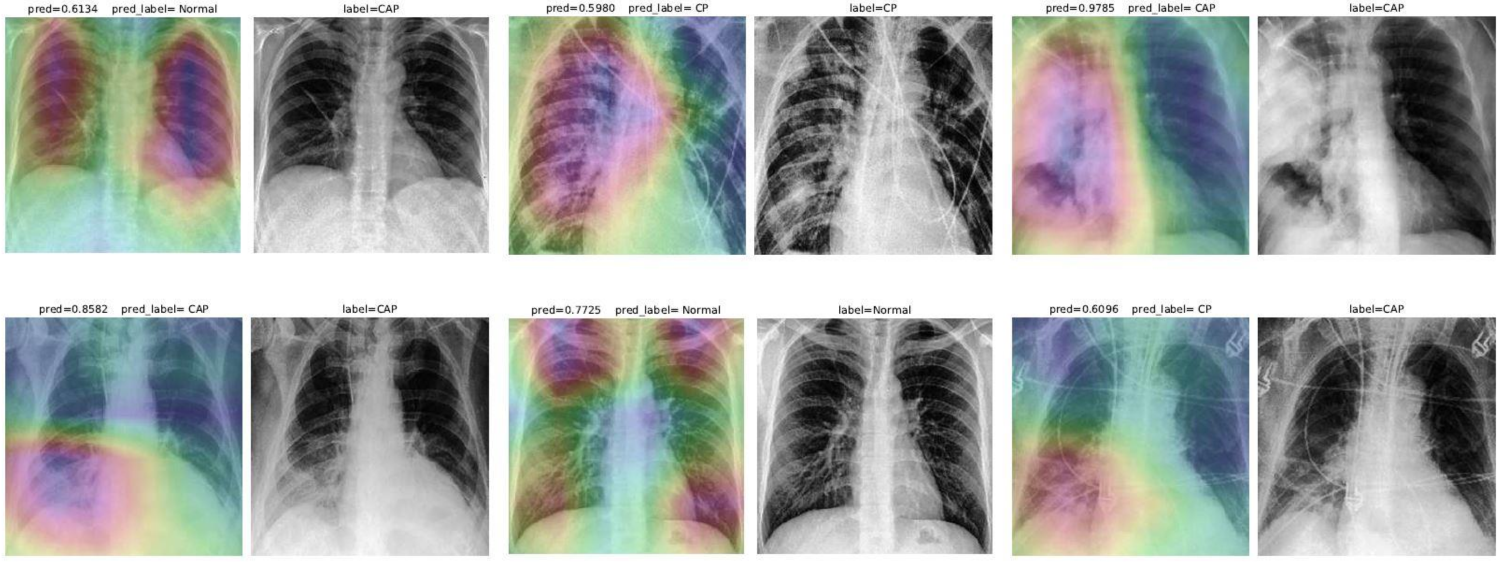
UNIVERSITY OF SASKATCHEWAN

Results

- Grad-CAM to visualize model output results, and to prevent “right decision with wrong reason” phenomenon commonly encountered when datasets are small
- Lung segmentation to enhance model accuracy and robustness to recurring text/signs
- Label smoothing to add uncertainty to the labelling
- A hierarchical approach to improve scores in discrimination between normal and non-COVID pneumonia classes

Final confusion matrix of COVID-CXNet

COVID-CXNet		**Predicted**		
		Normal	*CAP*	*CP*
Actual	*Normal*	689	32	3
	CAP	143	524	5
	CP	3	11	130

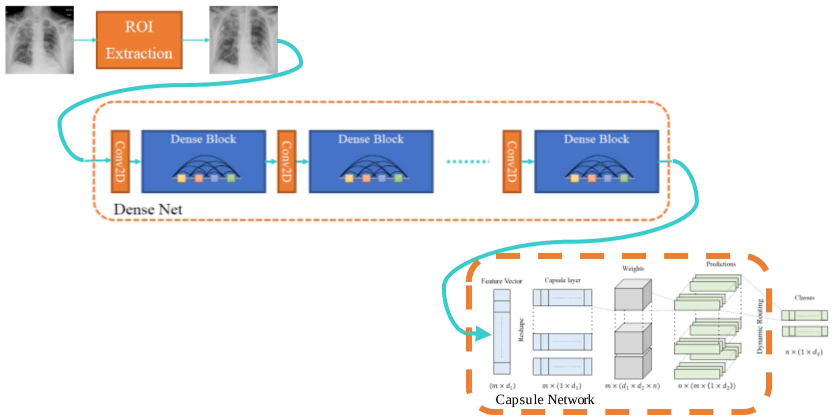


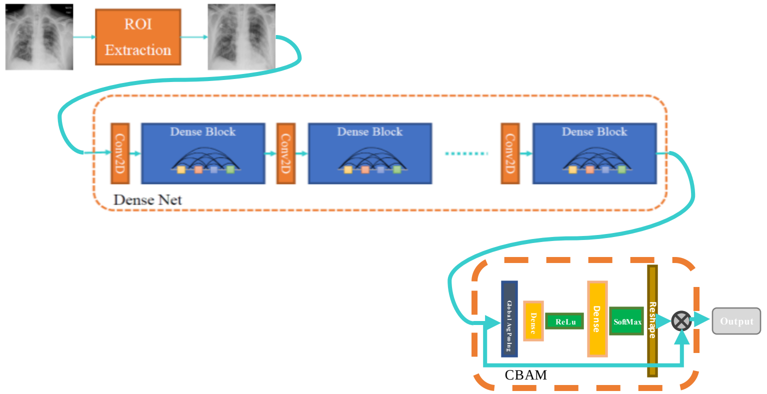
COVID-CXNet multiclass classification visualization results

A. Haghanifar, M. M. Majdabadi, Y. Choi, S. Deivalakshmi, and S. B. Ko, “Covid-CXNet: Detecting covid-19 in frontal chest x-ray images using deep learning,” Springer Multimedia Tools and Applications, 81, 30615-30645 (2022).

Enhancements on Model Architecture (ongoing)

- Since COVID-CXNet is a modification of CheXNet as the main model, there is room for increasing performance scores by benefiting from different methods as the classifier block instead of simple dense layers
- Capsule Network is a good option, which has previously demonstrated to be helpful in PaXNet project.
- Attention Modules are relatively new in this field. While channel attention modules are more popular in this area, there are convolutional attention modules as well. Convolutional Block Attention Module (CBAM)[7] is used for both channel and spatial attention dimensions and is general so that could be added to any CNN architecture. A huge motivation in terms of adding attention to the network is [8] where the authors have proved that using their developed uncertainty attention block, pneumonia classification in chest x-rays improved roughly 9%



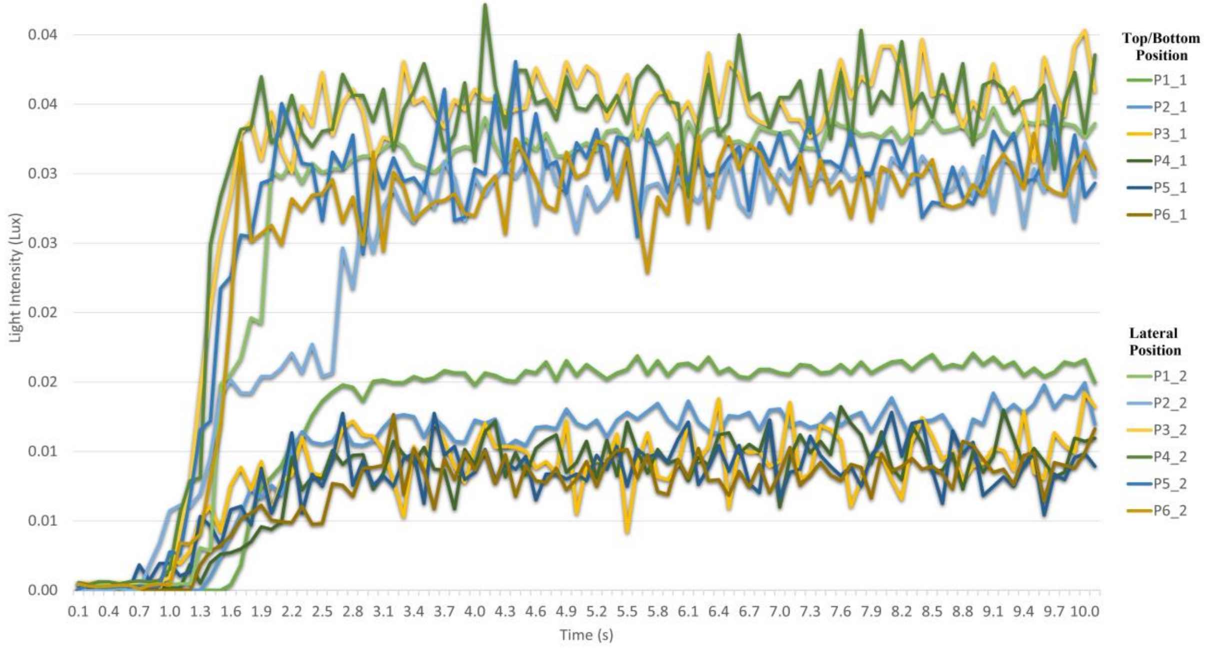


[7] S. Woo, J. Park, J Lee, and I. Kweon, "Cbam: convolutional block attention module," Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19.
[8] C. Wang, F. Su, T. Lee, Y. Tsai, and J. Chiang, "CUAB: Convolutional Uncertainty Attention Block Enhanced the Chest X-ray Image Analysis," arXiv preprint arXiv:2105.01840, 2021.

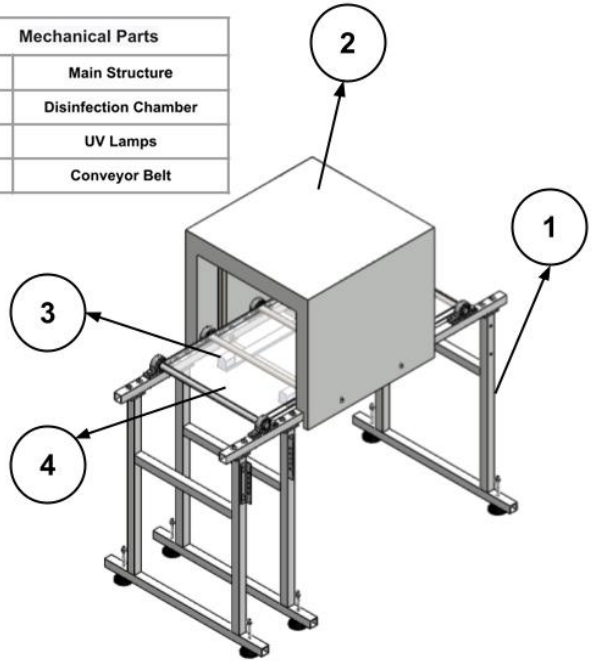
UV-SAFE: Advancing Market Product Disinfection with Automated UV-Technology Device

- An automated device for viruses and bacterial disinfection for use in market products published in proceedings of ETCM 2023, IEEE.
- Results of an experimental study, conducted with regular and irregular objects.
- The light incidence in the object was validated by experimental tests.
- The disinfection process takes 10 seconds per product, achieving a log-reduction dose.

Lamps configurations:



Mechanical Parts	
1	Main Structure
2	Disinfection Chamber
3	UV Lamps
4	Conveyor Belt



UV-Safe prototype

Academic Article Information Extraction

UNIVERSITY OF SASKATCHEWAN

Document Structure Extraction

- Automatically identifying and extracting areas of interest in academic works can help with further information extraction and analysis
- Datasets exist for training models on this task, but they often only annotate a handful of objects
- We create an expanding and open-source Densely Annotated Dataset (DAD) for academic article semantic segmentation
- Dataset publicly available on GitHub
  - https://github.com/LivingSkyTechnologies/Dense_Article_Dataset_DAD
- Presented dataset in 2020 Workshop on Scientific Document Analysis

L. Markewich, Y. Xing, H. Zhang, Z. Jiang, N. Lambert-Shirzad, R. Lee, Z. Li and S. Ko, "Document Structure Extraction: An Exploratory Study," *2020 Fourth International Workshop on Scientific Document Analysis (SCIDOCA2020)*, Nov. 2020.

Mitacs Living Sky Technologies

Document Structure Extraction

- DAD contains annotations for 42 document objects across 450 academic journals
- Other popular datasets like PubLayNet and DocBank contain annotations for 5 and 11 document objects respectively

Front Matters			Body Matters			Back Matters		
Class	Pages	Instances	Class	Pages	Instances	Class	Pages	Instances
abstract	455	466	abbreviation	49	49	acknowledgment	51	63
affiliation	477	449	caption	2754	4378	additional file	24	24
article history	382	385	citation	3742	37114	appendice	174	224
author name	449	449	code	49	49	author bio	115	147
contact info	537	539	core text	5000	13517	author contribution	121	126
copyright	480	639	figure	1940	2851	availability of data	138	138
date	446	473	index	704	3209	conflict int	279	282
doi	448	473	list	555	812	conset of publication	3	3
funding info	291	366	math formula	832	8634	corresponding author	38	39
highlights	31	31	nomenclature	18	18	editor	80	80
journal	442	677	section heading	2719	4342	ethics	7	7
keywords	434	435	subheading	2367	4212	note	354	1451
math subject class	3	3	table	1109	1459	publisher note	56	56
publisher	439	442				reference	968	1451
title	450	450				URLs to supplementary	132	150

Summary of annotated objects

Journal	#Articles
Ampersand	50
International Journal of Engineering Science	26
Results in Engineering	24
Language Testing in Asia	50
Research in Engineering Design	34
Protection and Control of Modern Power Systems	16
Sage Open	50
Advances in Mechanical Engineering	50
Asia & the Pacific Policy Studies	26
Brain and Behaviour	24
Advanced Science	50
Transactions on Computers	26
Transactions on Mobile Computing	18
Transactions on Affective Computing	6
Total	**450**

Summary of sourced academic articles

Academic Article Information Extraction

University of Saskatchewan

Document Structure Extraction

- We develop a new loss function for semantic segmentation, extracting the bounding boxes from inferenced segmentations and calculating the generalized intersection-over-union (gIOU) loss
- Additionally, we develop a custom weighting calculation that calculates class weights on every training batch, heavily weighting classes that are small/rarely appear
- Our methods result in a +1.99% F1 improvement with DeepLabV3+ on DAD
- DeepLabV3+ trained on DAD can be used for bootstrapped annotation, shows a 38% improvement in annotation speed
- Published in Springer IJDAR (2022)

Logan Markewich, Hao Zhang, Yubin Xing, Navid Lambert-Shirzad, Zhexin Jiang, Roy Ka-Wei Lee, Zhi Li, and Seok-Bum Ko. Segmentation for document layout analysis: not dead yet. International Journal on Document Analysis and Recognition (IJDAR), Jan 2022

	DAD							
	Gated-SCNN		FastFCN		DeepLabV3+		PSP-UNet	
Class	Pixel Acc.	mIOU	Pixel Acc.	mIOU	Pixel Acc.	mIOU	Pixel Acc.	mIOU
abstract	98.04%	71.02%	99.49%	94.39%	99.43%	96.22%	92.41%	86.93%
author name	98.43%	39.49%	99.93%	71.74%	99.66%	84.47%	95.08%	73.79%
background	83.91%	81.75%	94.68%	92.96%	97.48%	96.01%	94.97%	93.45%
caption	91.78%	34.38%	99.01%	72.88%	96.84%	89.92%	97.43%	82.69%
contact info	98.28%	39.88%	99.94%	55.74%	92.63%	65.39%	94.47%	41.78%
copyright	97.64%	38.54%	99.47%	70.81%	98.29%	89.76%	98.34%	75.47%
core text	78.78%	71.33%	94.50%	92.74%	96.44%	94.26%	93.69%	92.13%
doi	98.38%	30.71%	98.82%	47.08%	100.00%	84.79%	99.96%	56.58%
figure	84.47%	76.00%	97.34%	94.46%	98.60%	95.34%	97.69%	94.31%
funding info	70.31%	16.18%	85.71%	47.89%	70.86%	53.11%	74.60%	47.34%
keywords	96.24%	33.03%	99.80%	66.40%	96.22%	78.93%	92.79%	47.68%
list	43.70%	27.02%	95.57%	73.90%	72.14%	59.19%	77.94%	59.79%
math formula	76.34%	21.88%	97.61%	50.05%	81.89%	58.44%	90.71%	39.37%
note	98.55%	32.28%	99.88%	80.00%	73.92%	64.45%	74.61%	59.15%
publisher	100.00%	18.08%	100.00%	36.05%	93.33%	55.68%	93.33%	20.69%
reference	87.69%	78.20%	97.65%	93.69%	93.97%	90.47%	95.06%	93.23%
section heading	93.20%	31.35%	99.66%	57.27%	96.93%	79.25%	97.73%	58.95%
table	88.79%	69.78%	97.91%	91.56%	96.64%	93.36%	97.76%	90.78%
title	99.86%	54.62%	99.25%	82.14%	98.11%	90.80%	93.27%	86.54%
appendice*	35.04%	38.28%	96.43%	82.10%	46.34%	44.88%	33.69%	30.42%
author bio*	99.75%	61.67%	99.93%	94.67%	99.50%	95.42%	98.38%	95.14%
editor*	100.00%	10.15%	100.00%	33.31%	100.00%	57.57%	100.00%	19.87%
Average	87.24%	44.35%	**97.85%**	71.90%	90.87%	**78.08%**	90.18%	65.73%
F1	78.76%		95.46%		**96.26%**		93.60%	

Benchmark results on DAD w/ gIOU loss + loss weights

	PubLayNet							
	Gated-SCNN		FastFCN		DeepLabV3+		PSP-UNet	
Class	Pixel Acc.	mIOU	Pixel Acc.	mIOU	Pixel Acc.	mIOU	Pixel Acc.	mIOU
background	66.74%	64.64%	91.76%	89.62%	96.30%	95.44%	97.32%	96.04%
figure	67.79%	59.00%	98.84%	88.54%	99.17%	92.25%	98.34%	93.80%
list	48.84%	33.17%	94.87%	77.25%	96.85%	82.99%	94.78%	86.51%
table	95.45%	64.31%	99.34%	90.38%	99.38%	94.60%	98.90%	96.06%
text	71.37%	66.50%	93.58%	88.90%	95.88%	94.34%	97.28%	95.51%
title	65.25%	31.91%	98.53%	50.69%	99.01%	73.33%	99.05%	72.32%
Average	69.24%	54.72%	96.15%	80.90%	**97.76%**	88.83%	97.61%	**90.04%**
F1	73.08%		94.26%		97.11%		**97.66%**	

Benchmark results on PubLayNet w/ gIOU loss + loss weights

Mitacs Living Sky Technologies

Document Content Analysis

- Content analysis includes many subtasks
  - Sentiment Analysis
  - Named Entity Recognition
  - Topic Extraction
  - Keyword Extraction
  - **Relation Extraction (RE)**
- RE focuses on classifying relations between important nouns (named entities)

An example of relation extraction. Three entities are classified with specific relations.

Document Content Analysis

- Datasets exist for relation extraction, but contain limited categories for entities and relations

Dataset	# Entity Types	# Relation Types
CoNLL04 [74]	4	5
Re-TACRED [79]	17	40
DocRED [95]	6	96

- Models trained on these datasets will miss contextual information from the input

> Tesla has stopped building forward-facing radar sensors into its Model 3 sedans and Model Y SUVs in North America, after CEO Elon Musk publicly expressed a desire to rely on cameras to power the company's advanced driver assistance system, Autopilot. Tesla has been developing the vision-based version of Autopilot during the limited beta test of its "Full Self-Driving" software. But it's not completely done making sure Autopilot works without the radar sensors, though, as it is limiting or disabling some features on these vehicles for an indefinite amount of time. Autosteer — the Autopilot feature that can keep a Tesla centered in a lane, even around curves — will only be usable at 75 miles per hour and below. Tesla is also only making it available at an (unspecified) longer minimum following distance to any cars in front.

A state-of-the-art (SOTA) model only extracts background information.
No information about the actual content is identified.

REBEL
(Autopilot, developer, Tesla)
(Model 3, manufacturer, Tesla)
(Tesla, founded by, Elon Musk)
(Elon Musk, employer, Tesla)
(Autosteer, developer, Tesla)

Academic Article Information Extraction

UNIVERSITY OF SASKATCHEWAN

Document Content Analysis

- We create a new **D**escriptive **R**elation **D**ataset (DReD) for describing relations between general noun-phrases
  - 3,283 annotated paragraphs, 14,126 sentences
- Models trained on DReD can describe rich and contextual relations missed by current approaches
- Work is current in review with IEEE Transactions on Artificial Intelligence

Tesla has stopped building forward-facing radar sensors into its Model 3 sedans and Model Y SUVs in North America, after CEO Elon Musk publicly expressed a desire to rely on cameras to power the company's advanced driver assistance system, Autopilot. Tesla has been developing the vision-based version of Autopilot during the limited beta test of its "Full Self-Driving" software. But it's not completely done making sure Autopilot works without the radar sensors, though, as it is limiting or disabling some features on these vehicles for an indefinite amount of time. Autosteer — the Autopilot feature that can keep a Tesla centered in a lane, even around curves — will only be usable at 75 miles per hour and below. Tesla is also only making it available at an (unspecified) longer minimum following distance to any cars in front.

REBEL	$T5_{DReD}$	
(Autopilot, developer, Tesla)	Autopilot, which keeps a Tesla centered in a lane, is limited by Tesla's vision technology.	Tesla stopped building forward-facing radar sensors.
(Model 3, manufacturer, Tesla)	Tesla Model 3 has stopped building forward-facing radar sensors.	Cameras are now standard on the Tesla Model 3 sedan.
(Tesla, founded by, Elon Musk)	Tesla has stopped building forward-facing radar sensors after CEO Elon Musk expressed a desire to rely on cameras.	Tesla sedans now have cameras instead of radar sensors.
(Elon Musk, employer, Tesla)	Elon Musk is CEO of Tesla.	Tesla sedans no longer have forward-facing radar sensors.
(Autosteer, developer, Tesla)	Autosteer is a feature of Tesla that keeps it centered in a lane.	
SOTA model results	*Same relations described by a model trained on DReD*	*Relations found between general noun-phrases*

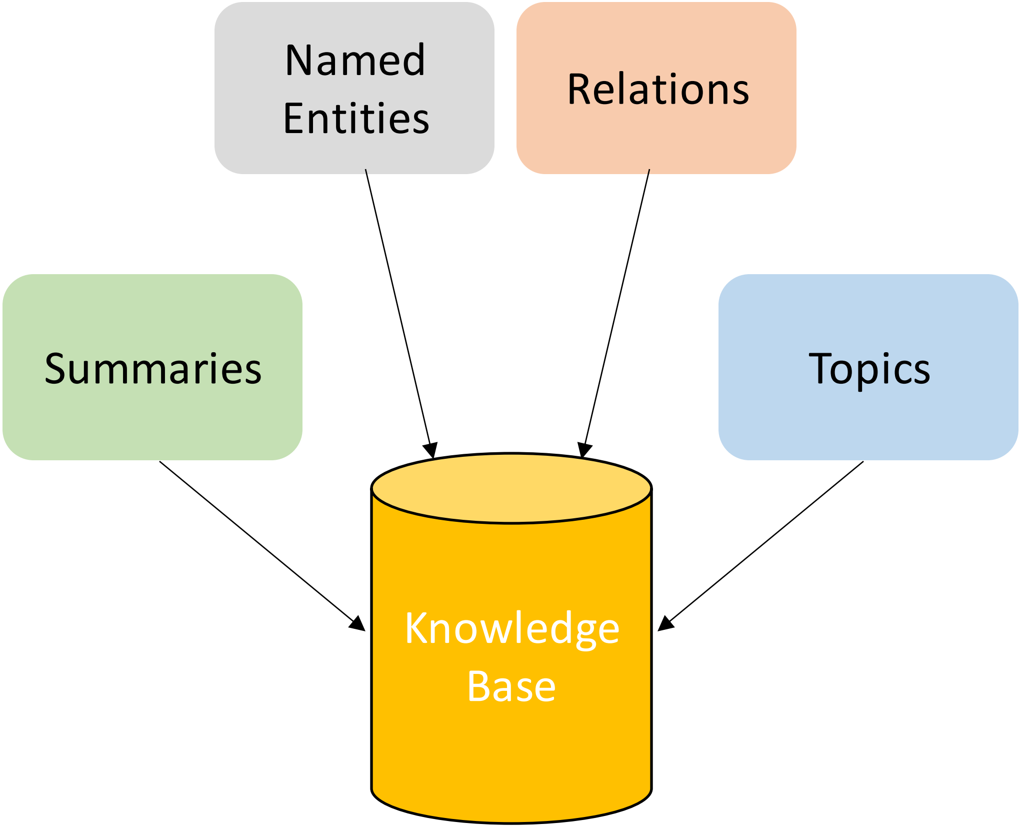
Mitacs Living Sky Technologies

Academic Article Information Extraction

UNIVERSITY OF SASKATCHEWAN

Document Content Analysis (Future)

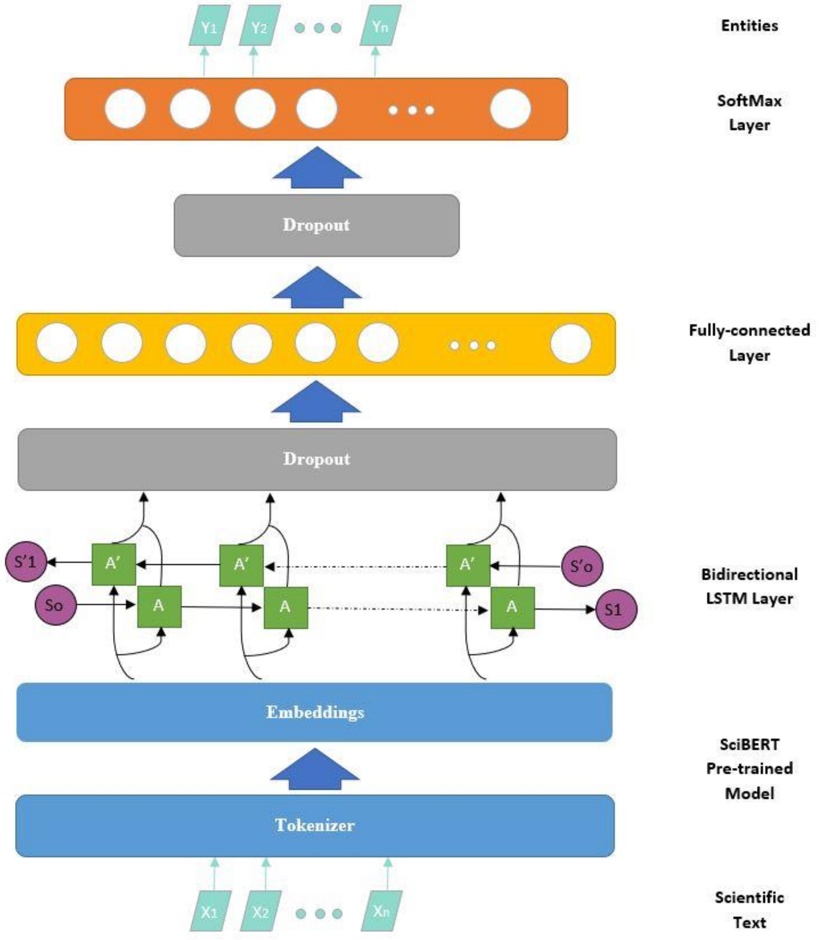
- On-going projects for document content analysis continue
  - Named entity recognition for scientific documents
  - New metrics for text summarization models
  - End-to-end argument mining and quality analysis
  - Topic extraction
- The goal is to combine all extracted facts from the content into a knowledge base



Mitacs Living Sky Technologies

Named-Entity Recognition

- Named-Entity Recognition over scientific texts is vital for extracting and understanding information.
- We introduce Sci-BERT+, an enhanced model to analyze scientific texts from the fields of Artificial Intelligence, Biomedical Engineering, and Natural Language Processing, by performing Named-Entity Recognition over them.
- Our solution utilizes a pre-trained scientific BERT-based language model connected to a bidirectional LSTM network.



Academic Article Information Extraction

University of Saskatchewan

Named-Entity Recognition

- Sci-BERT+ is capable of improving pattern recognition for scientific entity types with high accuracy.
- It was evaluated across three different datasets (SciERC, TDMSci, and NCBI-disease), emphasizing its ability to learn from and work with scientific articles' semantics and syntax.

TABLE I: Repeated holdout method results of our study and previous studies results

Reference Study	Dataset	Entities	BIO** F1	BIO** F1*	IO*** F1	IO*** F1*
SpERT [15]	SciERC	6	-	-	-	70.33
Sci-BERT+	SciERC	6	88.71 ± 0.47	71.11 ± 1.01	90.17 ± 0.34	**73.20 ± 1.06**
Flair-TDM [3]	TDMSci	3	76.47	-	-	-
Sci-BERT+	TDMSci	3	**92.49 ± 0.26**	70.91 ± 1.11	93.33 ± 0.26	76.17 ± 0.88
KeBioLM [16]	NCBI-disease	1	-	89.10	-	-
BLURB [44]	NCBI-disease	1	-	88.10	-	-
Sci-BERT+	NCBI-disease	1	98.16 ± 0.12	**89.59 ± 0.56**	98.18 ± 0.18	91.37 ± 0.85

* As stated this metric does not include No-Entities values into its calculation.
** Metrics values resulted from the outputs provided according to BIO encoding method.
*** Metrics values resulted from the outputs provided according to IO encoding method.

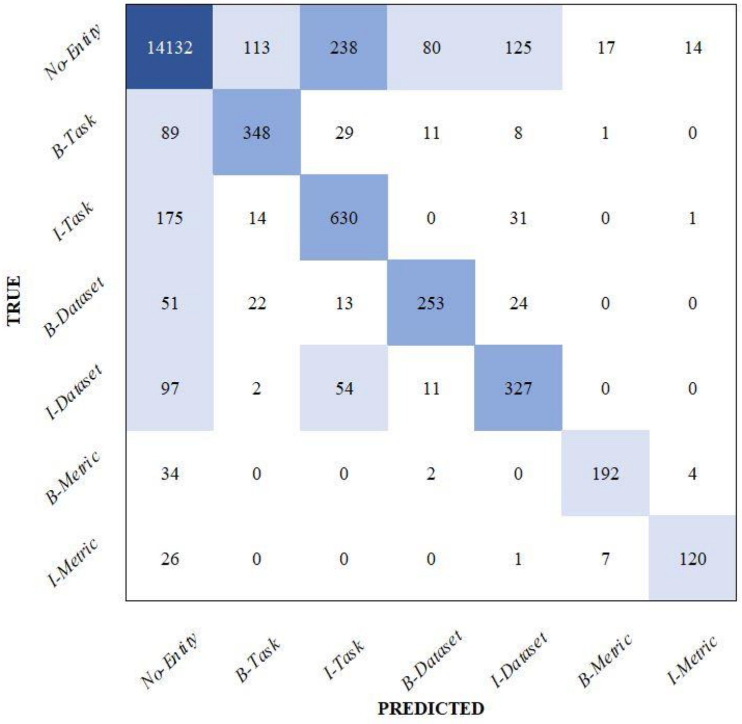
Mitacs Living Sky Technologies

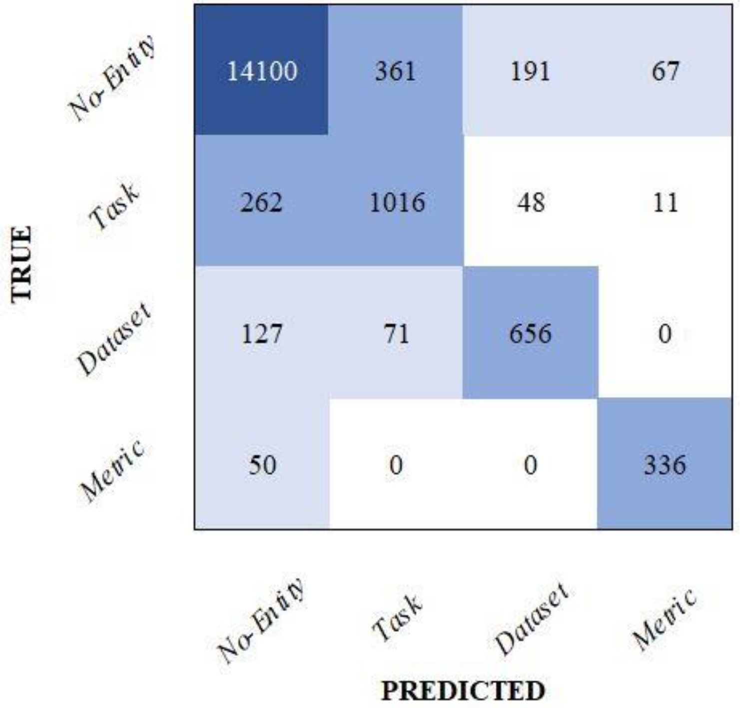
Academic Article Information Extraction

UNIVERSITY OF SASKATCHEWAN

Named-Entity Recognition

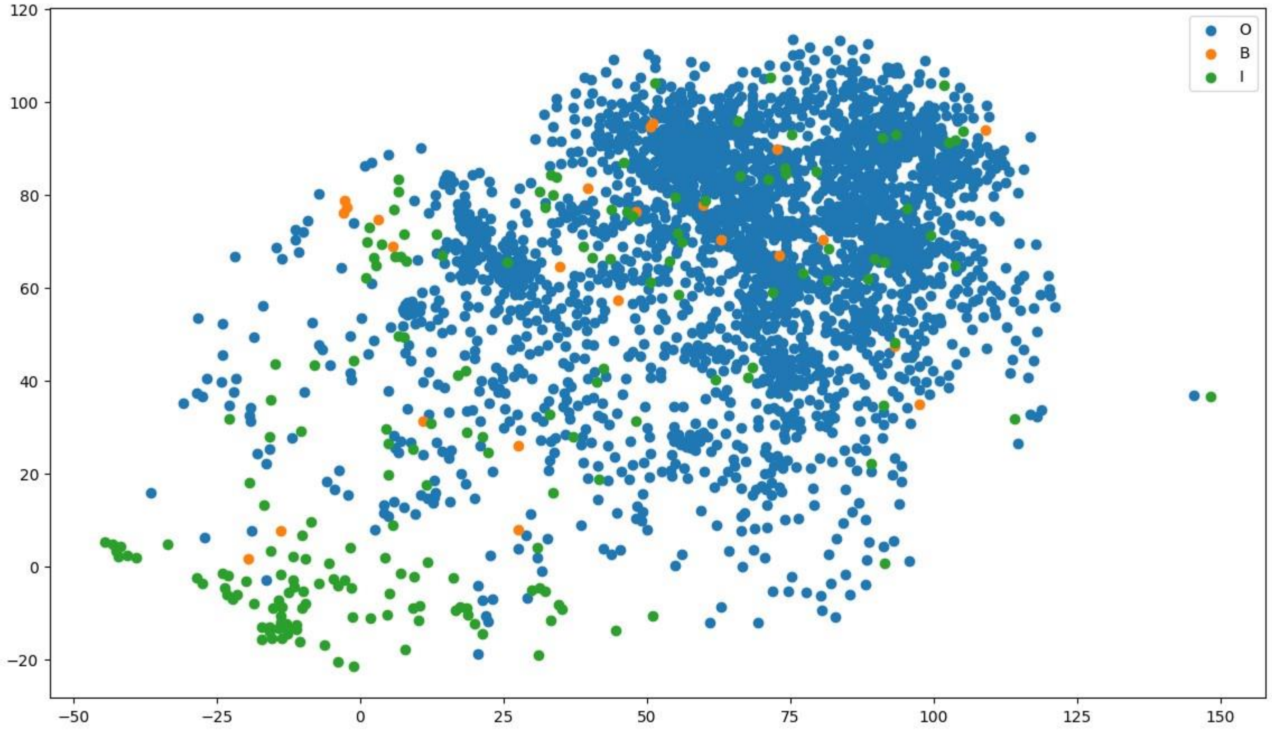
- Another advantage is seen while comparing the performance of our system between BIO and IO encoding methods. Our model could be used with either of these two encoding types, obtaining similar results (CM Example figures with the TDMSci Dataset).



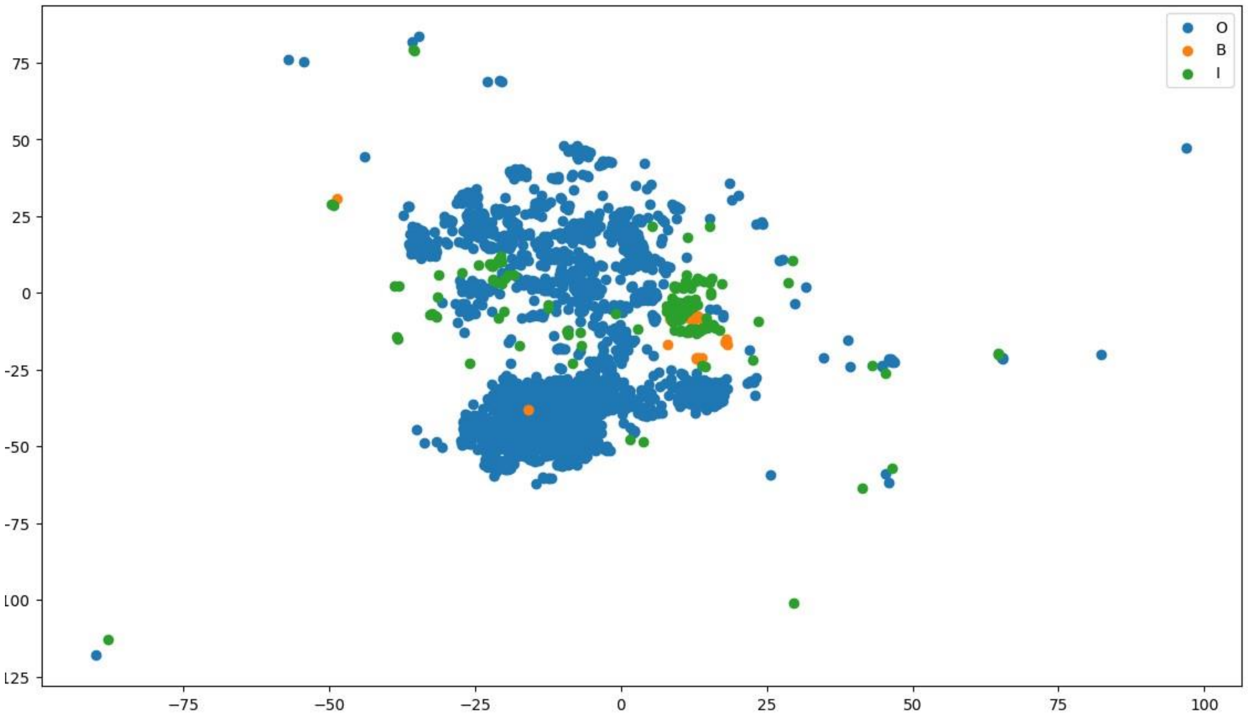


Named-Entity Recognition

- An implemented t-DSNE method shows the improvement obtained in classification while comparing the pure SciBERT model against our proposed solution, SciBERT^N.



SciBERT



SciBERT^N

Named-Entity Recognition

- It is important to emphasize that our model, SciBERT^N, refrains from fine-tuning the SciBERT layers and exclusively focuses on fine-tuning the LSTM layer. Consequently, our approach centers around creating a new foundational model with comparable attributes while achieving superior performance compared to a fine-tuned SciBERT model.

		SciBERT				SciBERT^N			
		Repeated Holdout		Cross-Validation		Repeated Holdout		Cross-Validation	
Encoding**	Datasets	F1	F1*	F1	F1*	F1	F1*	F1	F1*
BIO	SciERC***	89.07	73.25	96.71	93.18	89.18	72.12	96.39	91.17
	TDMSci***	92.69	72.50	98.78	94.94	92.75	72.02	97.93	91.70
	NCBI-Disease	97.78	88.00	98.97	94.68	98.28	90.15	99.19	95.66
	SciEnt	92.86	69.29	97.58	87.46	93.20	69.50	98.28	89.66
IO	SciERC	97.55	93.89	96.82	92.06	90.51	74.26	96.60	91.14
	TDMSci	93.38	76.76	98.16	93.13	93.59	77.05	98.85	95.68
	NCBI-Disease	98.17	91.39	99.40	97.12	98.36	92.22	99.31	96.74
	SciEnt	97.71	92.20	97.37	90.14	98.50	94.44	97.09	88.83

* Refer to Sect. IV-C2.
** BIO and IO encoding methods metrics values.
*** This models showed overfitting in their Repeated Holdout training evaluation method in both F1 metrics for SciBERT.
Note. Colored fields are used to show comparison metrics among models, and underlined values indicate the best score among these fields.

References

- Zhong, X., Tang, J., Yepes, A.J.: PubLayNet: Largest Dataset Ever for Document Layout Analysis. arXiv preprint, arXiv:1908.07836 (Aug 2019)
- He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. arXiv preprint, arXiv:1512.03385 (Dec 2015)
- Wick, C., Puppe, F.: Fully Convolutional Neural Networks for Page Segmentation of Historical Document Images. arXiv preprint, arXiv:1711.07695 (Nov 2017)
- Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI2015). pp. 234–241 (Nov 2015)
- Dong, R., Pan, X., Li, F.: DenseU-Net-Based Semantic Segmentation of Small Objects in Urban Remote Sensing Images. IEEE Access 7, 65347–65356 (May 2019)
- Dan Roth and Wen-tau Yih. A linear programming formulation for global inference in natural language tasks. In Proceedings of the Eighth Conference on Computational Natural Language Learning (CoNLL2004) at HLT-NAACL 2004, pages 1–8, Boston, Massachusetts, USA, May 6 - May 7 2004. Association for Computational Linguistics.
- George Stoica, Emmanouil Antonios Platanios, and Barnab'as P'oczos. Re-tacred: Addressing shortcomings of the tacred dataset. In AAAI, 2021.
- Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou, and Maosong Sun. Docred: A large-scale document-level relation extraction dataset. arXiv preprint arXiv:1906.06127, 2019.

Acknowledgements

University of Saskatchewan

NSERC CRSNG

NRC · CNRC

XILINX

Mitacs

KERI

intel

Jeonbuk National University

NVIDIA

CMC Microsystems

ETRI Electronics and Telecommunications Research Institute

ird International Road Dynamics Inc.

Keit 한국산업기술평가관리원 Korea Evaluation Institute of Industrial Technology

Microsoft

Living Sky Technologies