



New Technologies for
Cancer Detection

Machine Learning Drives
Medical Imaging Progress



CODSIT
Canadian Open Data Summit
Conference Highlights

Presenters

New Technologies for Canadian Observatories

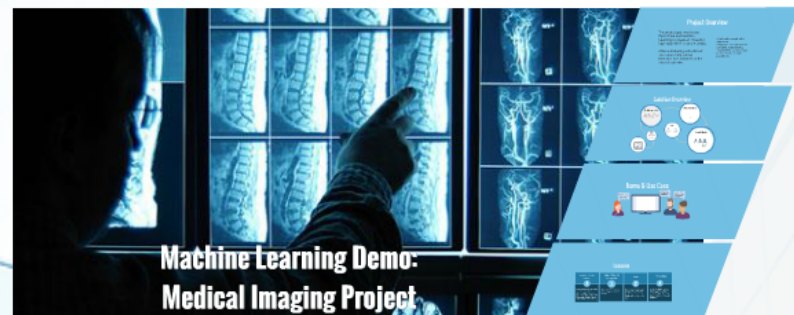
12.15.2017 AGM



LimbicConsulting

New Technologies for
Canadian Observatories

12.15.2017 AGM



Machine Learning Demo: Medical Imaging Project

Project Overview

The project goal was to use Deep Learning and AI to create a workflow to create an accurate, high value MRI in only 5 minutes. After collecting a list of 1000 cases, the current state of the project was as follows:

Advanced capabilities in AI/ML. The project has been successful in creating a workflow to create an accurate, high value MRI in only 5 minutes.

Solution Overview



Demo & Use Case



Lessons

Lesson	Impact	Key Takeaway
1. Data Quality	High	Ensure data is clean and representative.
2. Model Training	Medium	Use appropriate hyperparameters and regularization.
3. Model Deployment	High	Ensure the model is integrated into the clinical workflow.
4. Model Evaluation	High	Use appropriate metrics to evaluate model performance.

Project Overview

The project goal was to use Open Data and Machine Learning to create an impactful, high value MVP in only 6 weeks.

After considering a shortlist of use cases, lung cancer detection was selected as the project objective.

- Relatively unexplored in astronomy
- High impact requires more complete understanding
- Could drive future surveys based on data analysis procedures

- Relatively unexplored in astronomy
- High impact requires more complete understanding
- Could drive future surveys based on data analysis procedures

Project Overview

The project goal was to use Open Data and Machine Learning to create an impactful, high value MVP in only 6 weeks.

After considering a shortlist of use cases, lung cancer detection was selected as the project objective.

- Relatively unexplored in astronomy
- High impact requires more complete understanding
- Could drive future surveys based on data analysis procedures

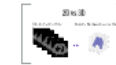
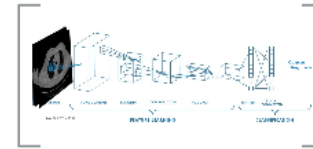
Solution Overview

Medical Imaging Data

- Low-dose computed tomography (CT) scan advantages:
- Low dose CT scans use less radiation
 - Low dose CT scans require no injections or dyes
 - Scans take less than a minute
 - CT scanners are relatively cheap



Neural Networks



Results



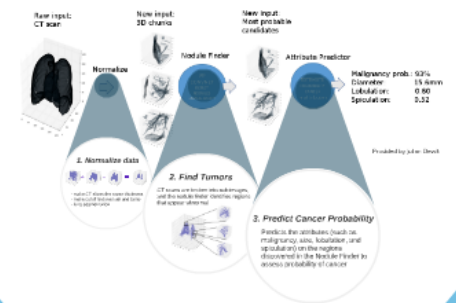
Project Approach

- 1 Understand the methods and reproduce results from Julian and Daniel's model
- 2 Develop a simple interface to create a diagnostic tool for doctors and radiologists

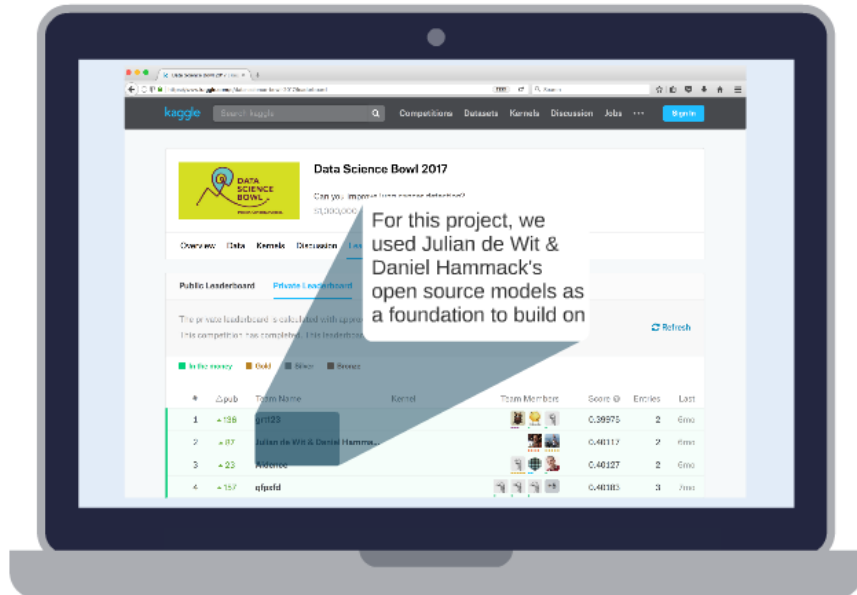
kaggle



How it Works



kaggle™





Data Science Bowl 2017

Can you improve lung cancer detection?

\$1,000,000

Overview | Data | Kernels | Discussion | Leaderboard

Public Leaderboard | Private Leaderboard

The private leaderboard is calculated with approximately 10% of the data. This competition has completed. This leaderboard is frozen.

Refresh

In the money Gold Silver Bronze

#	Δ pub	Team Name	Kernel	Team Members	Score	Entries	Last
1	▲136	grt123			0.39975	2	6mo
2	▲87	Julian de Wit & Daniel Hamma...			0.40117	2	6mo
3	▲23	Aidence			0.40127	2	6mo
4	▲157	qfpxfd		+5	0.40183	3	7mo

For this project, we used Julian de Wit & Daniel Hammack's open source models as a foundation to build on

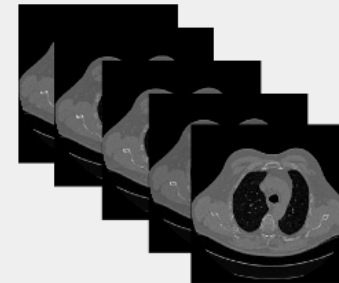
Project Approach

- 1** Understand the methods and reproduce results from Julian and Daniel's model
- 2** Develop a simple interface to create a diagnostic tool for doctors and radiologists

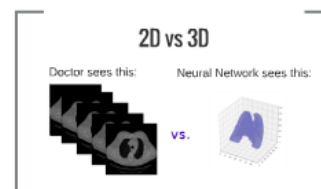
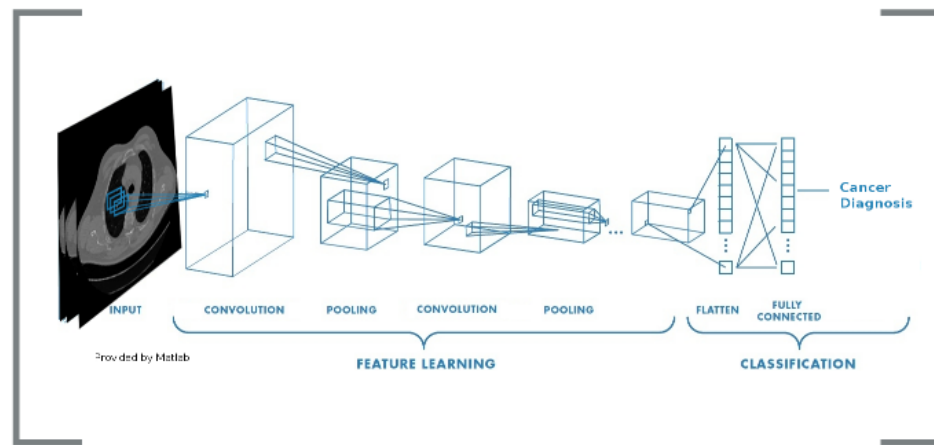
Medical Imaging Data

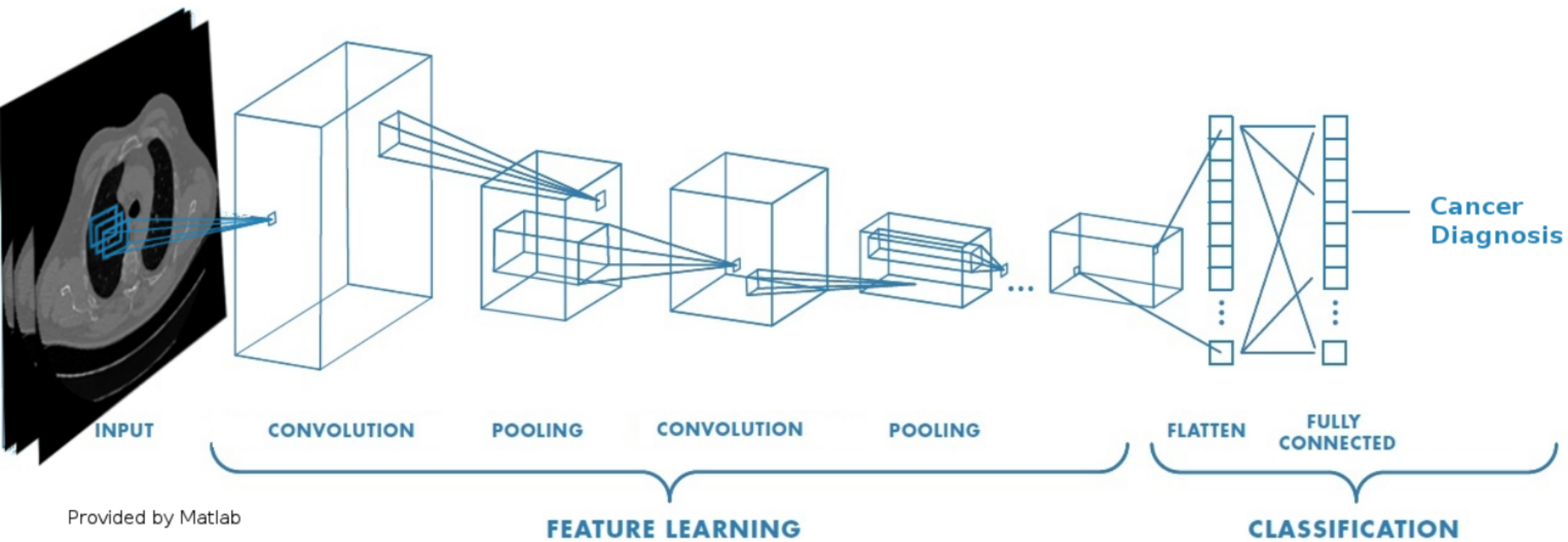
Low-dose computed tomography (CT) scan advantages:

- Low dose CT scans use less radiation
- Low dose CT scans require no injections or dyes
- Scans take less than a minute
- CT scanners are relatively cheap



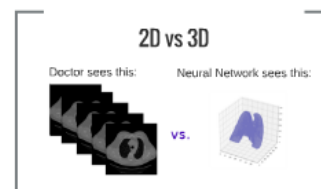
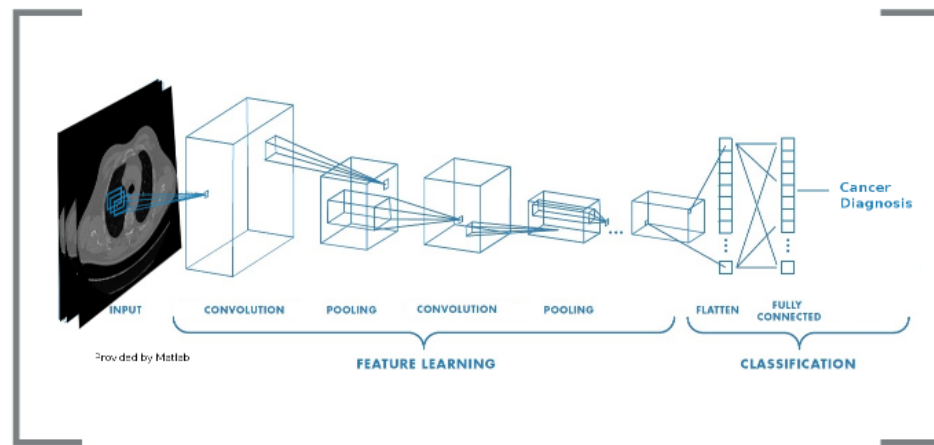
Neural Networks





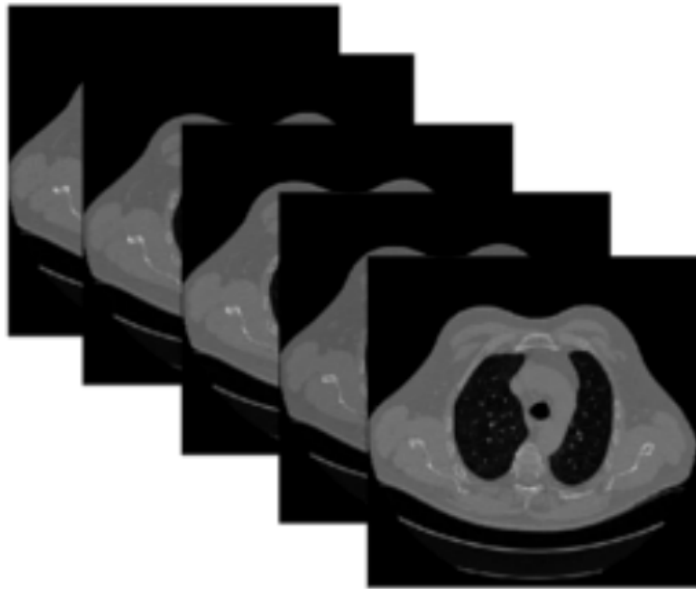
Provided by Matlab

Neural Networks



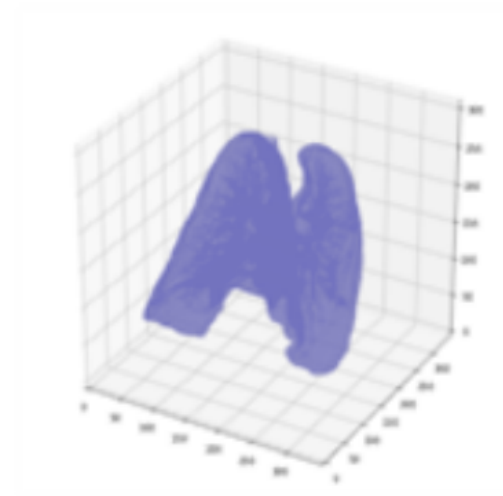
2D vs 3D

Doctor sees this:

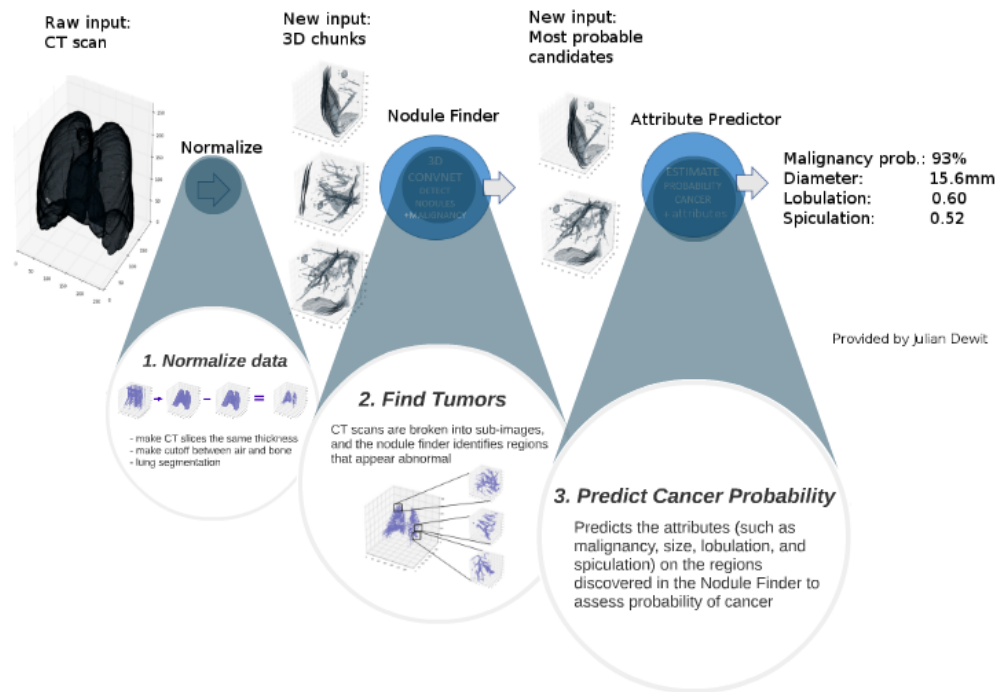


Neural Network sees this:

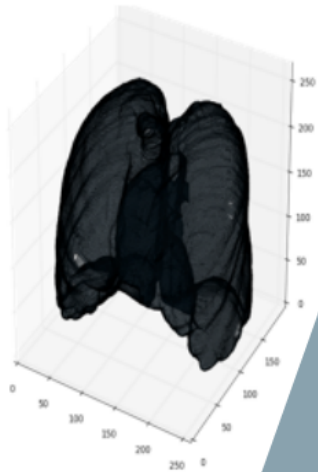
vs.



How it Works



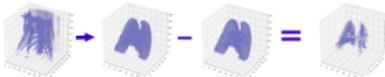
Raw input:
CT scan



Normalize

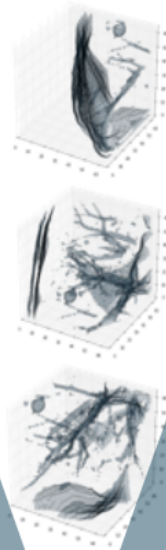


1. Normalize data

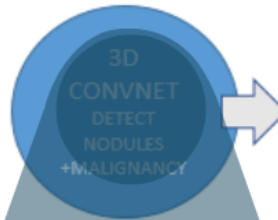


- make CT slices the same thickness
- make cutoff between air and bone
- lung segmentation

New input:
3D chunks

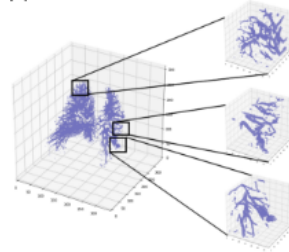


Nodule Finder

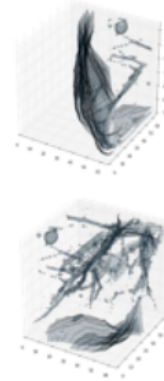


2. Find Tumors

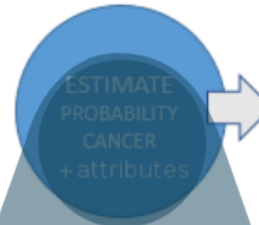
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



New input:
Most probable
candidates



Attribute Predictor



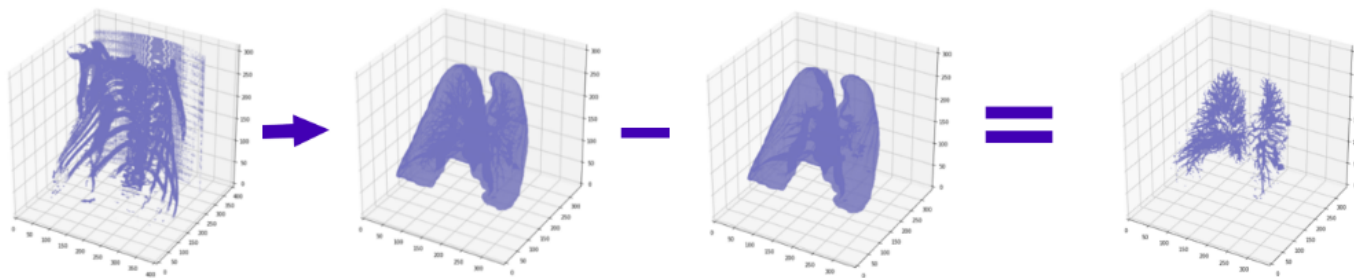
3. Predict Cancer Probability

Predicts the attributes (such as malignancy, size, lobulation, and spiculation) on the regions discovered in the Nodule Finder to assess probability of cancer

Malignancy prob.: 93%
Diameter: 15.6mm
Lobulation: 0.60
Spiculation: 0.52

Provided by Julian Dewit

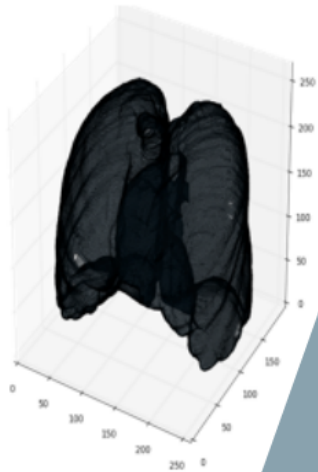
1. Normalize data



- make CT slices the same thickness
- make cutoff between air and bone
- lung segmentation

CT scan
and the
that a

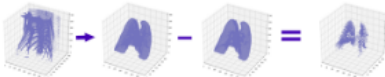
Raw input:
CT scan



Normalize

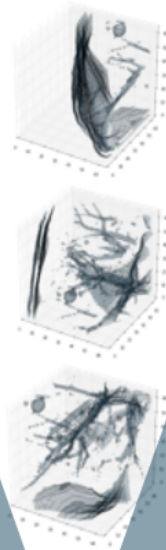


1. Normalize data

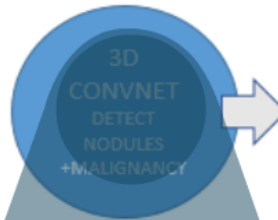


- make CT slices the same thickness
- make cutoff between air and bone
- lung segmentation

New input:
3D chunks

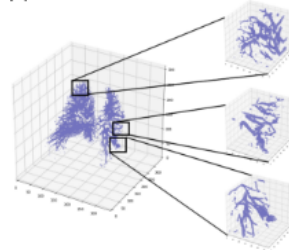


Nodule Finder

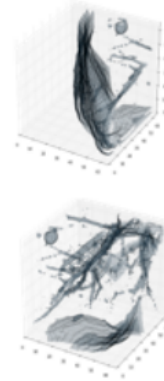


2. Find Tumors

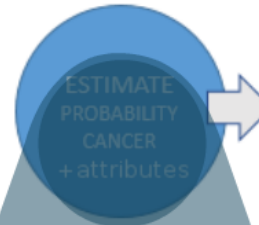
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



New input:
Most probable
candidates



Attribute Predictor



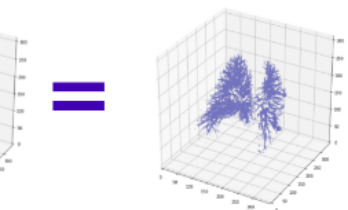
3. Predict Cancer Probability

Predicts the attributes (such as malignancy, size, lobulation, and spiculation) on the regions discovered in the Nodule Finder to assess probability of cancer

Malignancy prob.: 93%
Diameter: 15.6mm
Lobulation: 0.60
Spiculation: 0.52

Provided by Julian Dewit

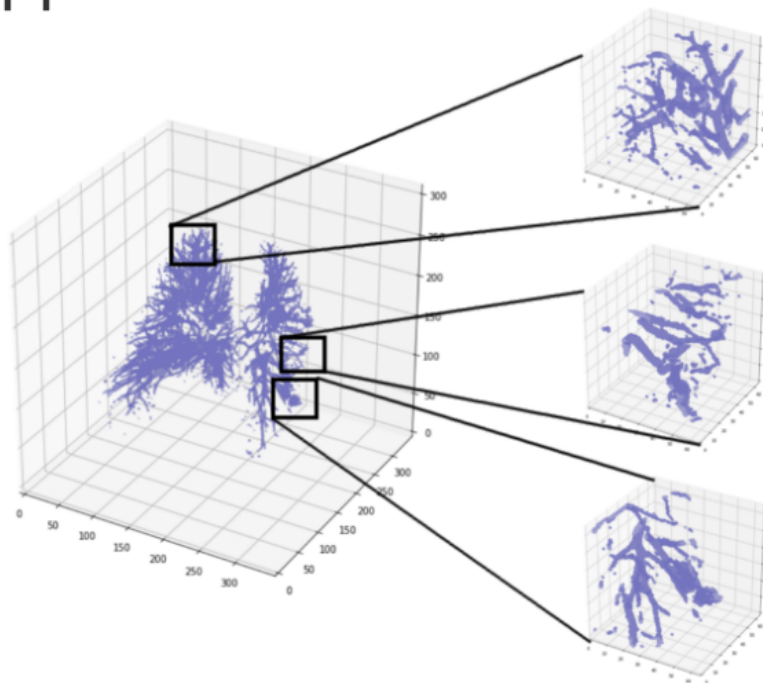
data



the thickness
and bone

2. Find Tumors

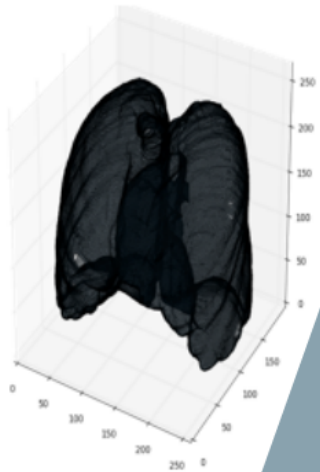
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



3. Predict

Predicts
malignant
spiculation
discover
assess p

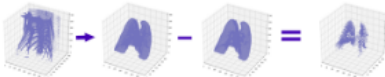
Raw input:
CT scan



Normalize

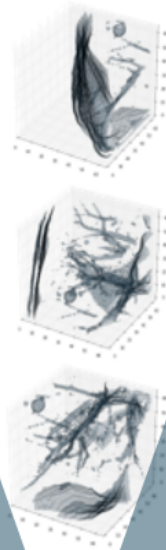


1. Normalize data

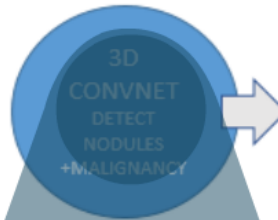


- make CT slices the same thickness
- make cutoff between air and bone
- lung segmentation

New input:
3D chunks

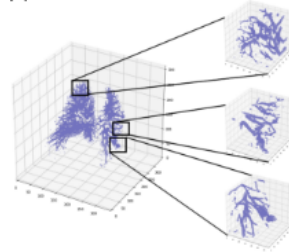


Nodule Finder

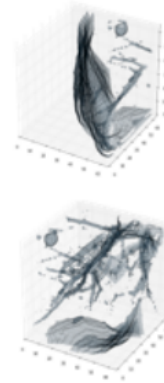


2. Find Tumors

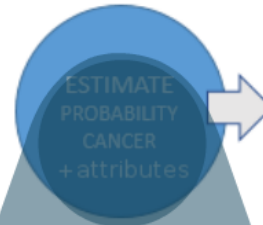
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



New input:
Most probable
candidates



Attribute Predictor



3. Predict Cancer Probability

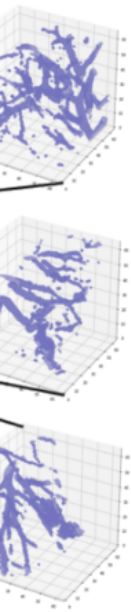
Predicts the attributes (such as malignancy, size, lobulation, and spiculation) on the regions discovered in the Nodule Finder to assess probability of cancer

Malignancy prob.: 93%
Diameter: 15.6mm
Lobulation: 0.60
Spiculation: 0.52

Provided by Julian Dewit

ors

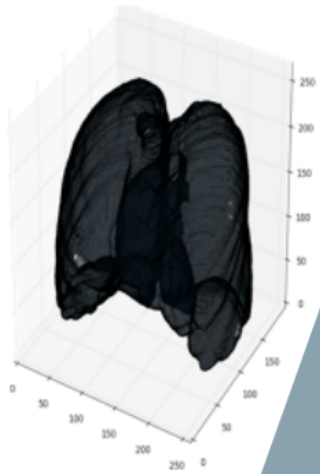
sub-images,
identifies regions



3. Predict Cancer Probability

Predicts the attributes (such as malignancy, size, lobulation, and spiculation) on the regions discovered in the Nodule Finder to assess probability of cancer

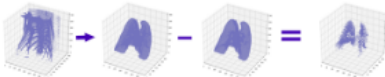
Raw input:
CT scan



Normalize

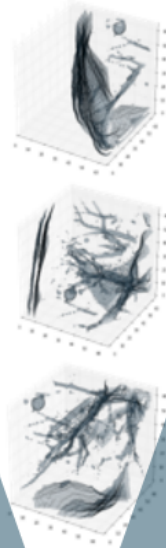


1. Normalize data

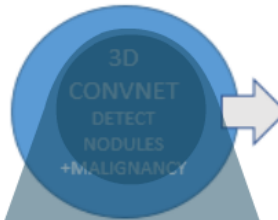


- make CT slices the same thickness
- make cutoff between air and bone
- lung segmentation

New input:
3D chunks

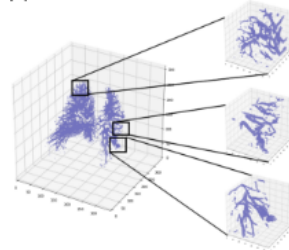


Nodule Finder

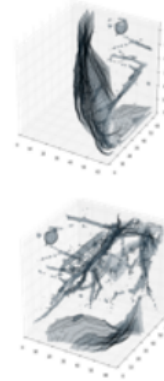


2. Find Tumors

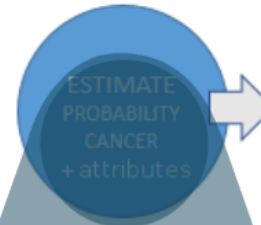
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



New input:
Most probable
candidates



Attribute Predictor



3. Predict Cancer Probability

Predicts the attributes (such as malignancy, size, lobulation, and spiculation) on the regions discovered in the Nodule Finder to assess probability of cancer

Malignancy prob.: 93%
Diameter: 15.6mm
Lobulation: 0.60
Spiculation: 0.52

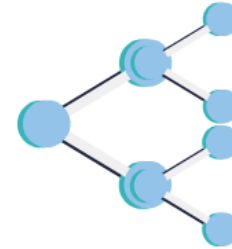
Provided by Julian Dewit

Results



Human (trained professional)

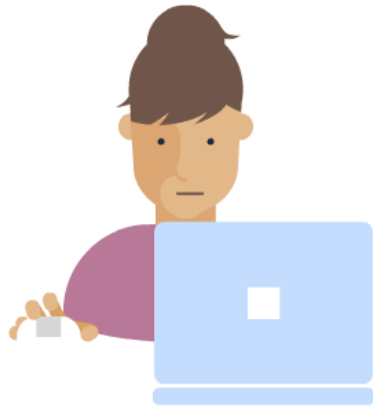
- Non-Cancerous Cases: 66%
- Cancerous Cases: 89%



Convolutional Neural Network

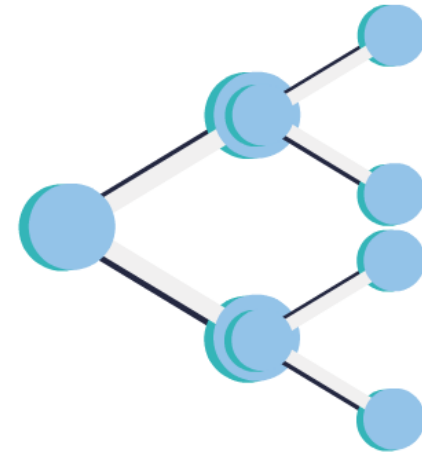
- ✓ • Non-Cancerous Cases: 69%
- ✓ • Cancerous Cases: 93%

RESULTS



Human (trained professional)

- Non-Cancerous Cases: 66%
- Cancerous Cases: 89%



Convolutional Neural Network

- ✓ • Non-Cancerous Cases: 69%
- ✓ • Cancerous Cases: 93%

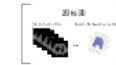
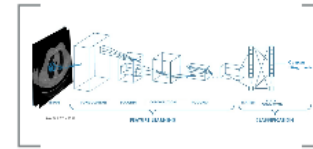
Solution Overview

Medical Imaging Data

- Low-dose computed tomography (CT) scan advantages:
- Low dose CT scans use less radiation
 - Low dose CT scans require no injections or dyes
 - Scans take less than a minute
 - CT scanners are relatively cheap



Neural Networks



Results



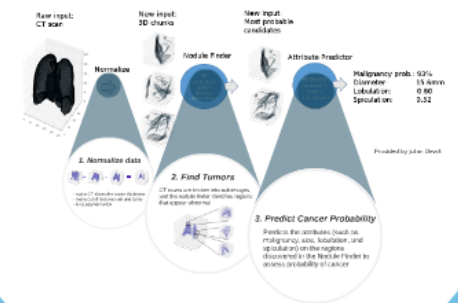
Project Approach

- 1 Understand the methods and reproduce results from Julian and Daniel's model
- 2 Develop a simple interface to create a diagnostic tool for doctors and radiologists

kaggle

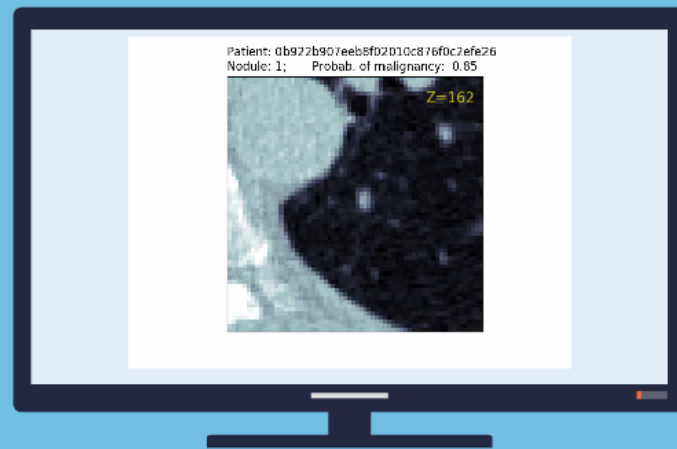


How it Works



Demo & Use Case

Based on this data and our observations, we recommend immediate treatment.



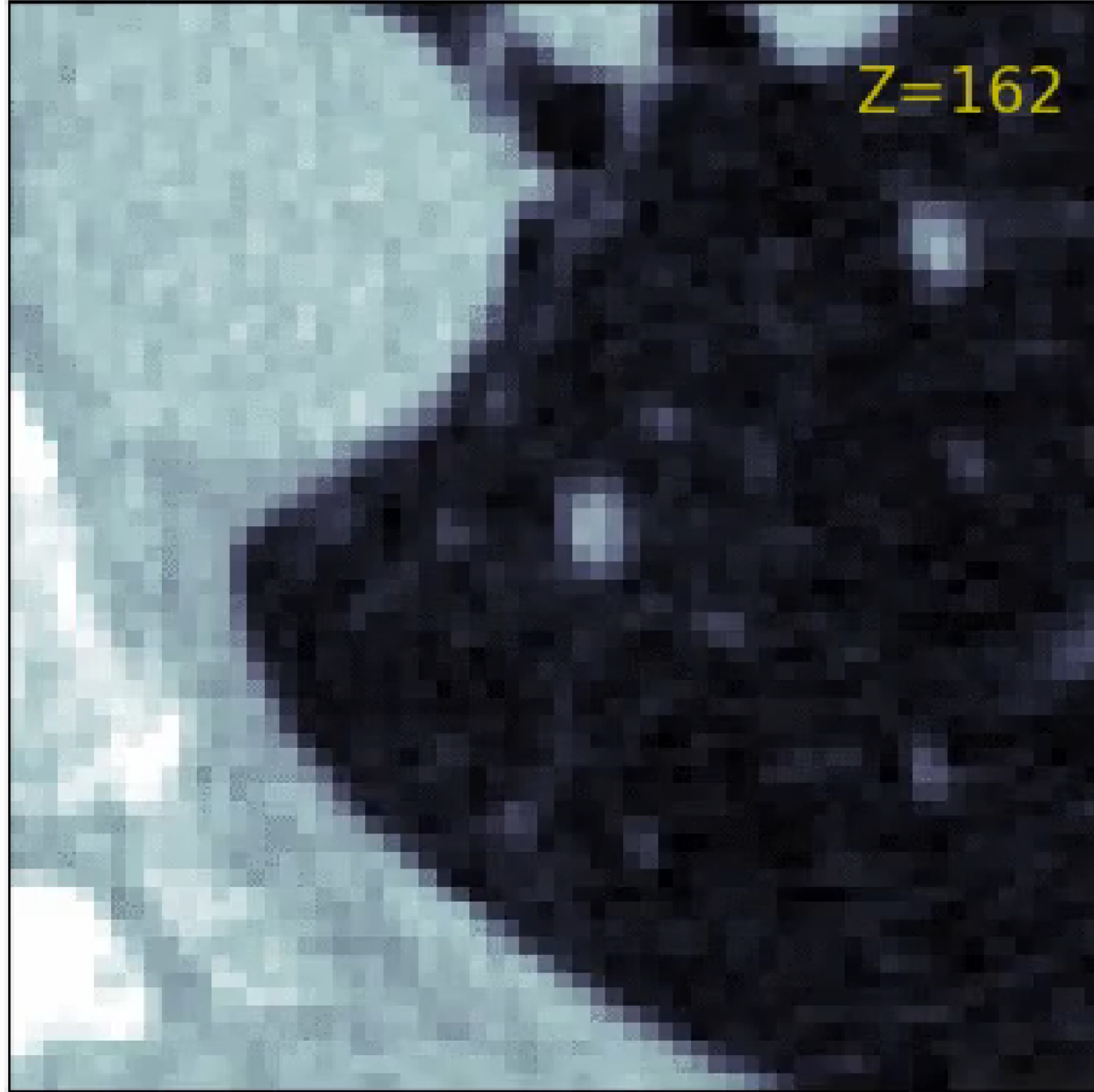
Let's isolate and examine just the red area, which shows the highest probability of a malignant tumor.



The system has flagged this patient's recent CT scans. Let's take a look and visually assess the areas the Neural Net has identified.

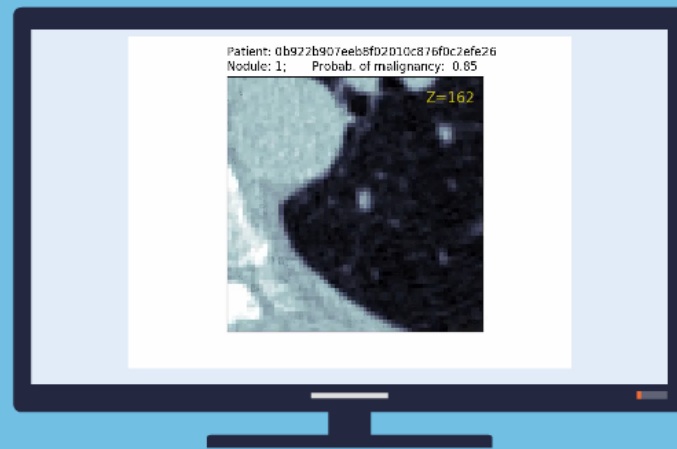


Patient: 0b922b907eeb8f02010c876f0c2efe26
Nodule: 1; Probab. of malignancy: 0.85



Demo & Use Case

Based on this data and our observations, we recommend immediate treatment.



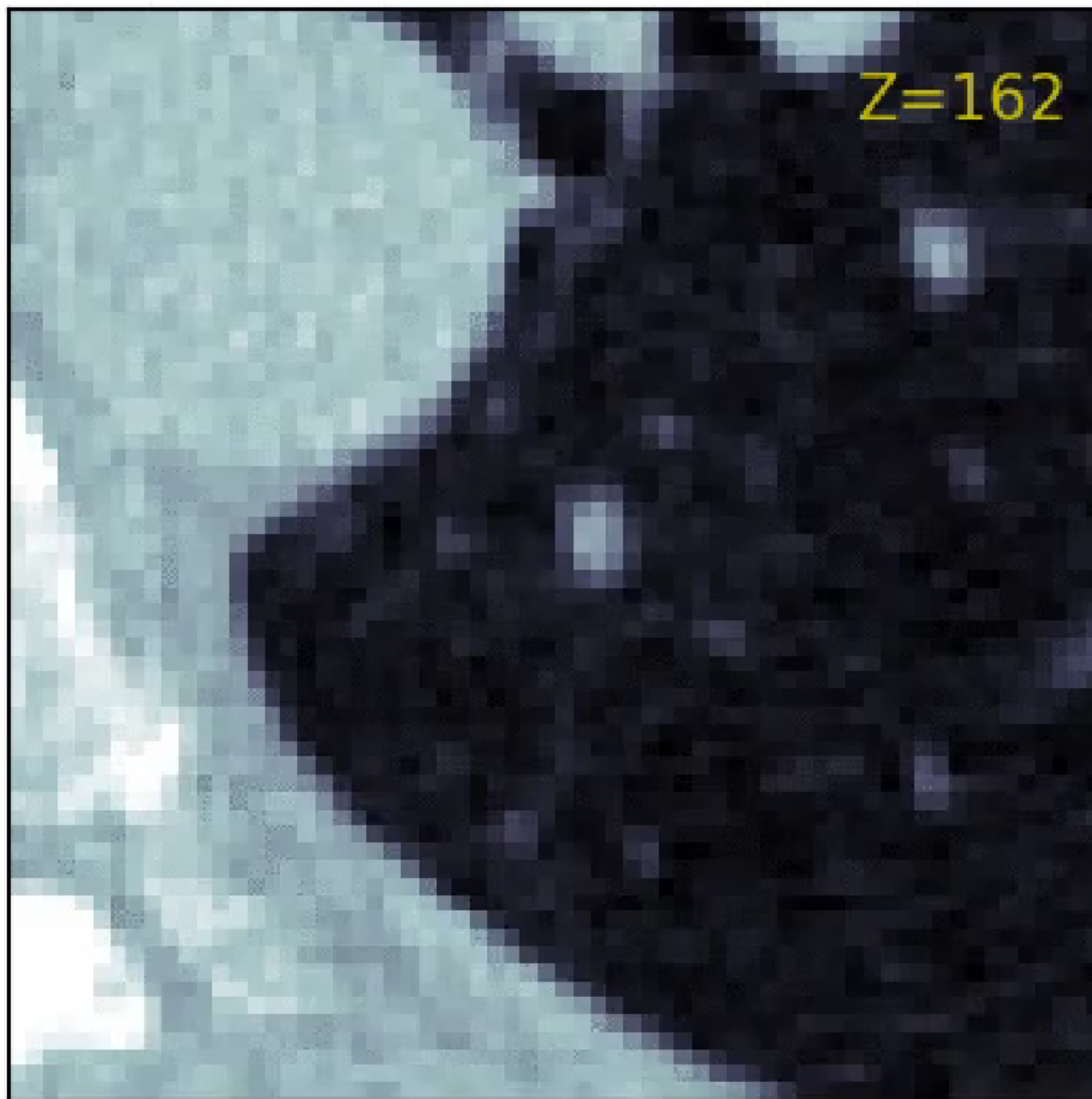
Let's isolate and examine just the red area, which shows the highest probability of a malignant tumor.



The system has flagged this patient's recent CT scans. Let's take a look and visually assess the areas the Neural Net has identified.

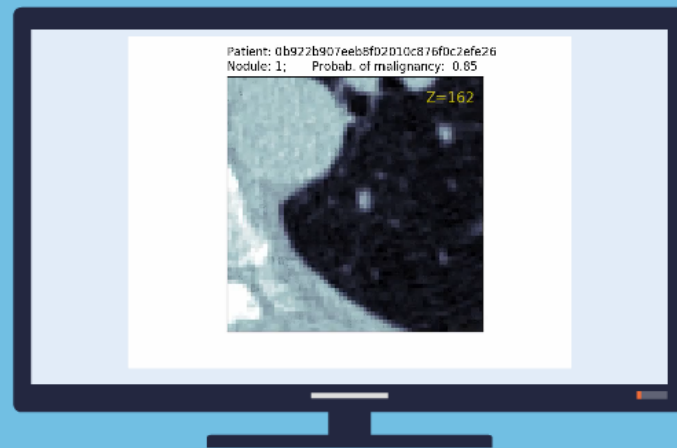


Patient: 0b922b907eeb8f02010c876f0c2efe26
Nodule: 1; Probab. of malignancy: 0.85



Demo & Use Case

Based on this data and our observations, we recommend immediate treatment.



Let's isolate and examine just the red area, which shows the highest probability of a malignant tumor.



The system has flagged this patient's recent CT scans. Let's take a look and visually assess the areas the Neural Net has identified.



Lessons

Impactful Use cases

1

Not all potential use cases are relevant; select important problems and provide actionable insights that would be hard to get without Machine Learning

DEEP Neural Networks

2

State of the art models are getting deeper and deeper.

Data

3

The research and rapid implementation of this project was enabled by large, well-curated data sets.

Flexibility

4

Agility, flexibility, and an R&D mindset enable successful outcomes for data-driven, AI-powered projects.

Machine Learning Demo: Medical Imaging Project

Project Overview

The project goal was to use Deep Learning and AI to create a workflow to create an accurate, high value MRI in only 5 minutes. After collecting a list of 1000 cases, the current state of the art was used as the project objective.

Advanced capabilities in AI/ML. The project was a success in terms of accuracy and speed. The project was a success in terms of accuracy and speed.

Solution Overview



Demo & Use Case



Lessons

Lesson	Impact	Key Takeaway
1. Data Quality	High	Ensure data is clean and representative.
2. Model Selection	Medium	Choose the right model for the task.
3. Model Deployment	High	Ensure the model is deployed in a scalable and secure environment.
4. Evaluation	High	Use a variety of metrics to evaluate model performance.

New Technologies for
Canadian Observatories

12.15.2017 AGM

