
Machine Learning Demo:
Medical Imaging Project
Meetup News
CODSIT
Canadian Open Data Summit:
Conference Highlights
Presenters

New Technologies for Canadian Observatories

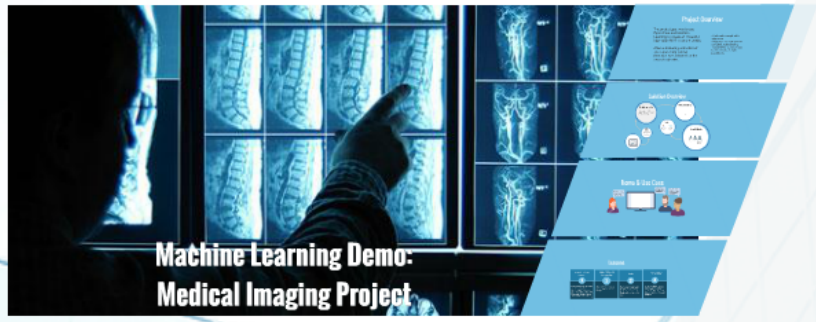
12.15.2017 AGM

LimbicConsulting

New Technologies for
Canadian Observatories

12.15.2017 AGM

LimbicConsulting




Project Overview
Solution Overview
Demo & Use Case
Lessons
Machine Learning Demo:
Medical Imaging Project

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

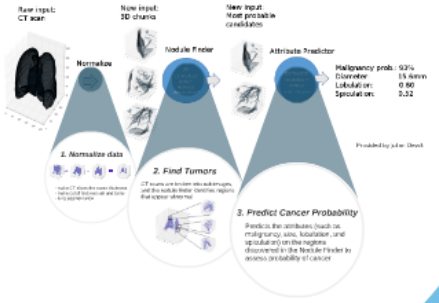
Human (trained professional)

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

Convolutional Neural Network

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

How it Works



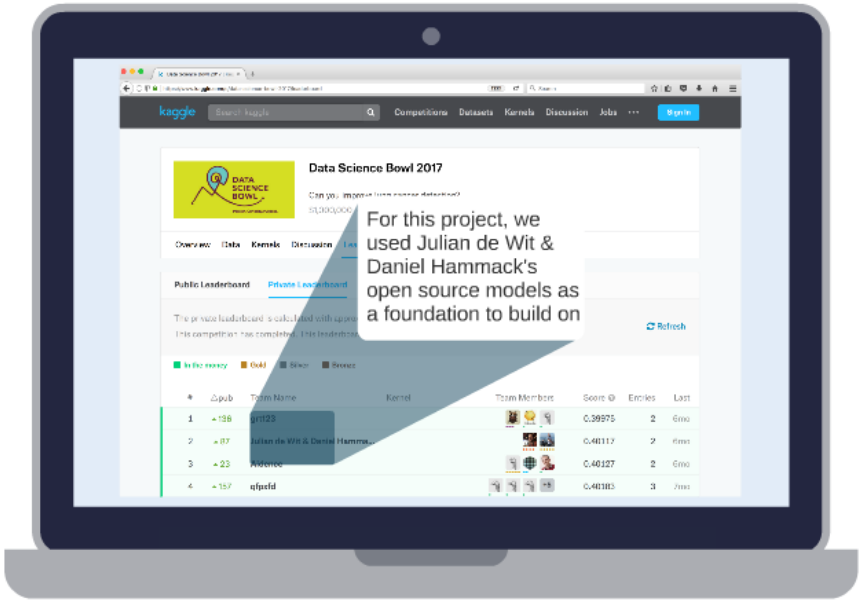
Project Approach

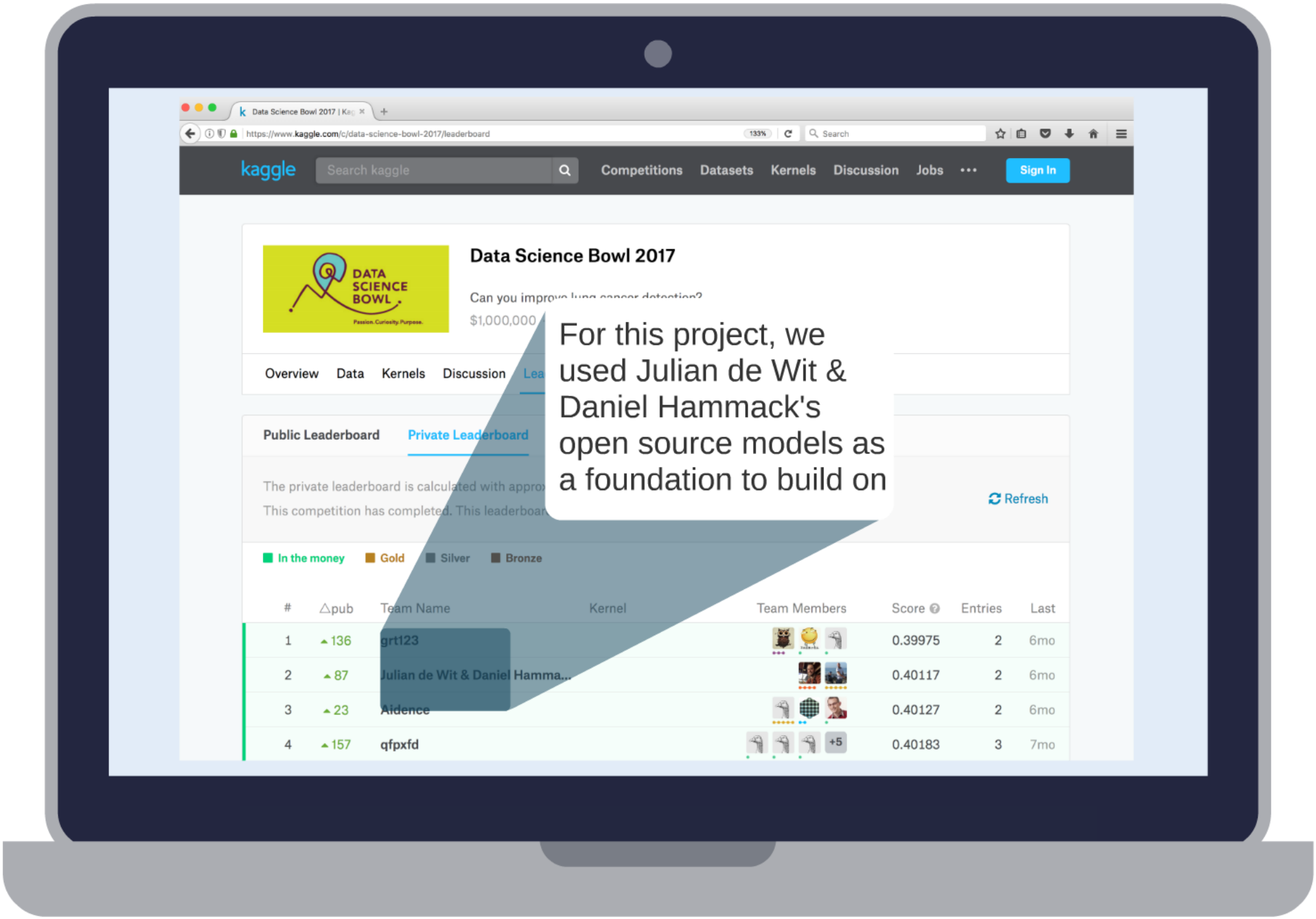
1. Understand the methods and reproduce results from Julian and Daniel's model
2. Develop a simple interface to create a diagnostic tool for doctors and radiologists

kaggle

2 D
to
fo

kaggle™




Data Science Bowl 2017 | Kaggle
https://www.kaggle.com/c/data-science-bowl-2017/leaderboard
kaggle
Search kaggle
Competitions
Datasets
Kernels
Discussion
Jobs
Sign In
DATA SCIENCE BOWL
Passion. Curiosity. Purpose.
Data Science Bowl 2017
Can you improve lung cancer detection?
$1,000,000
Overview
Data
Kernels
Discussion
Public Leaderboard
Private Leaderboard
The private leaderboard is calculated with approx
This competition has completed. This leaderboar
Refresh
In the money
Gold
Silver
Bronze
△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

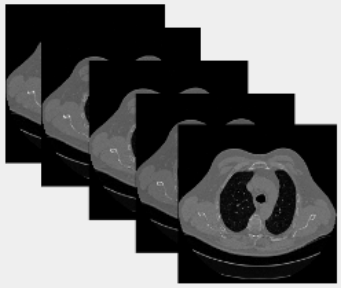
1. Understand the methods and reproduce results from Julian and Daniel's model

2. Develop a simple interface to create a diagnostic tool for doctors and radiologists

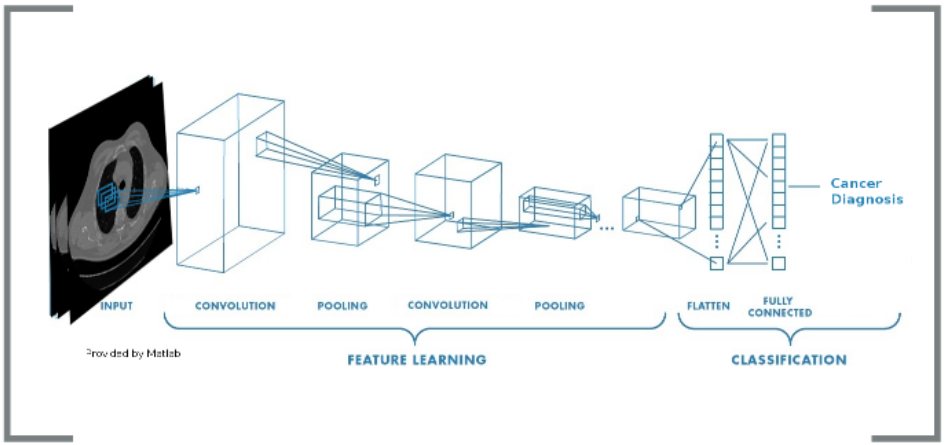
Medical Imaging Data

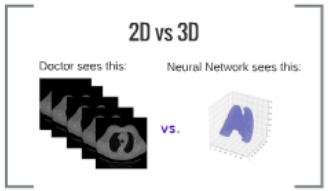
Low-dose computed tomography (CT) scan advantages:

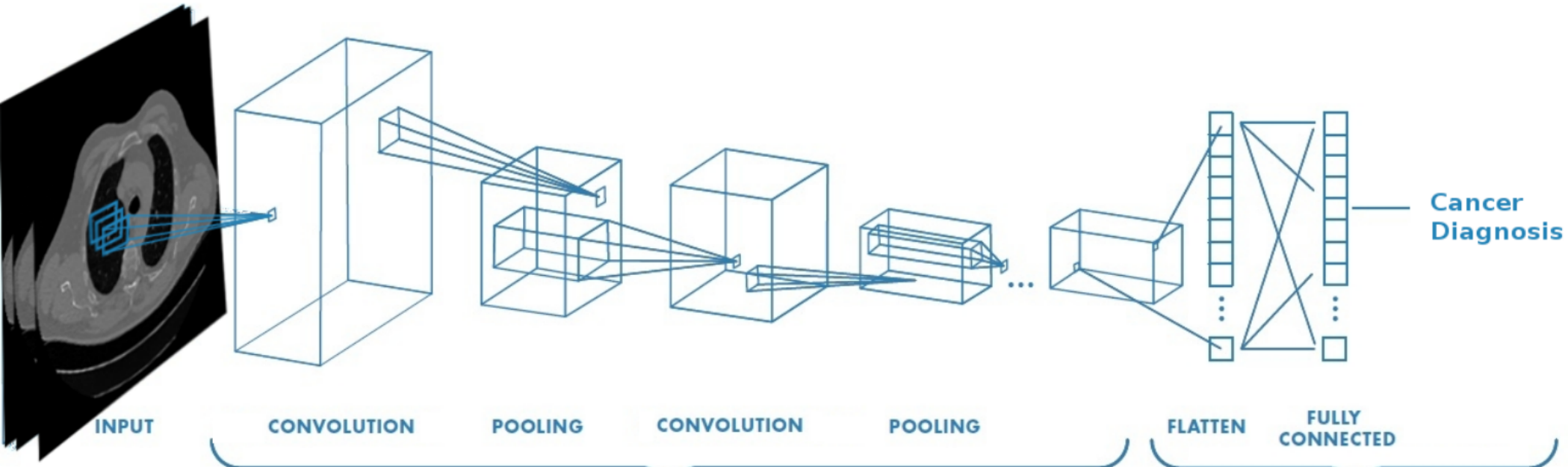
- Low dose CT scans use less radiation
- Low dose CT scans require no injections or dyes
- Scans take less than a minute
- CT scanners are relatively cheap



Neural Networks

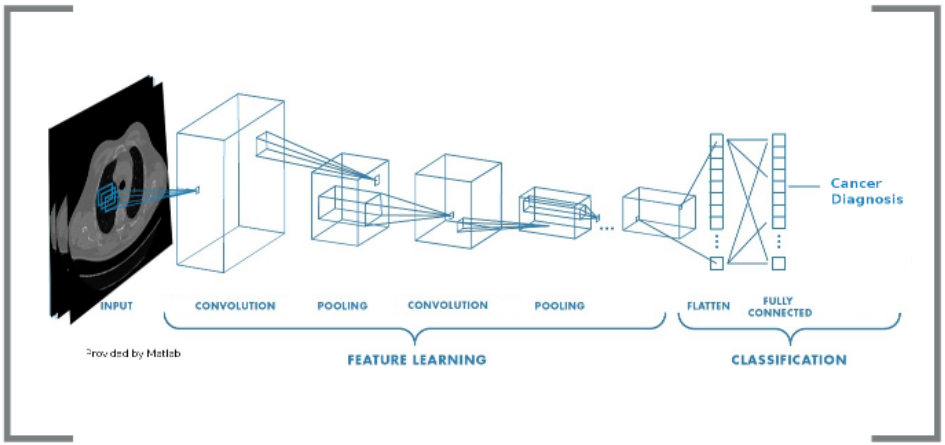


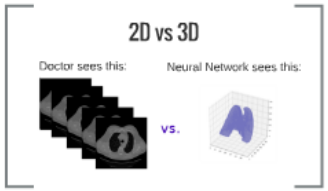




Provided by Matlab

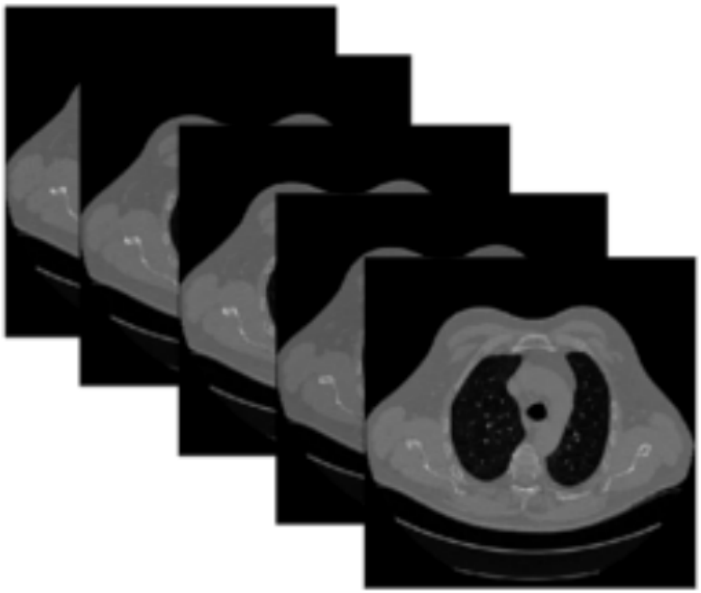
Neural Networks





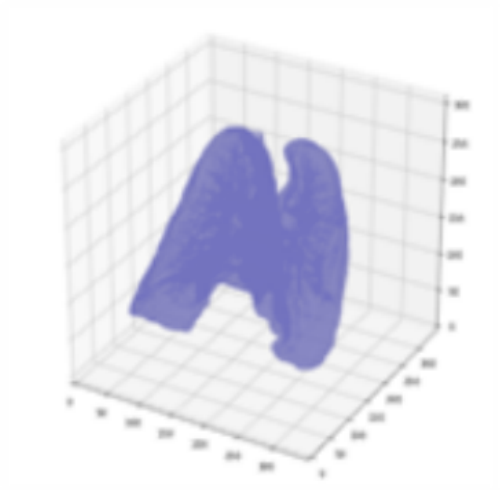
2D vs 3D

Doctor sees this:

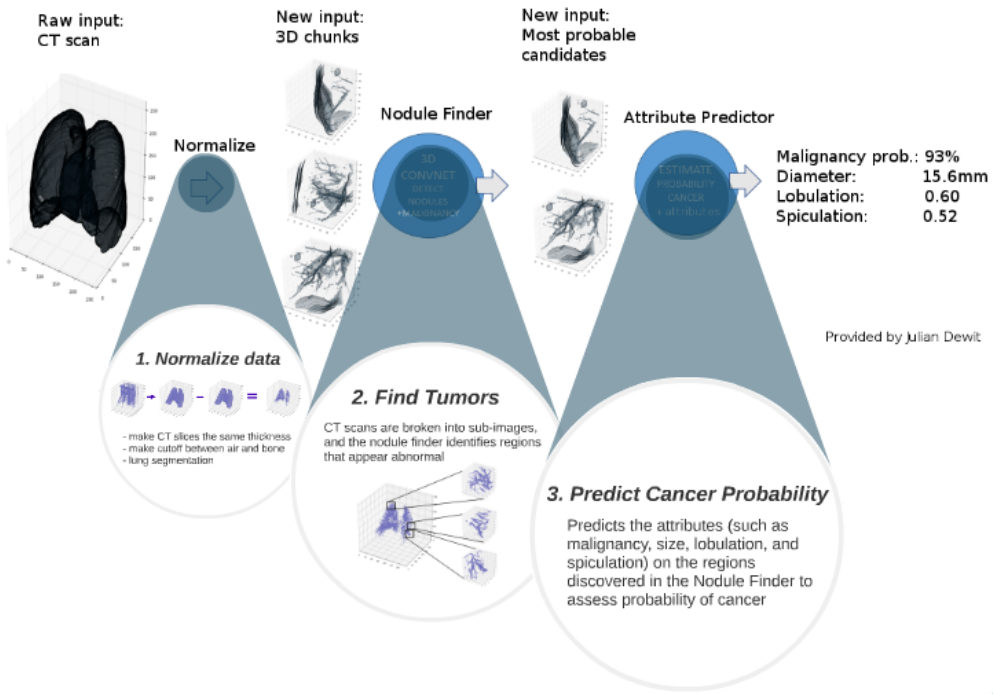


vs.

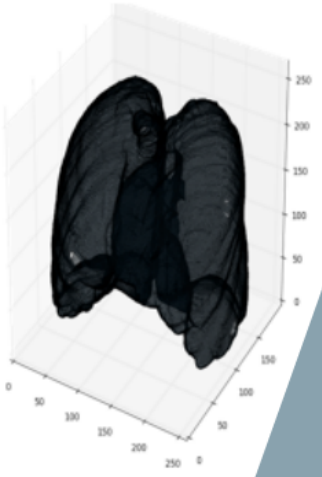
Neural Network sees this:



How it Works

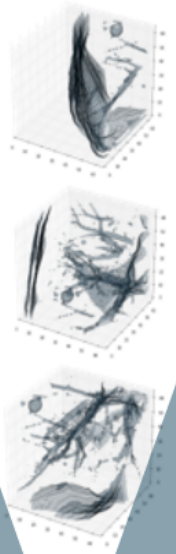

Raw input:
CT scan
Normalize
New input:
3D chunks
Nodule Finder
3D CONVNET
DETECT NODULES +MALIGNANCY
New input:
Most probable candidates
Attribute Predictor
ESTIMATE PROBABILITY CANCER +attributes
Malignancy prob.: 93%
Diameter: 15.6mm
Lobulation: 0.60
Spiculation: 0.52
Provided by Julian Dewit
1. Normalize data
- make CT slices the same thickness
- make cutoff between air and bone
- lung segmentation
2. Find Tumors
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal
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

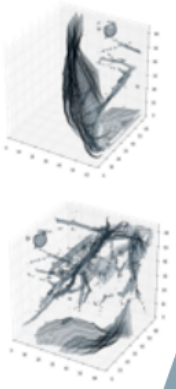
New input:
3D chunks



Nodule Finder

3D CONVNET
DETECT NODULES
+MALIGNANCY

New input:
Most probable candidates



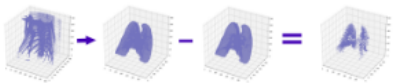
Attribute Predictor

ESTIMATE PROBABILITY CANCER
+attributes

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

Provided by Julian Dewit

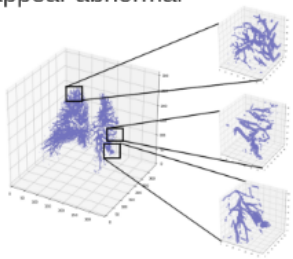
1. Normalize data



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

2. Find Tumors

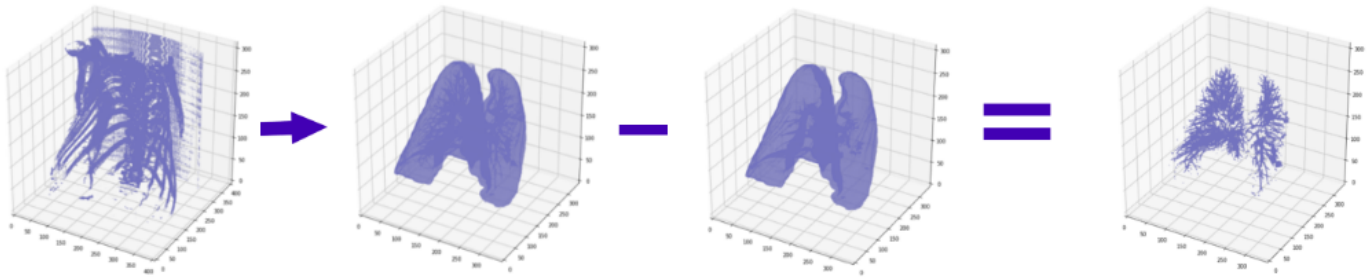
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



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

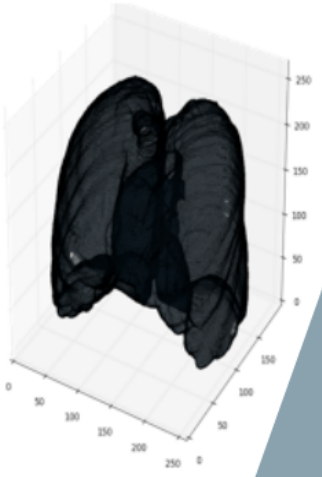
1. Normalize data



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

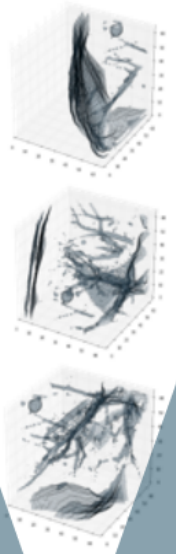
CT sc
and th
that a

Raw input:
CT scan



Normalize

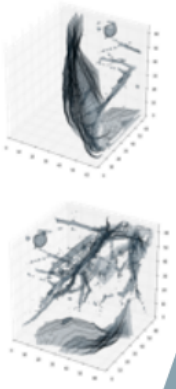
New input:
3D chunks



Nodule Finder

3D CONVNET
DETECT NODULES
+MALIGNANCY

New input:
Most probable candidates



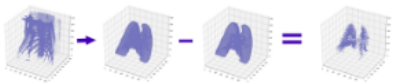
Attribute Predictor

ESTIMATE PROBABILITY CANCER
+attributes

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

Provided by Julian Dewit

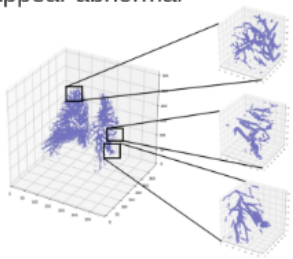
1. Normalize data



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

2. Find Tumors

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

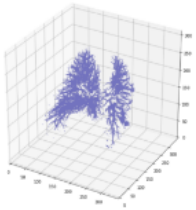


3. Predict Cancer Probability

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

e data

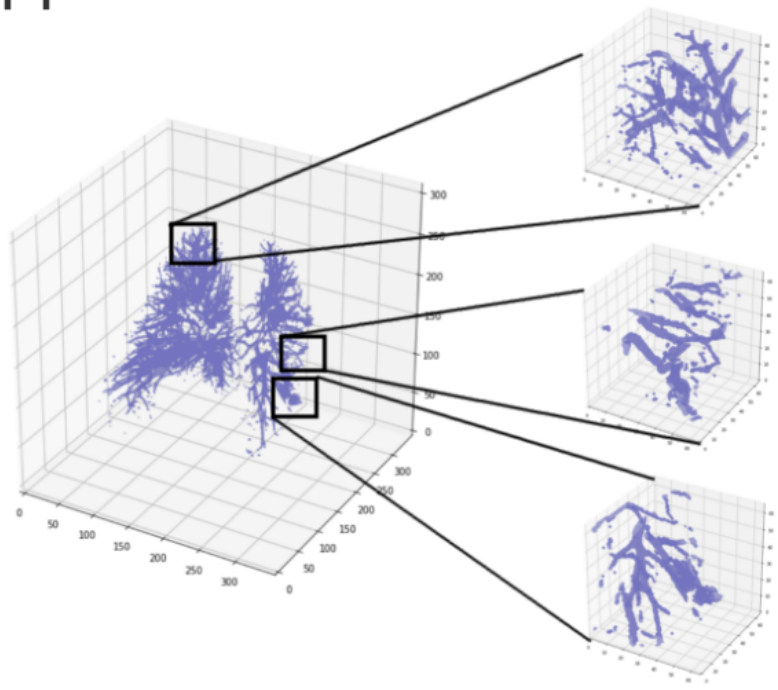
=



ne thickness
r and bone

2. Find Tumors

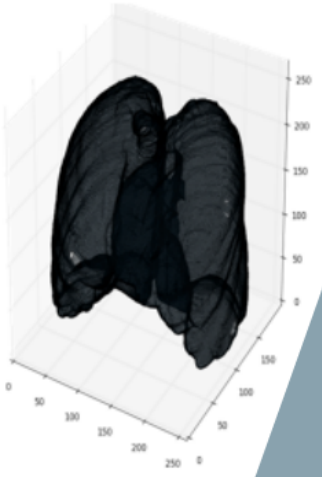
CT scans are broken into sub-images, and the nodule finder identifies regions that appear abnormal



3. Predi

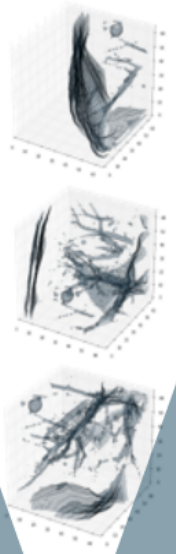
Predicts
maligna
spiculati
discover
assess

Raw input:
CT scan



Normalize

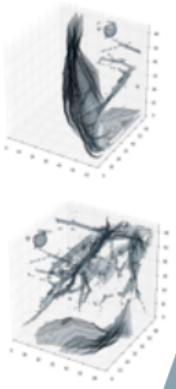
New input:
3D chunks



Nodule Finder

3D CONVNET
DETECT NODULES
+MALIGNANCY

New input:
Most probable candidates



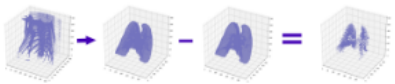
Attribute Predictor

ESTIMATE PROBABILITY CANCER
+attributes

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

Provided by Julian Dewit

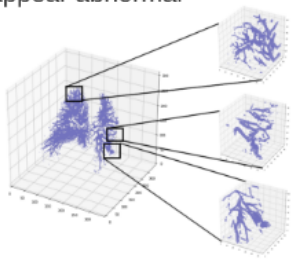
1. Normalize data



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

2. Find Tumors

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

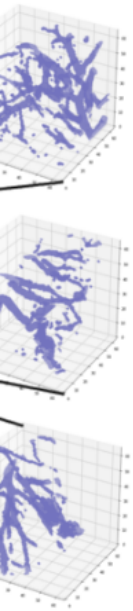


3. Predict Cancer Probability

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

ors

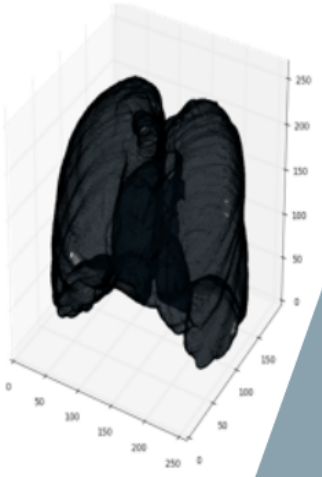
sub-images,
tifies regions



3. Predict Cancer Probability

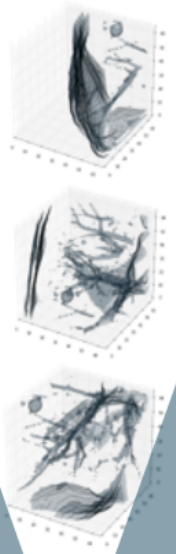
Predicts the attributes (such as malignancy, size, lobulation, and spiculation) on the regions discovered in the Nodule Finder to assess probability of cancer

Raw input:
CT scan



Normalize

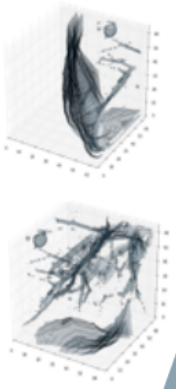
New input:
3D chunks



Nodule Finder

3D CONVNET
DETECT NODULES
+MALIGNANCY

New input:
Most probable candidates



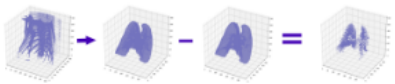
Attribute Predictor

ESTIMATE PROBABILITY CANCER
+attributes

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

Provided by Julian Dewit

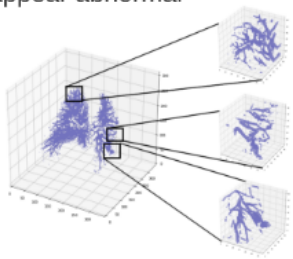
1. Normalize data



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

2. Find Tumors

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



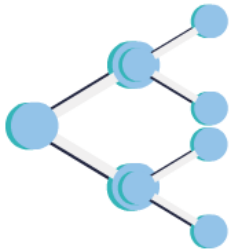
3. Predict Cancer Probability

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

Results

Human (trained professional)

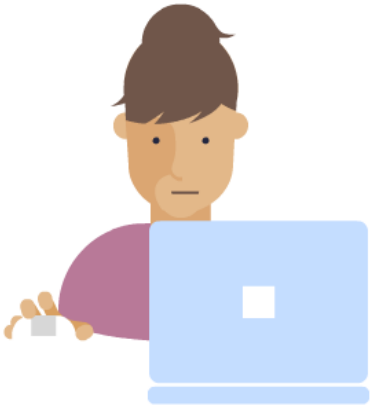
- Non-Cancerous Cases: 66%
- Cancerous Cases: 89%



Convolutional Neural Network

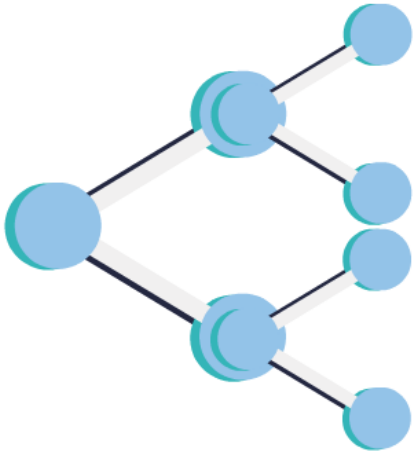
- ✔ Non-Cancerous Cases: 69%
- ✔ Cancerous Cases: 93%

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

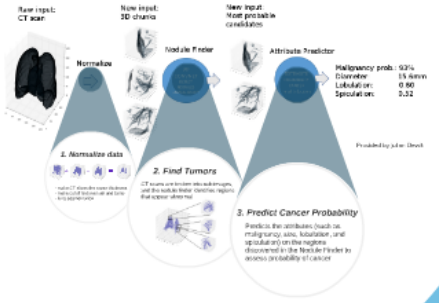
Human (trained professional)

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

Convolutional Neural Network

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

How it Works

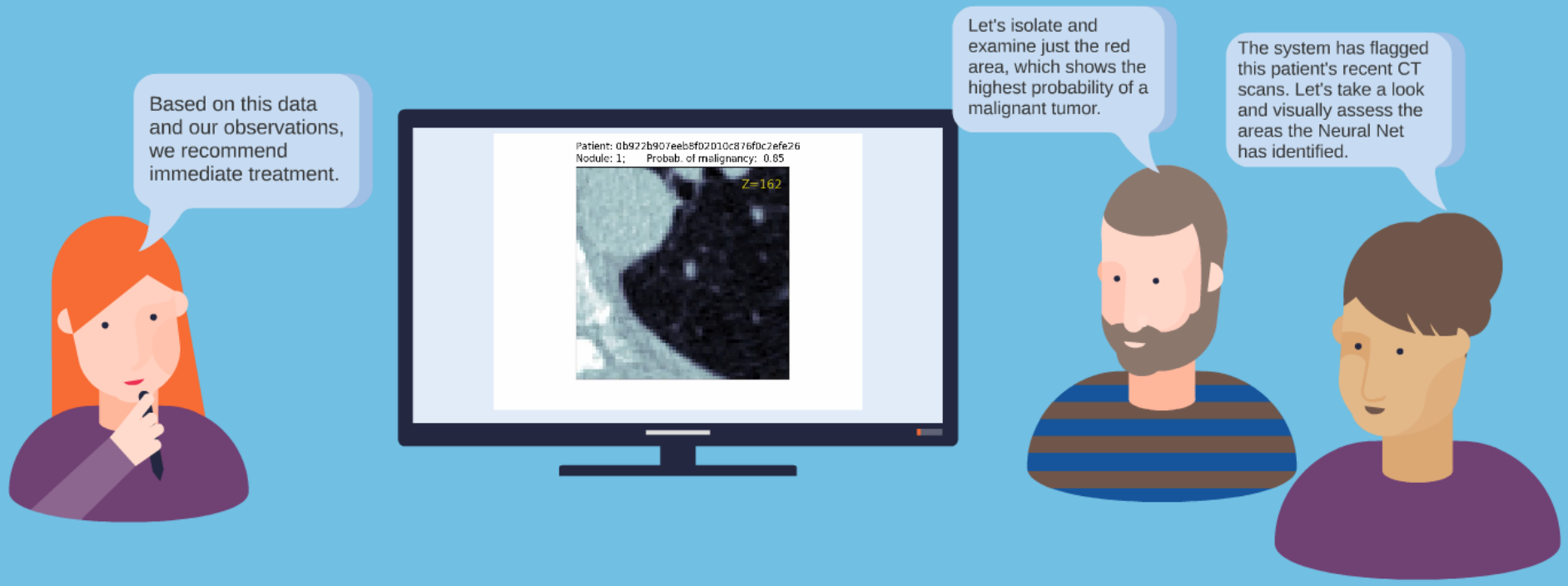


Project Approach

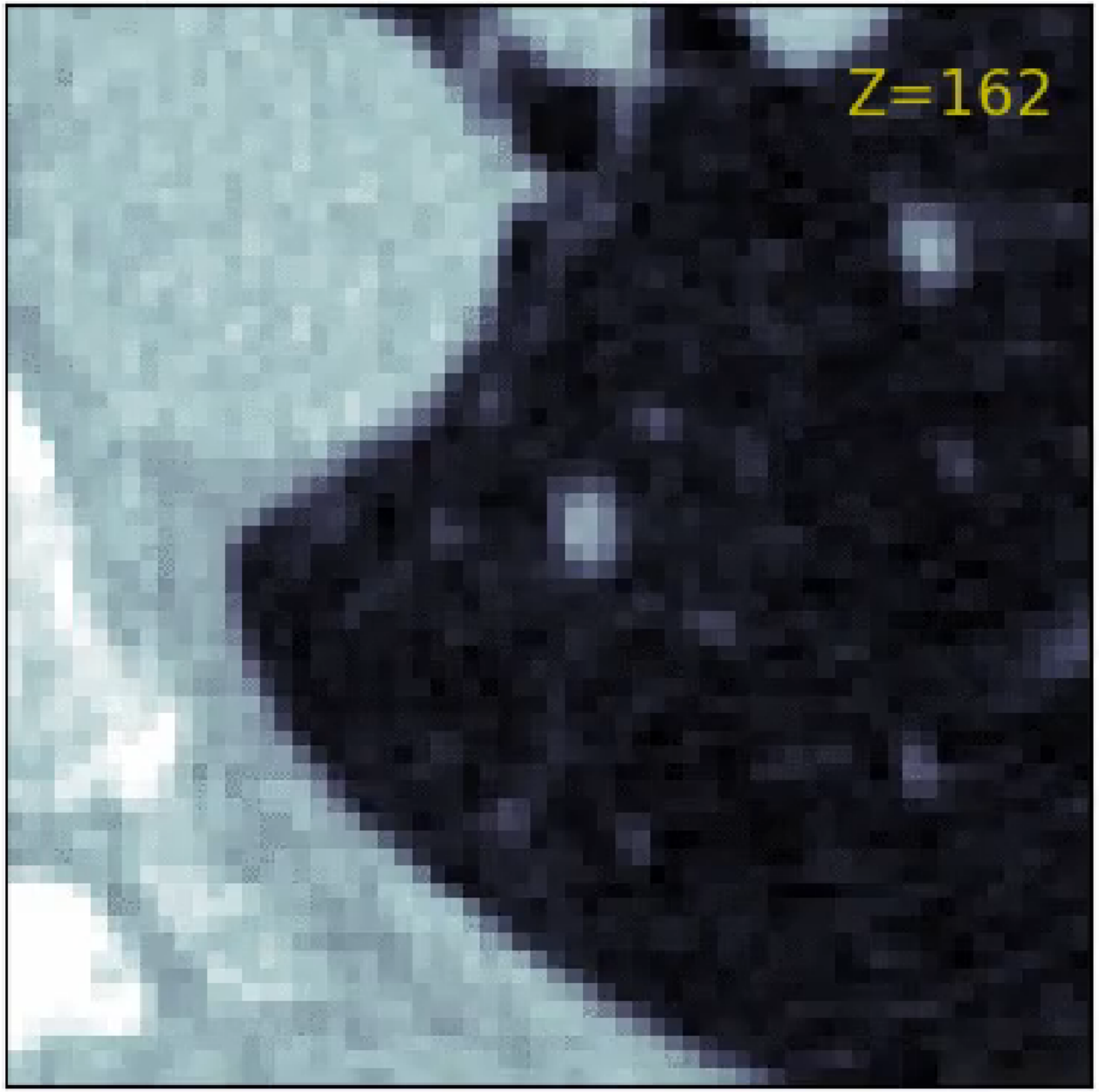
1. Understand the methods and reproduce results from Julian and Daniel's model
2. Develop a simple interface to create a diagnostic tool for doctors and radiologists

kaggle

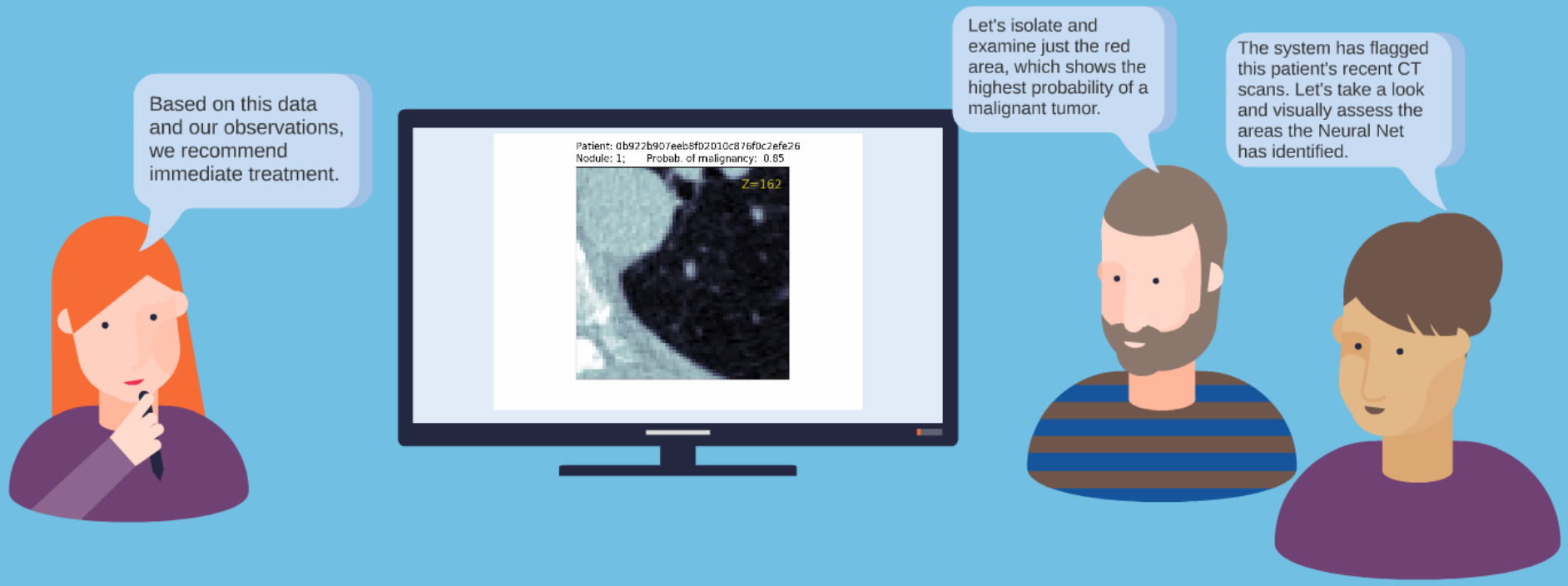
Demo & Use Case



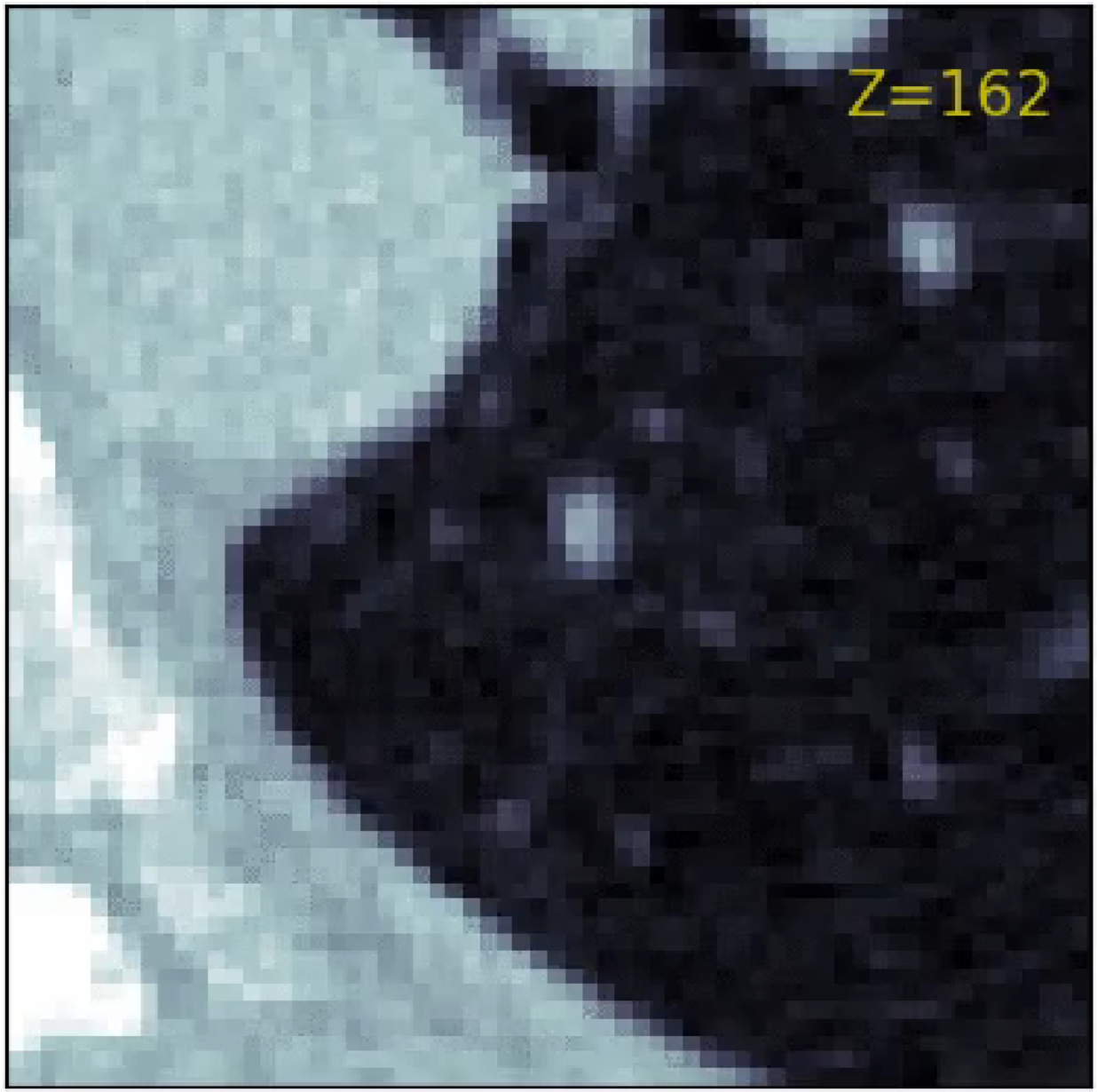
Patient: 0b922b907eeb8f02010c876f0c2efe26
Nodule: 1; Probab. of malignancy: 0.85



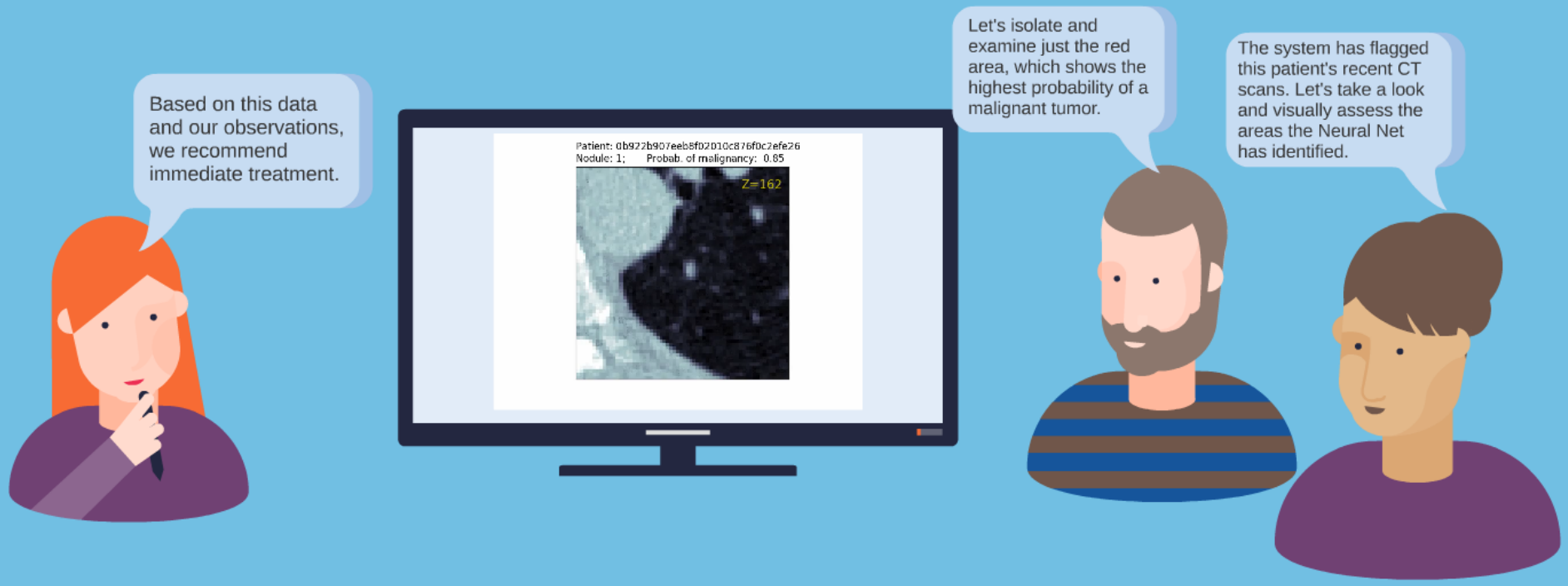
Demo & Use Case



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



Demo & Use Case



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 sucessful outcomes for data-driven, AI-powered projects.


Project Overview
Solution Overview
Demo & Use Case
Lessons
Machine Learning Demo:
Medical Imaging Project

New Technologies for
Canadian Observatories

12.15.2017 AGM

LimbicConsulting