Chest X-ray Abnormalities Detection

수학과 오서영

Overview

Introduction

Chest radiograph is difficult task for radiologist.
The interpretation of chest X-rays can lead to medical misdiagnosis, even for the best practicing doctor.
Computer-aided detection and diagnosis systems would help reduce the pressure on doctors at metropolitan hospitals and improve diagnostic quality in rural areas.

1. Data Exploration and Visualization

```
print("Number of rows in train dataframe: {}".format(train.shape[0]))
print("Number of Unique images in train set: {}".format(train.image_id.nunique()))
print("Number of Classes: {}₩n".format(train.class_name.nunique()))
print("Class Names: {}".format(list(train.class_name.unique())))
```

```
Number of rows in train dataframe: 67914
Number of Unique images in train set: 15000
Number of Classes: 15

Class Names: ['No finding', 'Cardiomegaly', 'Aortic enlargement', 'Pleural thickening',
'ILD', 'Nodule/Mass', 'Pulmonary fibrosis', 'Lung Opacity', 'Atelectasis', 'Other lesion',
'Infiltration', 'Pleural effusion', 'Calcification', 'Consolidation', 'Pneumothorax']
```

14 critical radiographic findings
+
'No finding'

```
0 - Aortic enlargement
1 - Atelectasis
2 - Calcification
3 - Cardiomegaly
4 - Consolidation
5 - ILD
6 - Infiltration
7 - Lung Opacity
8 - Nodule/Mass
9 - Other lesion
10 - Pleural effusion
11 - Pleural thickening
12 - Pneumothorax
13 - Pulmonary fibrosis
```

1. Data Exploration and Visualization

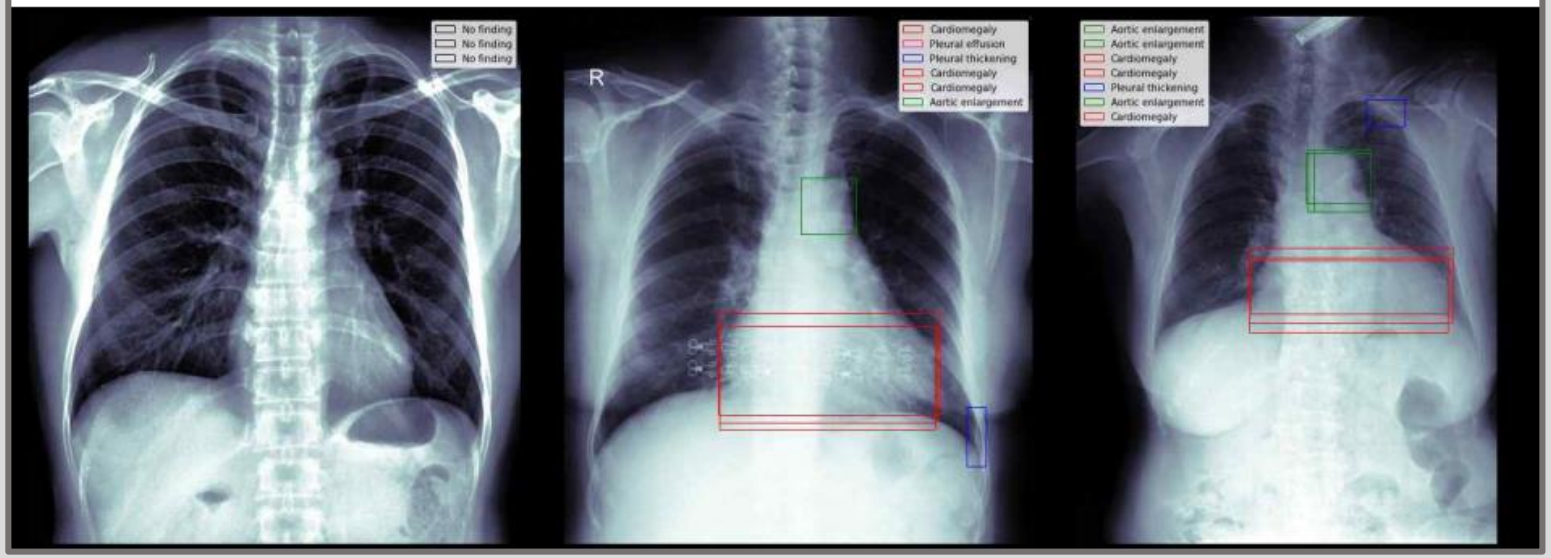
```
train.head()
```

	image_id	class_name	class_id	rad_id	x_min	y_min	x_max	y_max
0	50a418190bc3fb1ef1633bf9678929b3	No finding	14	R11	NaN	NaN	NaN	NaN
1	21a10246a5ec7af151081d0cd6d65dc9	No finding	14	R7	NaN	NaN	NaN	NaN
2	9a5094b2563a1ef3ff50dc5c7ff71345	Cardiomegaly	3	R10	691.0	1375.0	1653.0	1831.0
3	051132a778e61a86eb147c7c6f564dfe	Aortic enlargement	0	R10	1264.0	743.0	1611.0	1019.0
4	063319de25ce7edb9b1c6b8881290140	No finding	14	R10	NaN	NaN	NaN	NaN

(x_min, y_min, x_max, y_max) : **bounding box**
We have to predict (class, confidence, x_min, y_min, x_max, y_max)

1. Data Exploration and Visualization

Original



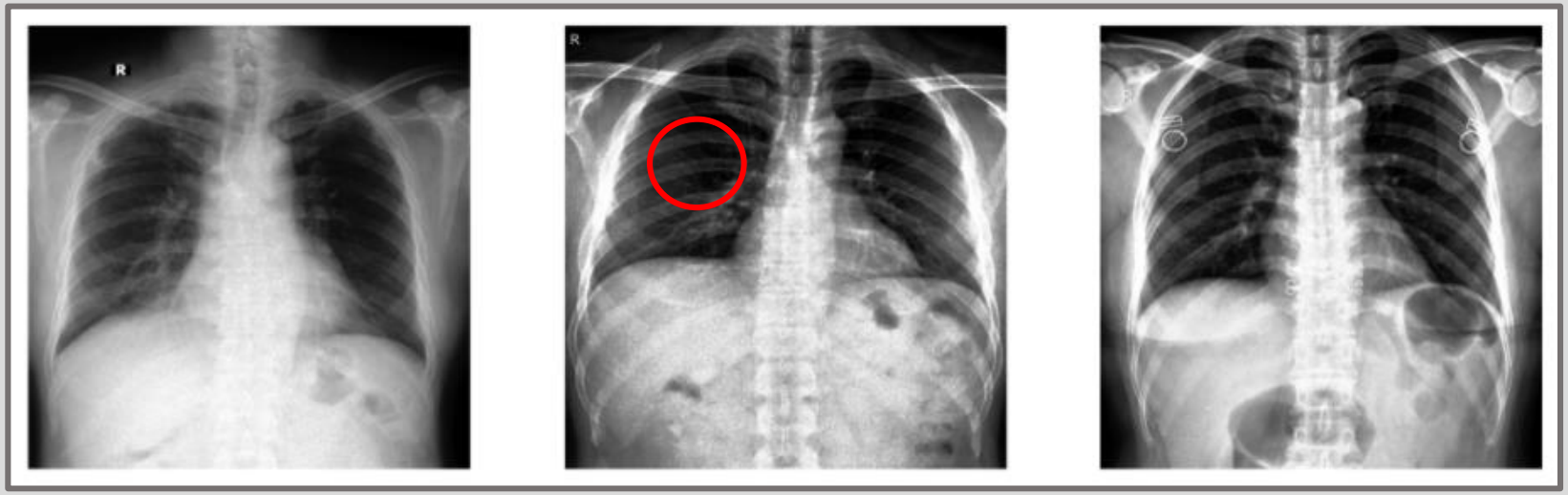
Bounding Box

2. Data Preprocessing

Original

Invert : 255 - img

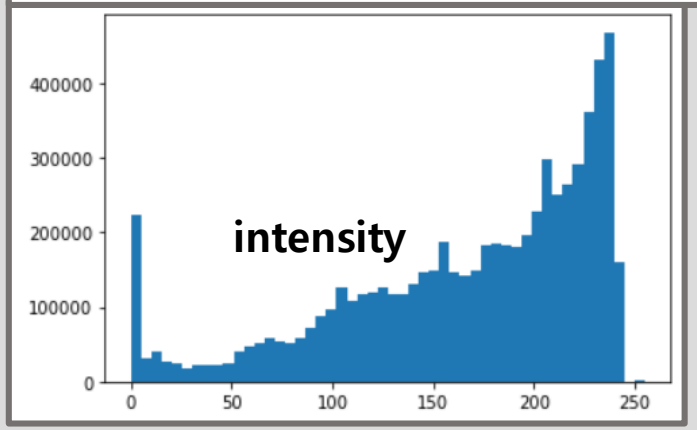
2. Data Preprocessing

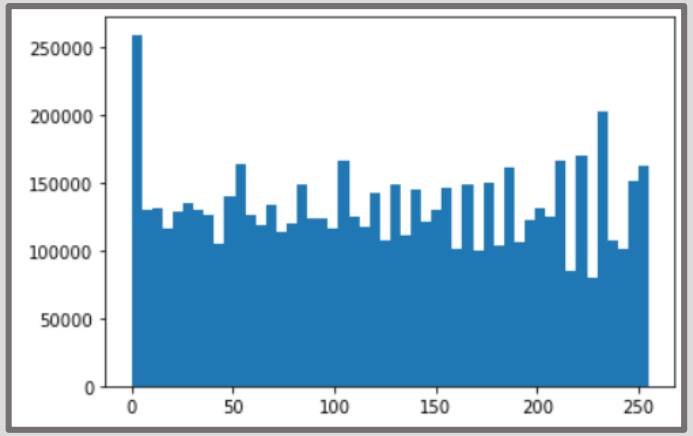


Histogram Equalization

The contrast increases.
Details of the dark regions appear clearly.

```
hist_eq = [cv2.equalizeHist(img) for img in imgs]
```



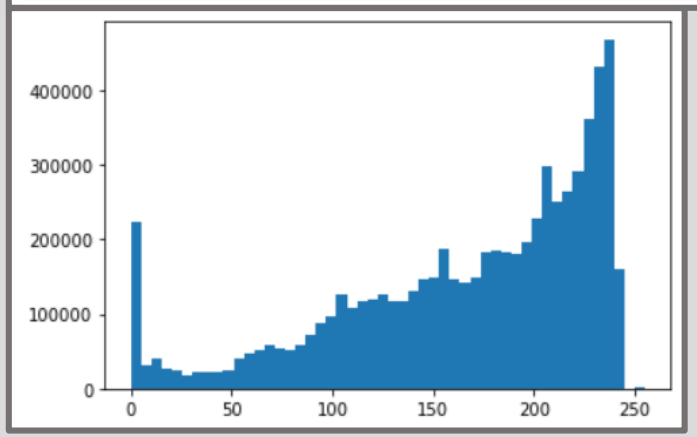


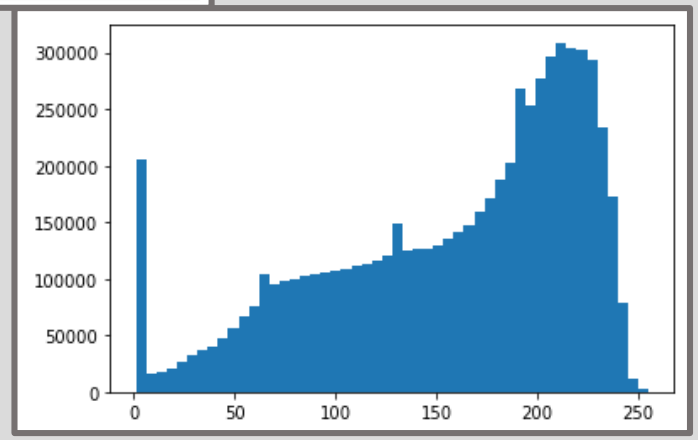
2. Data Preprocessing

Contrast Limited Adaptive Histogram Equalization

Image is divided into small blocks called **"tiles"**

```
clahe = cv2.createCLAHE(clipLimit = 2., tileGridSize = (10, 10))
```





3. Modeling – Baseline CNN

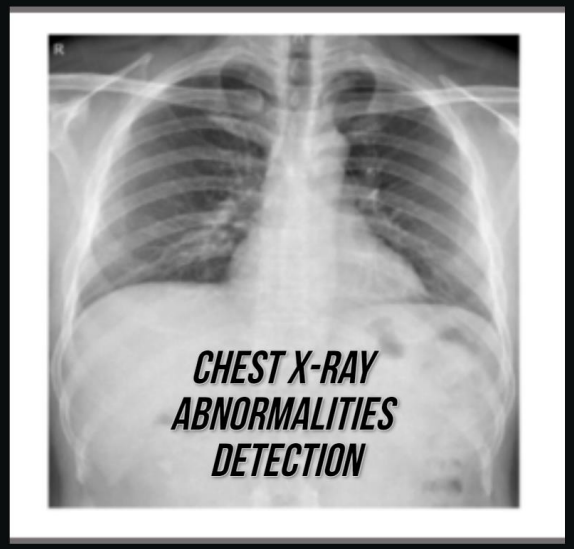
```
in1 = layer.Input(shape=(256,256,1))

out1 = layer.Conv2D(32,(3,3),activation="relu")(in1)
out1 = layer.Conv2D(32,(3,3),activation="relu")(out1)
out1 = layer.MaxPooling2D((2,2))(out1)

out1 = layer.Conv2D(64,(3,3),activation="relu")(out1)
out1 = layer.Conv2D(64,(3,3),activation="relu")(out1)
out1 = layer.MaxPooling2D((2,2))(out1)

out1 = layer.Conv2D(128,(3,3),activation="relu")(out1)
out1 = layer.Conv2D(128,(3,3),activation="relu")(out1)
out1 = layer.MaxPooling2D((2,2))(out1)
out1 = layer.Flatten()(out1)

out2 = layer.Dense(50,activation="relu",kernel_initializer="lecun_normal")(out1)
out2 = layer.Dense(30,activation="relu",kernel_initializer="lecun_normal")(out2)
out2 = layer.Dense(15,activation="sigmoid",kernel_initializer="lecun_normal",name='class_out')(out2)

out3 = layer.Dense(50,activation="relu",kernel_initializer="lecun_normal")(out1)
out3 = layer.Dense(30,activation="relu",kernel_initializer="lecun_normal")(out3)
out3 = layer.Dense(56,activation="relu",kernel_initializer="lecun_normal",name="box_out")(out3)

model = tf.keras.Model(inputs=in1,outputs=[out2,out3])
model.compile(loss={'class_out':'categorical_crossentropy','box_out':'mse'},optimizer="adam")
```

Class, Confidence

Bounding Box

mean average precision
0.05

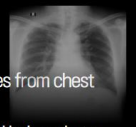
Infographic



인공지능
흉부 X선 진단

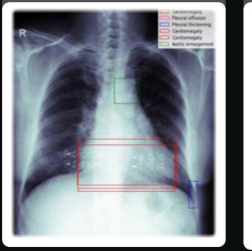
Introduction

Chest radiograph



- Automatically localize and classify thoracic abnormalities from chest radiographs
- Computer-aided detection and diagnosis systems would help reduce the pressure on doctors at metropolitan hospitals and improve diagnostic quality in rural areas.

Bounding Box



15000 Training Images

Labeling

```
0 - Aortic enlargement
1 - Atelectasis
2 - Calcification
3 - Cardiomegaly
4 - Consolidation
5 - ILD
6 - Infiltration
7 - Lung Opacity
8 - Nodule/Mass
9 - Other lesion
10 - Pleural effusion
11 - Pleural thickening
12 - Pneumothorax
13 - Pulmonary fibrosis
```

15 Classes

Data Preprocessing

Medical Image Analysis

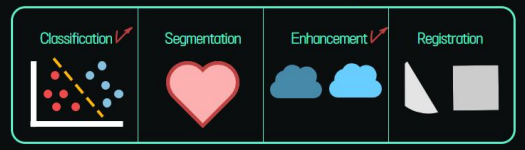


Image Enhancement

Invert | Histogram Equalization | CLAHE

Modeling

1 Conv Layer

Flatten

2 Dense sigmoid | 3 Dense relu

01 **Convolutional Layer**
이미지 특징 추출

02 **Class out**
Sigmoid 확률 계산 및 분류

03 **Bounding box out**
Bounding box 좌표 예측

CNN 기반 객체 탐지 진행

Kaggle Competition

[1] VinBigData Chest X-ray Abnormalities Detection,
https://www.kaggle.com/c/vinbigdata-chest-xray-abnormalities-detection/data

References

[1] Chest X-ray abnormalities: Baseline[TF.Keras],
https://www.kaggle.com/bibhash123/chest-x-ray-abnormalities-baseline-tf-keras/comments

[2] x-ray image Enhancement test,
https://www.kaggle.com/kuuuuub/x-ray-image-enhancement-test