```
# FDA  Submission

**Your Name:** Rob Straker

**Name of your Device:** Convolutional Neural Netwo

## Algorithm Description

### 1. General Information

**Intended Use Statement:** This algorithm is inten

**Indications for Use:** This algorithm is indicate

**Device Limitations:** The algorithm performs mode

**Clinical Impact of Performance:** The algorithm i


### 2. Algorithm Design and Function

![Algorithm Flowchart](/algorithm_flowchart.png)

**DICOM Checking Steps:** The following checks wre
1. Confirm body part is 'Chest'.
2. Confirm modality is 'DX', which stands for Digit
3. Confirm patient position is either 'AP' (Anterio

**Preprocessing Steps:** Preprocessing was complete

```

```
*Metadata:*
1. Column Creation. Split "Finding Labels" column
2. Pneumonia Class. Create a column for 'Pneumonia

*Images:*
1. Image resizing. Resize images to consistent squ
2. Image normalization. Subtract mean intensity fr

**CNN Architecture:** A pre-trained VGG-16 model w

![Pre-Trained VGG-16 Model](/architecture_vgg16.pn

![Fully Connected Layer Extension](/architecture_c


### 3. Algorithm Training

**Parameters:**
* Types of augmentation used during training:
    - rescale=1. / 255.0
    - horizontal_flip = True
    - vertical_flip = False
    - height_shift_range= 0.1
    - width_shift_range=0.1
    - rotation_range=20
    - shear_range = 0.1
    - zoom_range=0.1)

* Batch size = 64

```

```
61 * Optimizer learning rate = 1.0e-6
62
63 * Layers of pre-existing architecture that were fr
64     - block1_conv1 False
65     - block1_conv2 False
66     - block1_pool False
67     - block2_conv1 False
68     - block2_conv2 False
69     - block2_pool False
70     - block3_conv1 False
71     - block3_conv2 False
72     - block3_conv3 False
73     - block3_pool False
74     - block4_conv1 False
75     - block4_conv2 False
76     - block4_conv3 False
77     - block4_pool False
78     - block5_conv1 False
79     - block5_conv2 False
80
81 * Layers of pre-existing architecture that were fi
82     - block5_conv3 True
83     - block5_pool True
84
85 * Layers added to pre-existing architecture:
86
87     Model: "sequential_2"
88     _________________________________________________
89     Layer (type)                 Output Shape
90
```

```
    =================================================================
    model_2 (Model)              (None, 7, 7, 512)
    _________________________________________________________________
    flatten_2 (Flatten)          (None, 25088)
    _________________________________________________________________
    dropout_4 (Dropout)          (None, 25088)
    _________________________________________________________________
    dense_4 (Dense)              (None, 1024)
    _________________________________________________________________
    dropout_5 (Dropout)          (None, 1024)
    _________________________________________________________________
    dense_5 (Dense)              (None, 512)
    _________________________________________________________________
    dropout_6 (Dropout)          (None, 512)
    _________________________________________________________________
    dense_6 (Dense)              (None, 1)
    
    =================================================================
    Total params: 40,931,137
    Trainable params: 28,576,257
    Non-trainable params: 12,354,880
    _________________________________________________________________


![Algorithm Training Performance History](/trainin

![Precision-Recall Curve](/p_r_curve.png)

![AUC Curve](/auc_curve.png)

```

```
121 ![F1 Threshold Curve](/f1_thresh_curve.png)
122
123 **Final Threshold and Explanation:**
124 The final threshold was determined by finding that
125
126 Final Threshold = 0.38
127
128
129 ### 4. Databases
130
131 **Description of Training Dataset:**
132 The dataset was split 80:20, with 80% of images as
133
134 **Description of Validation Dataset:**
135 Efforts were made to ensure representative proport
136
137
138 ### 5. Ground Truth
139
140 The ground truth for this algorithm is the Silver
141
142
143 ### 6. FDA Validation Plan
144
145 **Patient Population Description for FDA Validatio
146
147 **Ground Truth Acquisition Methodology:** Recruit
148
149 **Algorithm Performance Standard:** The key metric
150 - *F1 = 2 * (precision x recall) / (precision + re
```

- Also known as the harmonic mean of precision and
- This is based on real radiologist performance.