



CHESTLENS AI

Testing Document

Group 18

Supervisor	
Name: Email: Organization:	Dr. Mehdi Moradi moradm4@mcmaster.ca McMaster University, CAS
Team Member 1	
Name: Email: Student Number:	Abhishek Sharma shara109@mcmaster.ca 400322503
Team Member 2	
Name: Email: Student Number:	Anthony Vu vua11@mcmaster.ca 400306059
Team Member 3	
Name: Email: Student Number:	Hussein Saad saadh@mcmaster.ca 400307995
Team Member 4	
Name: Email: Student Number:	Nathan Starr starrn@mcmaster.ca 400323095
Team Member 5	
Name: Email: Student Number:	Yuvraj Jain jainy3@mcmaster.ca 400259484

Table of Contents:

Frontend Testing.....	2
Backend Testing.....	5
Average Runtime for Backend.....	5
Key Testing Terms.....	5
Our Model Validation Results.....	6
Our Model Testing Results.....	7
Expected AUCs vs Actual AUCs For Our Model.....	8
Individual Image Testing For Our Model.....	8
Testing Results for Atelectasis.....	8
Analysis of Atelectasis Highlighting.....	12
Testing Results for Cardiomegaly.....	12
Analysis of Cardiomegaly Highlighting.....	15
Testing Results for Consolidation.....	15
Analysis of Consolidation Highlighting.....	19
Testing Results for Edema.....	19
Analysis of Edema Highlighting.....	22
Testing Results for No Finding.....	22
Analysis of No Finding Highlighting.....	27
Testing Results for Pleural Effusion.....	28
Analysis of Pleural Effusion Highlighting.....	30
Single Image Testing Analysis.....	30
Pre-Trained Model Validation Results.....	31
Pre-Trained Model Testing Results.....	32

Frontend Testing

In our System Verification and Validation Plan, we mentioned using the unittest library which was used for our login, doctor, and patient/index pages. In particular, we wanted to extensively test the functions related to uploading images and running the model. For performance testing, more details on the model itself are part of our backend testing, but for the end user, results are displayed within 5-10 seconds (within our initial goal of 10-15 seconds).

Tested fetching an image (fetchimage route).

```
def test_fetch_image(self):
    response = self.app.get('/fetchimage')
    self.assertEqual(response.status_code, 200)
    self.assertEqual(response.get_json(), {"filename": "image.jpg"})
```

Tested deleting an image (deleteimage route).

```
def test_delete_image(self):
    response = self.app.get('/deleteimage')
    self.assertEqual(response.status_code, 200)
    self.assertEqual(response.get_json(), {"true": True})
```

The following tests are for the login page. As stated in the V&V plan, we wanted to ensure that only users with valid credentials would be led to the doctor's page.

```

# Testing for login() or base route ('/')
def test_login_get_request(self):
    response = self.app.get('/')
    self.assertEqual(response.status_code, 200)
    self.assertIn(b'Login', response.data)

def test_login_post_request_existing_doctor(self):
    response = self.app.post('/', data={'username': 'doctor3@gmail.com', 'password': 'password3'})
    self.assertEqual(response.status_code, 302)

def test_login_post_request_non_existent_doctor(self):
    response = self.app.post('/', data={'username': 'doctor2@gmail.com', 'password': 'password3'})
    self.assertEqual(response.status_code, 200)
    self.assertNotIn(b'Invalid username or password', response.data)

```

The following tests are for the doctor's page.

```

# Testing for doctor() or route ('/doctor')
def test_doctor_authenticated(self):
    with self.app as client:
        with client.session_transaction() as sess:
            sess['d_id'] = 1
        response = client.get('/doctor')
        doctor_id = "d1"
        patients = Patient.query.filter_by(d_id=doctor_id).all()
        self.assertEqual(response.status_code, 200)
        self.assertEqual(response.content_type, 'text/html; charset=utf-8')
        self.assertTrue(patients)
        self.assertIn(b'List of Patients', response.data)

def test_doctor_unauthenticated(self):
    with self.app as client:
        response = client.get('/doctor')
        self.assertEqual(response.status_code, 302)
        self.assertEqual(response.location, '/')

```

The following tests are for our comments feature.

```

# Testing for comments() or route ('/comments')
def test_comments_update_comment(self):
    with self.app as client:
        response = client.post('/comments', json={'scan_id': '50712315', 'textarea_content': 'Updated comment'})
        scan = NewScan.query.filter_by(s_id='50712315').first()
        self.assertEqual(response.status_code, 200)
        self.assertEqual(response.content_type, 'application/json')
        self.assertIsNotNone(scan)
        self.assertEqual(scan.s_comment, 'Updated comment')

def test_comments_missing_data(self):
    with self.app as client:
        response = client.post('/comments', json={})
        self.assertEqual(response.status_code, 200)
        self.assertEqual(response.content_type, 'application/json')
        response_data = response.get_json()
        self.assertIn('Error', response_data)

```

The following tests are for the index/patient's page.

```
# Testing for index(patient_id) or route ('/index/<patient_id>')
def test_index_patient_found(self):
    with self.app as client:
        response = client.get('/index/17007063')
        patient = Patient.query.get('17007063')
        scans = NewScan.query.filter_by(p_id='17007063').all()
        self.assertEqual(response.status_code, 200)
        self.assertEqual(response.content_type, 'text/html; charset=utf-8')
        self.assertTrue(patient)
        self.assertTrue(scans)
        self.assertIn(b'Patient Name', response.data)
        self.assertIn(b'Select Previous Scans', response.data)

def test_index_patient_not_found(self):
    with self.app as client:
        response = client.get('/index/999')
        self.assertEqual(response.status_code, 302)
        self.assertEqual(response.location, '/doctor')
```

Tested uploading an image for the prebuilt model.

```
# Testing for upload() or route ('/upload')
def test_upload_new_image(self):
    with open('../demo_images/Steven.dcm', 'rb') as img:
        data = {
            'xrayImage': (img, 'Steven.dcm'),
            'uploadedImages': '17007063'
        }
        response = self.app.post('/upload', content_type='multipart/form-data', data=data)
    self.assertEqual(response.status_code, 200)
    response_data = json.loads(response.data)
    self.assertEqual(response_data['Study_ID'], '50712315')
    self.assertEqual(response_data['Patient_ID'], '17007063')
    self.assertEqual(response_data['Patient_Name'], 'Steven Sanders')
    self.assertEqual(response_data['Patient_Sex'], 'M')
    self.assertEqual(response_data['Patient_Birth_Date'], '1989-01-19')
    self.assertEqual(response_data['Acquisition_Date'], '2009-11-01')
    self.assertEqual(response_data['View_Position'], 'AP')
    self.assertEqual(response_data['Patient_Age_at_Time_of_Acquisition'], '20 years, 9 months, 13 days')
    self.assertEqual(response_data['comment'], 'Updated comment')
```

Tested uploading an image for our model.

```
# Testing for upload_our_model() or route ('/upload-our-model')
def test_upload_our_model_new_image(self):
    with open('../demo_images/Steven.dcm', 'rb') as img:
        data = {
            'xrayImage': (img, 'Steven.dcm'),
            'uploadedImages': '17007063'
        }
        response = self.app.post('/upload-our-model', content_type='multipart/form-data', data=data)
    self.assertEqual(response.status_code, 200)
    response_data = json.loads(response.data)
    self.assertEqual(response_data['Study_ID'], '50712315')
    self.assertEqual(response_data['Patient_ID'], '17007063')
    self.assertEqual(response_data['Patient_Name'], 'Steven Sanders')
    self.assertEqual(response_data['Patient_Sex'], 'M')
    self.assertEqual(response_data['Patient_Birth_Date'], '1989-01-19')
    self.assertEqual(response_data['Acquisition_Date'], '2009-11-01')
    self.assertEqual(response_data['View_Position'], 'AP')
    self.assertEqual(response_data['Patient_Age_at_Time_of_Acquisition'], '20 years, 9 months, 13 days')
    self.assertEqual(response_data['comment'], 'Updated comment')
```

Result of above tests:

```
Ran 13 tests in 10.273s  
OK
```

Note: The above tests inherently tests the JavaScript functions as well, hence adding npm for Jest was deemed unnecessary. As for load testing, we would ideally like to test for many concurrent uploads but due to limited storage on our hosted database, we can only store around 20-25 images.

Backend Testing

Average Runtime for Backend

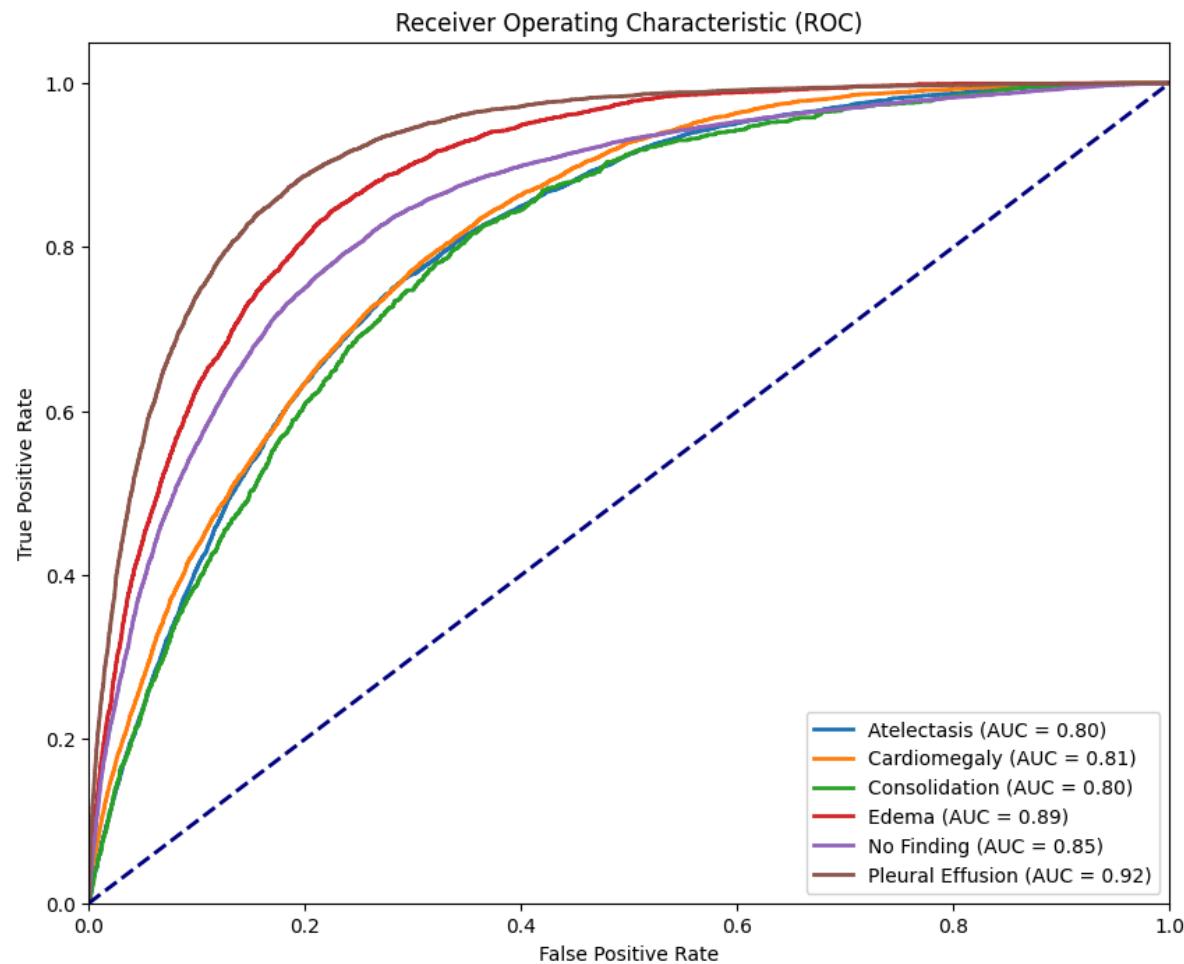
```
import time  
  
total_time = 0  
  
img = "/home/group18/Data/mimic/test/p18_p19/a73f67cb-036cb1ae-34b24fc2-725883fd-b5961466.jpg"  
  
for x in range (0,20):  
    start_time = time.time()  
    run_with_no_csv(img)  
    end_time = time.time()  
    total_time += end_time - start_time  
  
print("Average Time in Seconds", total_time / 20)  
✓ 9.7s  
  
Average Time in Seconds 0.48916534185409544
```

We used the Python Time library to test the run time of our model. We took an average of running each model 20 times on an image and we got an average runtime of 0.4892 seconds for our model and 1.1268725633621215 for the pre-trained model. This is under our requirement of 10 seconds.

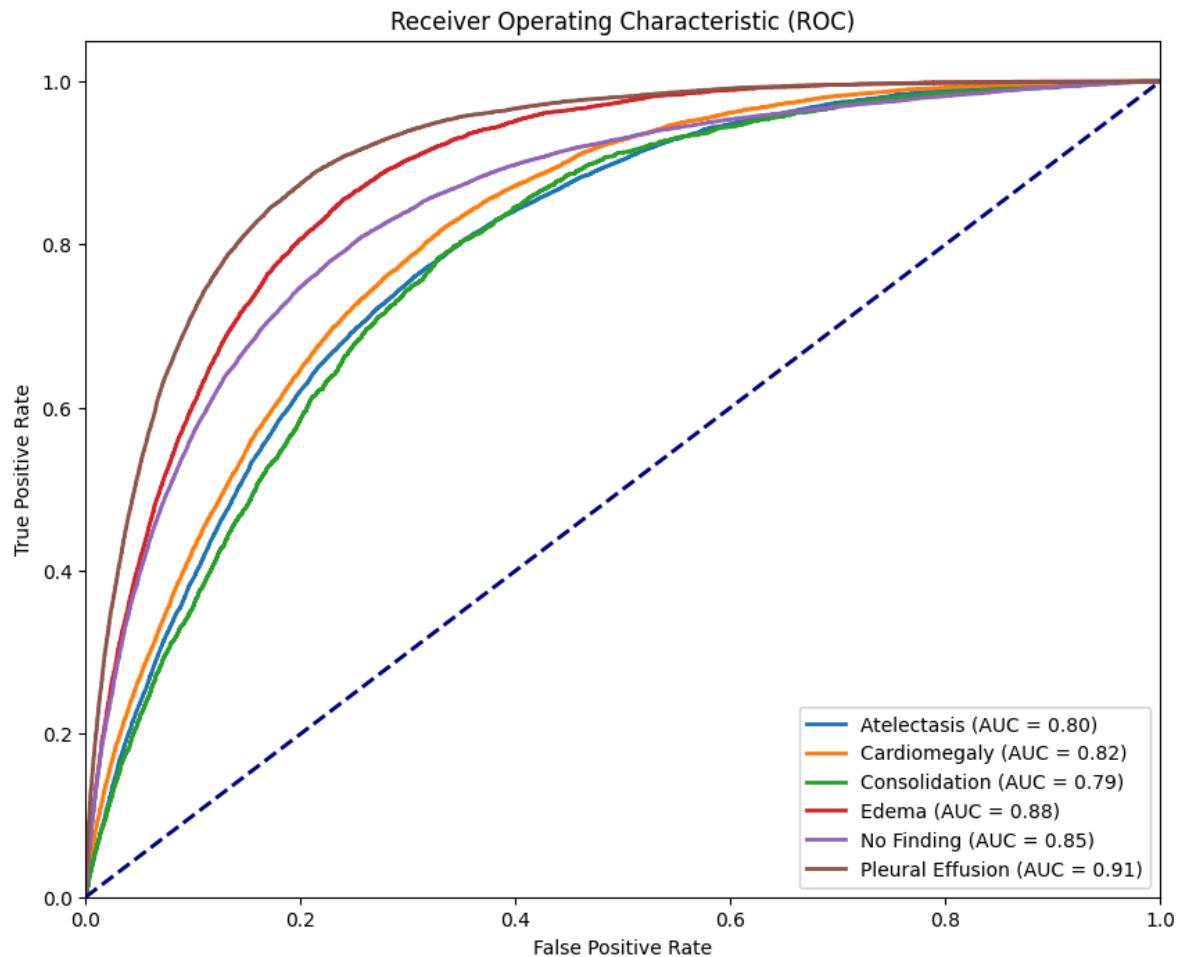
Key Testing Terms

- **ROC:** This chart shows how well a model can tell the difference between two groups (like "present" and "not present") by displaying its accuracy under various conditions.
- **AUC:** Measures the area under the ROC curve and it shows a model's ability to differentiate between categories. A higher value indicates a better performance.
- **Best Threshold:** This represents the optimal point for balancing true and false positives, enhancing model accuracy.
- **Precision:** This measures the accuracy of the model in correctly identifying "positive" cases, focusing on whether the model's "yes" answers are actually correct.
- **Recall:** This assesses the model's ability to capture all the "positive" cases, emphasizing the importance of not overlooking any "yes" instances.
- **F1 Score:** A metric that merges precision and recall, providing a single value that reflects the model's effectiveness in accurately identifying positives without missing any.

Our Model Validation Results



Our Model Testing Results



	Atelectasis	Cardiomegaly	Consolidation	Edema	No Finding	Pleural Effusion
AUC	0.7969	0.8154	0.7893	0.8831	0.8483	0.9110
Best Threshold	0.1711	0.1580	0.0385	0.0980	0.3500	0.1623
Precision	0.3395	0.3372	0.0837	0.2730	0.6581	0.5287
Recall	0.7731	0.8190	0.7979	0.8564	0.8095	0.8859
F1 Score	0.4718	0.4777	0.1515	0.4140	0.7260	0.6622

For F1, precision, and recall we did not define any specific values for testing in our SRS document. That being said, our F1 scores are relatively low. F1 scores are dependent on Recall and Precision and based on these scores Precision is low which is bringing down the F1 score while recall values are high. This means that our model does well in identifying the positive cases, however it has a tendency to over predict some negative cases as positive. This is preferred for disease identification because falsely identifying a disease and following up is better than missing a disease. Because of this despite the relatively low F1 and precision we believe our model passes these tests due to its ability to correctly identify positive diseases.

Expected AUCs vs Actual AUCs For Our Model

Finding	AUC Defined in P0	Validation AUC	Final Testing AUC
Atelectasis	0.80	0.8029	0.7969
Cardiomegaly	0.85	0.8130	0.8154
Consolidation	0.85	0.7951	0.7893
Edema	0.85	0.8861	0.8831
No Finding	0.85	0.8487	0.8483
Pleural Effusion	0.92	0.9175	0.9110

The only AUC that surpasses our P0 requirement is Edema while Atelectasis, No Finding and Pleural Effusion are all very close. Also between validation and testing the AUC of Cardiomegaly increased while the others decreased by under a percent. This means our model does suffer from slight underfitting that is within reason.

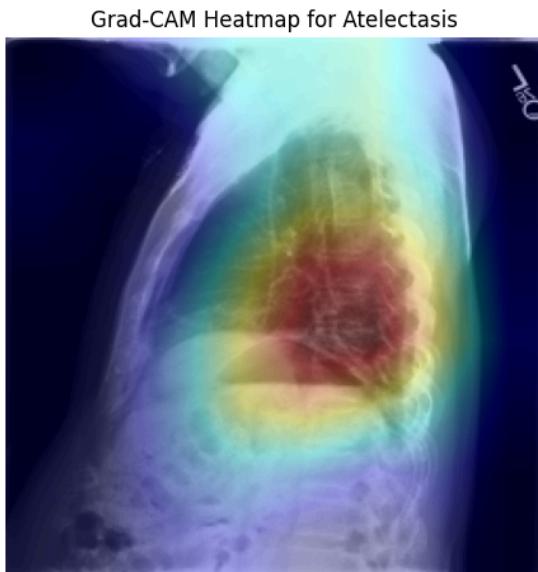
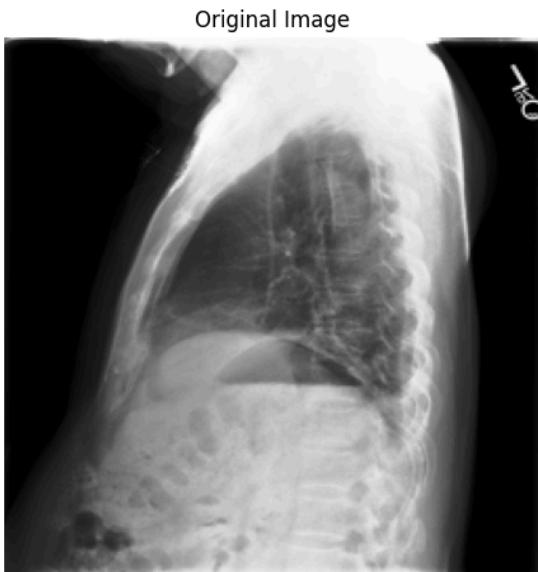
Individual Image Testing For Our Model

Testing Results for Atelectasis

Atelectasis Test 1

p18_p19/p19/p19109135/s51197801.txt

img = "20d9383c-3fa80c3c-94218c7f-15020bd1-e47ed769"

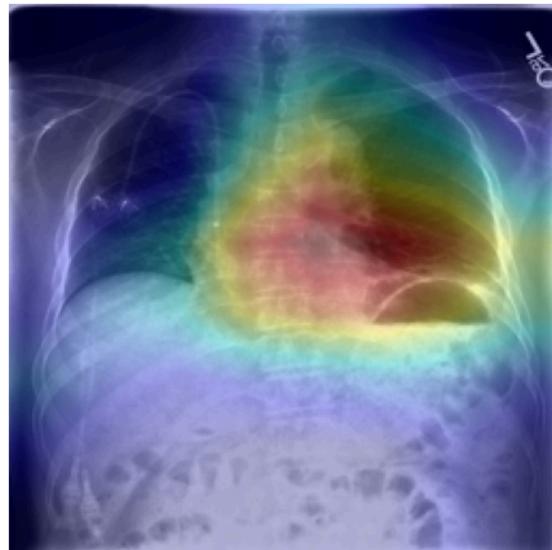


img = "ae4d45fc-815b6cd8-d29c078f-ad849410-cbb8cf47"

Original Image



Grad-CAM Heatmap for Atelectasis



Finding From Report:

- There is bronchovascular crowding in the left lower lobe consistent with atelectasis.
- IMPRESSION: Atelectasis in the left lower lobe.

Atelectasis Testing Image 2

p18_p19/p19/p19173988/s54718065.txt

img = "52b360b9-9fe4bb09-98bffb7f-3d8e4391-863e32e3"

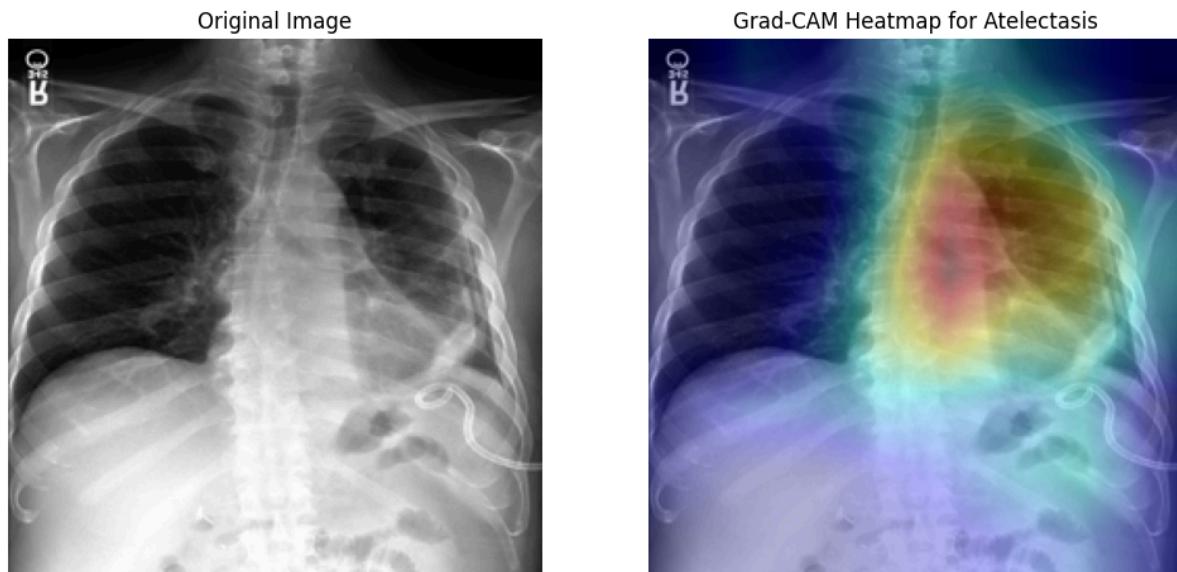
Original Image



Grad-CAM Heatmap for Atelectasis



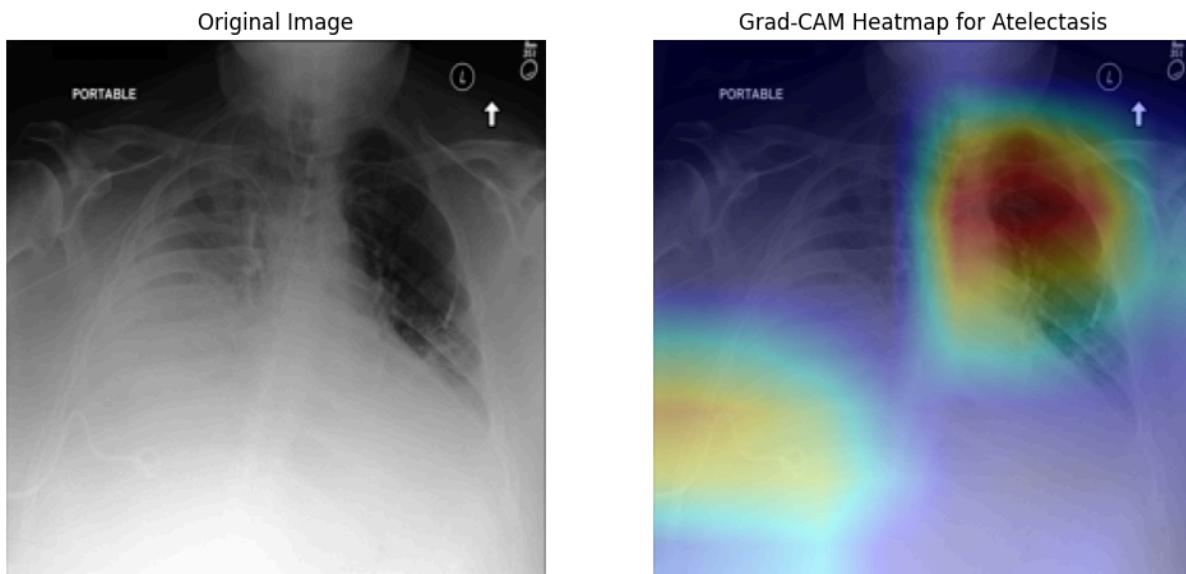
img = "f556ae7b-066cd0b1-954eaf59-4c30d965-f86260ff"



Finding From Report: This is also reflected in extensive atelectasis in the left mid and lower lung. The right lung is clear.

Atelectasis Testing Image 3
p18_p19/p19/p19806884/s59414102.txt

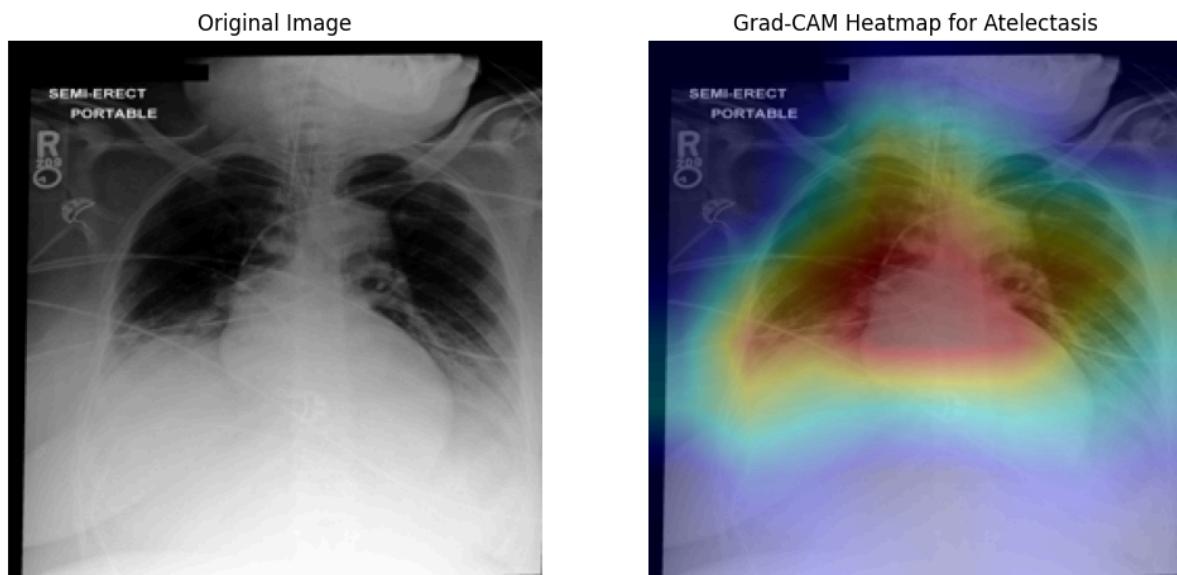
img = "d944bdaf-e3e6f59b-f08a84f6-86f07ba4-6c783f84"



Finding From Report: Finding from severe atelectasis in the right lung and some compromise of aeration in the left lower lobe as well.

Atelectasis Testing Image 4
p18_p19/p18/p18741255/s51561605.txt

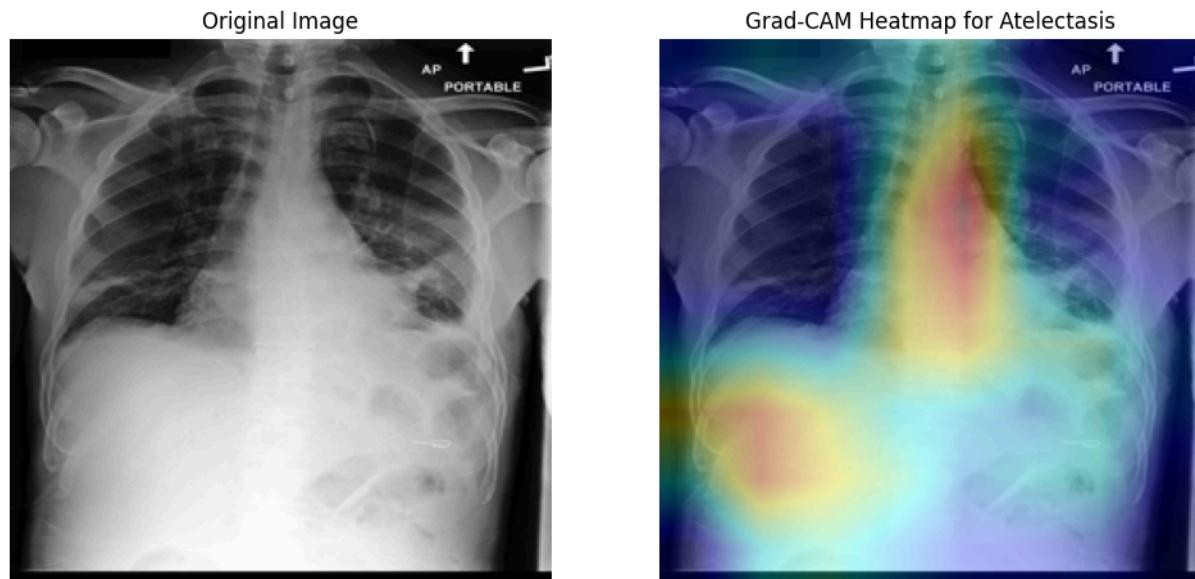
img = "34589968-8b864391-ae8fda46-91ab2791-4daf87af"



Finding From Report: Atelectasis in the right lower lobe is still severe, minimally improved.

Atelectasis Testing Image 5
p18_p19/p19/p19229277/s53608284.txt

img = "cf875bfb-1ea9aff7-cb1c26b0-e1598bfb-b193ff62"



Finding From Report: Moderate to severe left basal atelectasis has worsened. Atelectasis in the right lower lung is also more pronounced.

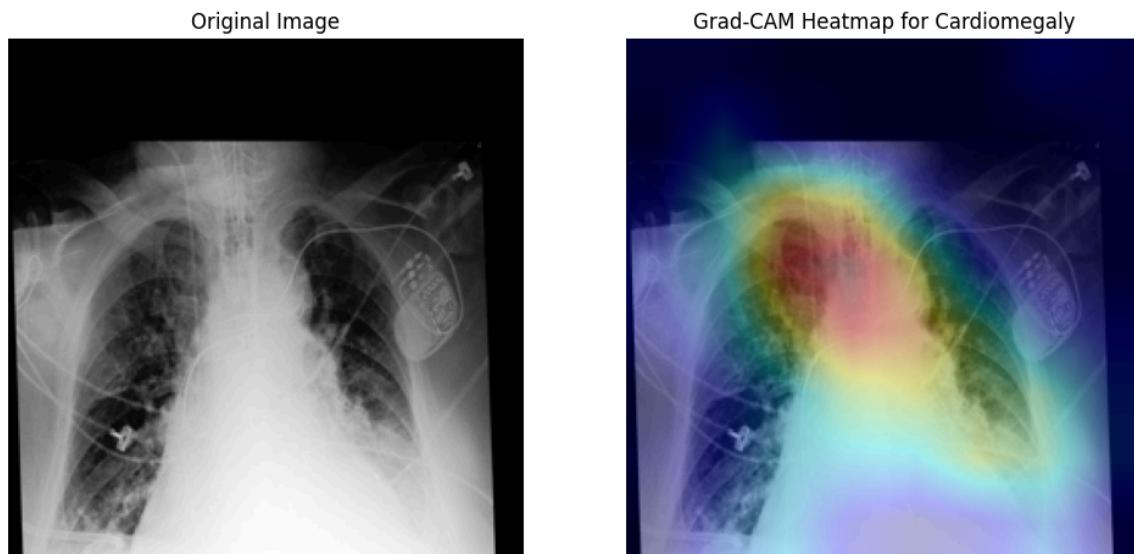
Analysis of Atelectasis Highlighting

Overall the highlighting produced for Atelectasis is generally in the correct area with image 3 being the only notable exception. Some of the highlights are either higher or lower than expected while still covering the correct area.

Testing Results for Cardiomegaly

Cardiomegaly Testing Image 1
p18_p19/p18/p18396526/s52262111.txt

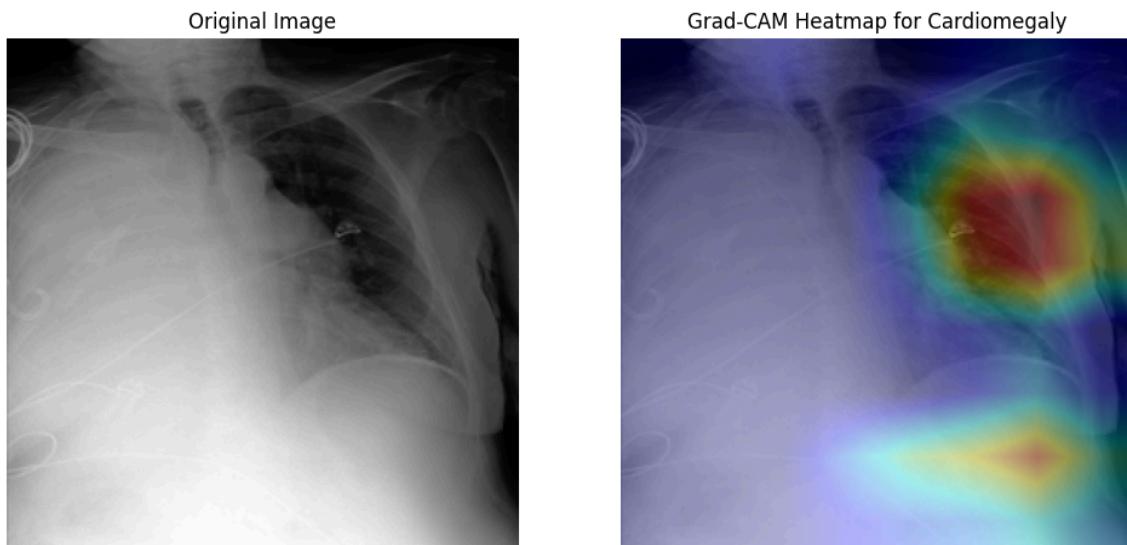
img = 3f2023b8-39fdbbf0-56b564b1-725766fd-4521856f



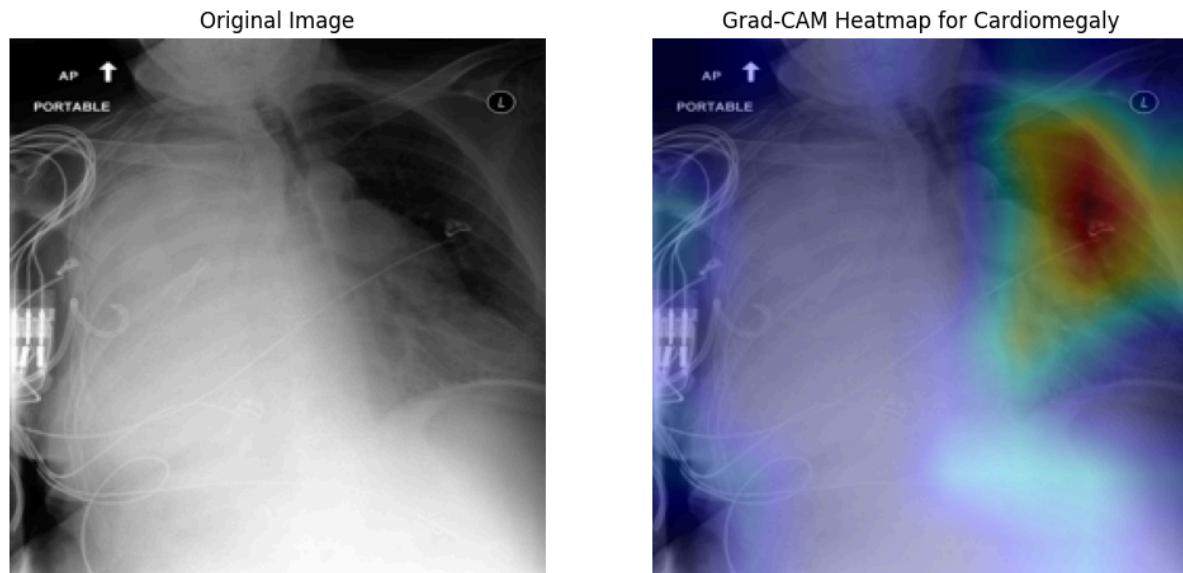
Finding From Report: severe chronic cardiomegaly unchanged.

Cardiomegaly Testing Image 2
p18_p19/p18/p18712968/s58288058.txt

img = "01a9c196-4d0a3697-86bc7015-19c697c3-8a6fd8ea"



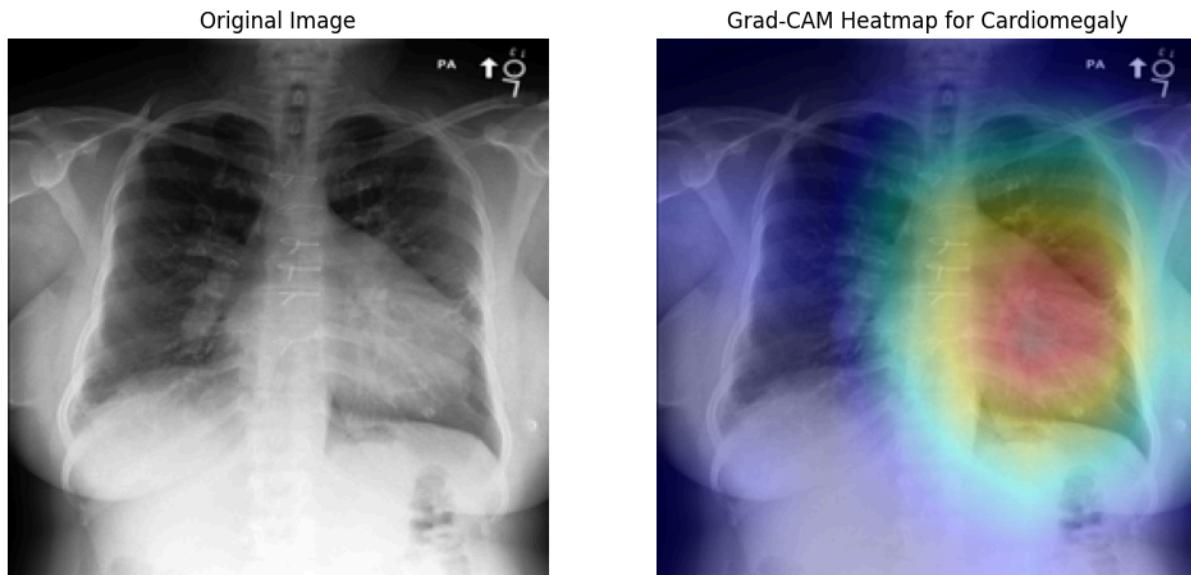
img = "e29e98e9-2c21732f-ce097e4f-90427d27-6af093c"



Finding From Report: No reason to suspect that moderate chronic cardiomegaly has improved.

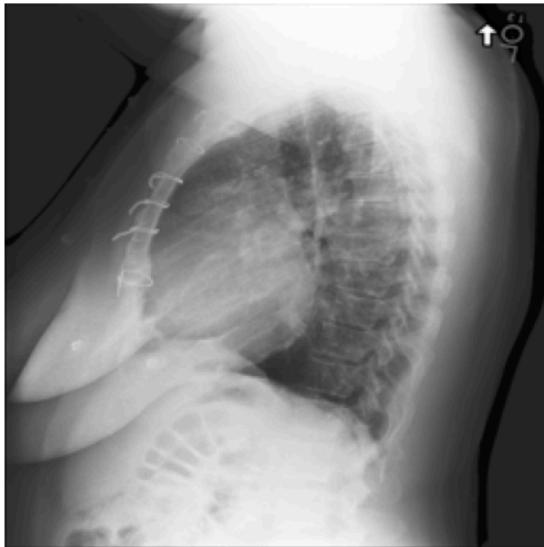
Cardiomegaly Testing Image 3
p18_p19/p19/p19807980/s57843918.txt

img = 379d3644-53dcc490-f52c384f-4905bd21-9813c730

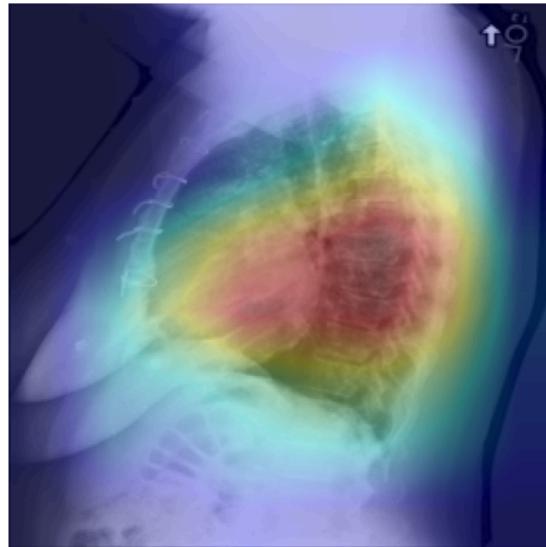


img = 2886d54b-ffa8bbe0-506cdef6-b8dd2ae4-c533ed1c

Original Image



Grad-CAM Heatmap for Cardiomegaly



Finding From Report: Severe cardiomegaly is unchanged

Cardiomegaly Testing Image 4

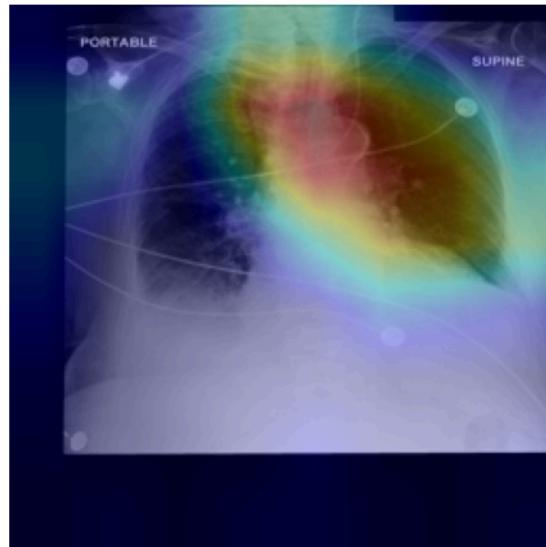
p18_p19/p19/p19817306/s53766057.txt

img = "e724f8b3-d68a8347-10f8d635-445486dd-5e1f1517"

Original Image



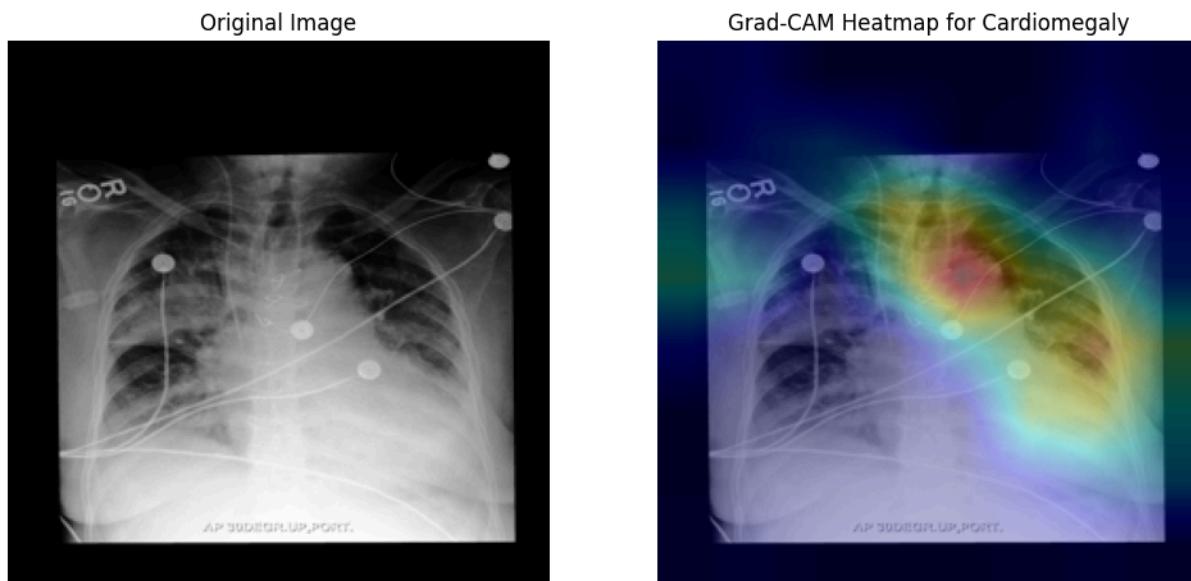
Grad-CAM Heatmap for Cardiomegaly



Finding From Report: Moderate cardiomegaly and aortic knob calcifications.

Cardiomegaly Testing Image 5
p18_p19/p19/p19840732/s58791803.txt

img = "a424fc43-96f3e8ef-34323d26-38a8ad6b-13998689"



Finding From Report: Mild cardiomegaly is not appreciably changed.

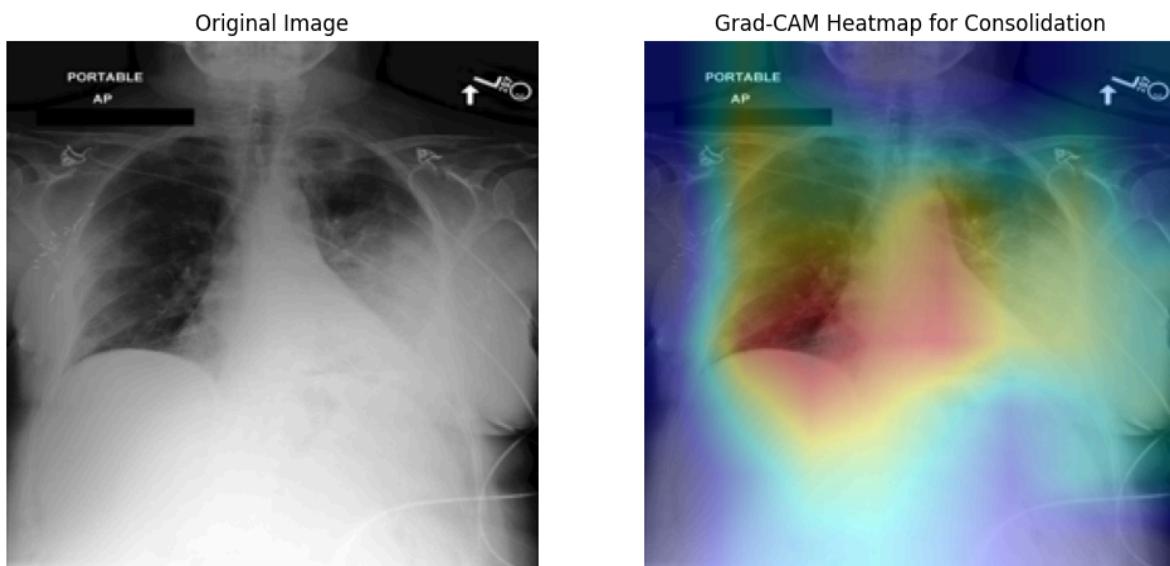
Analysis of Cardiomegaly Highlighting

Cardiomegaly is a disease of the heart so it has no location specified in doctors notes. In general the highlight is too high with some images being completely outside the heart.

Testing Results for Consolidation

Consolidation Testing Image 1
p18_p19/p18/p18722792/s54662820.txt

img = "1b854fc7-523ffa19-f61d6f35-a08a3097-a6286649"



Finding From Report:

- Consolidation seen within though left mid to lower lung.
- IMPRESSION: Confluent region of consolidation in the left lung, which could represent pneumonia in the proper clinical setting.

Consolidation Testing Image 2

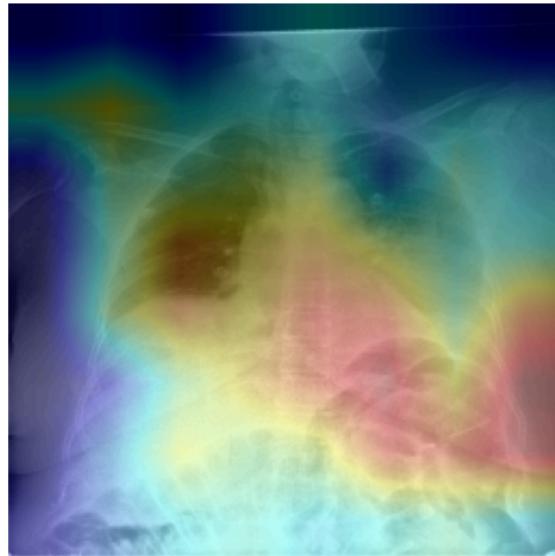
p18_p19/p18/p18716038/s52770632.txt

img = "d5f42cd8-b55716cb-6df1e516-7e45a869-8d6cca95"

Original Image



Grad-CAM Heatmap for Consolidation

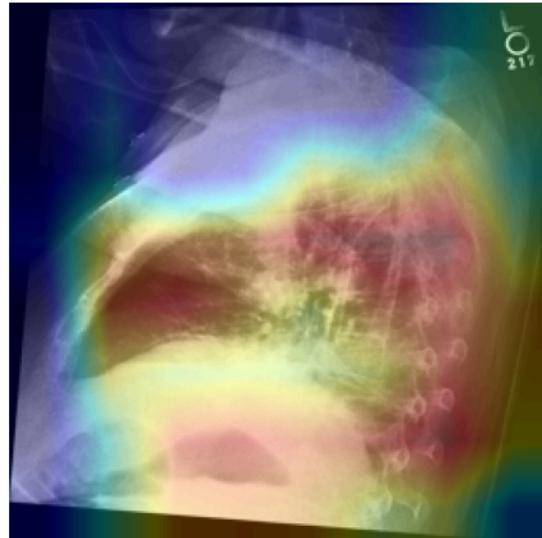


img = "f97eed61-812d-052b-6bf05c8b-072a8076-c83d9d3f"

Original Image



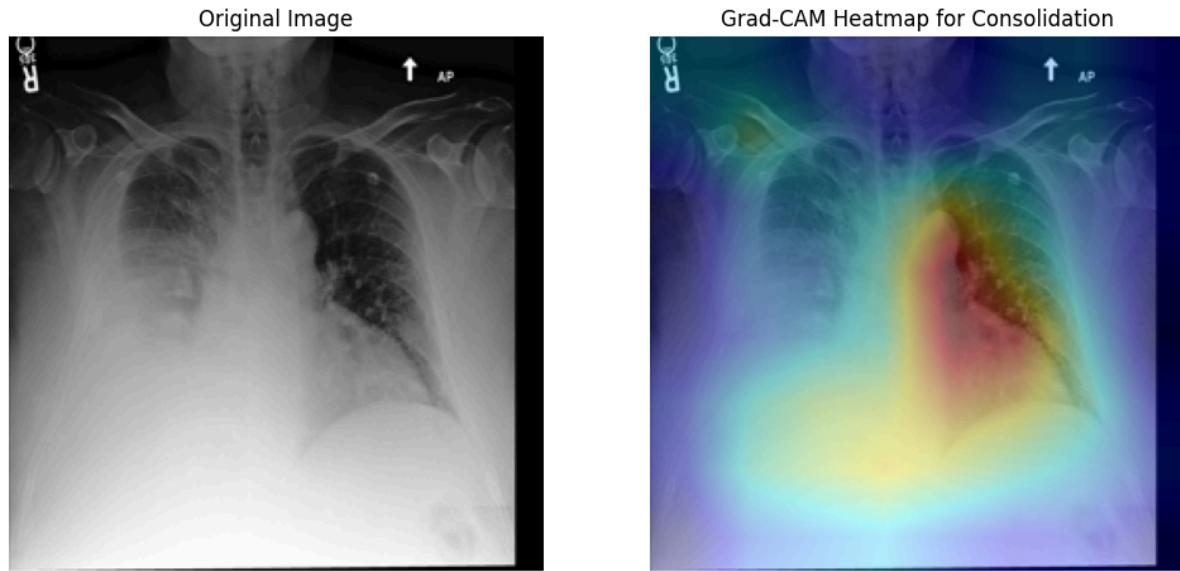
Grad-CAM Heatmap for Consolidation



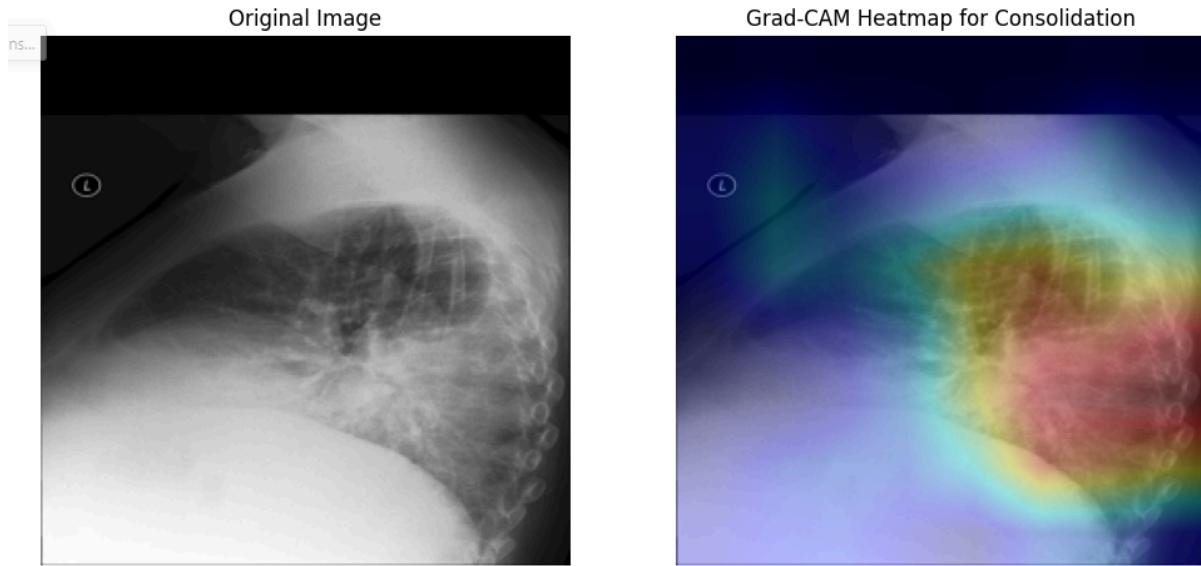
Finding From Report: There is consolidation in the left lower lobe, which could represent aspiration or pneumonia.

Consolidation Testing Image 3
p18_p19/p19/p19774163/s58753570.txt

img = "7474687d-324ff331-2e95f8d3-b7914d7c-c3d1de52"



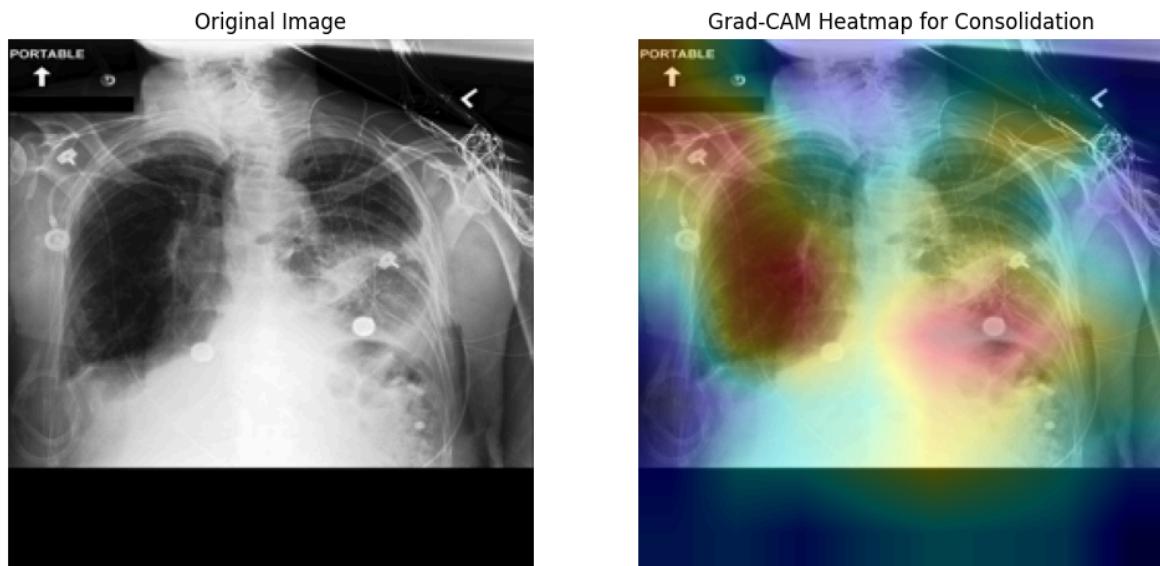
img = "a2212511-518016f2-5565443e-14f03252-12c867d1"



Finding From Report: consolidation in the right lower lung.

Consolidation Testing Image 4
p18_p19/p18/p18821140/s58495446.txt

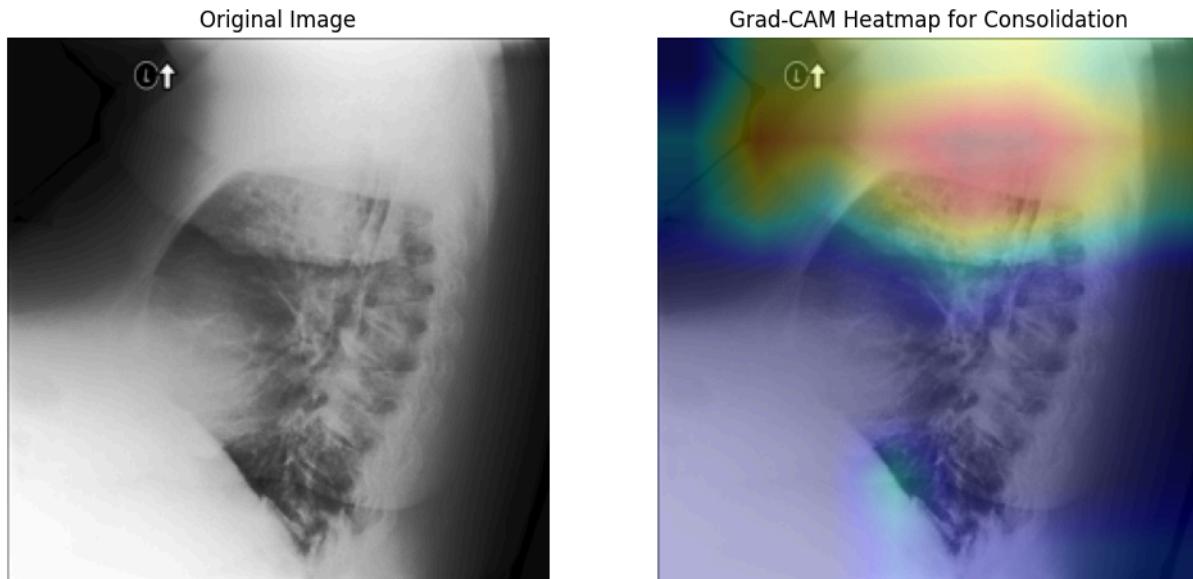
img = "ebc940d3-b17d2c78-26fe5f74-2031a732-6e1ccd22.jpg"



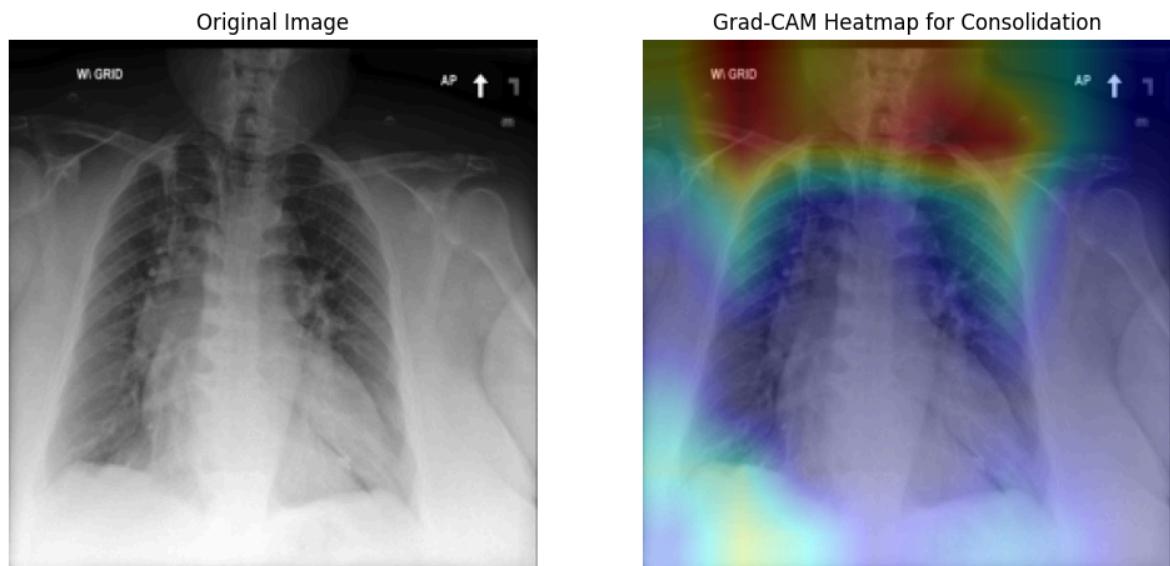
Finding From Report: PERSISTENT MODERATE TO SEVERE CONSOLIDATION, LEFT MIDLUNG, COULD BE ATELECTASIS

Consolidation Testing Image 5
p18_p19/p19/p19723160/s57120919.txt

img = "7b385b7f-04ab764c-05c36ce0-5d9ed945-f3476595.jpg"



img = "b3d4be3e-4e06b970-b7b96c8e-c431edfb-19aa5f41.jpg"



Finding From Report: There is right upper lobe consolidation which could represent a small focus of pneumonia.

Analysis of Consolidation Highlighting

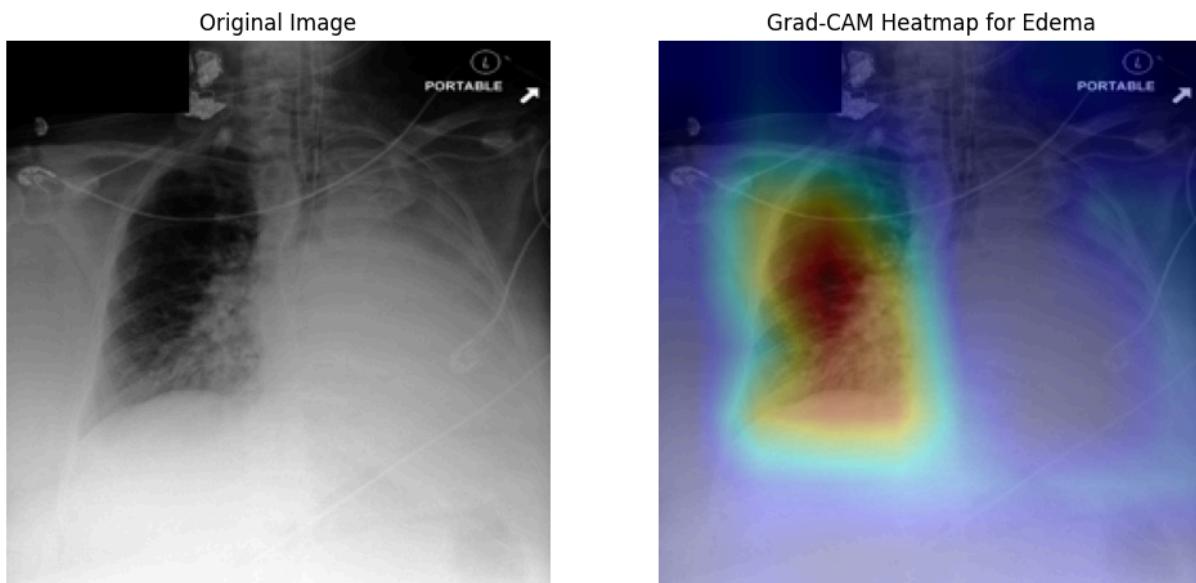
Consolidation highlighting is underperforming with the majority of highlightings incorrect.

Testing Results for Edema

Edema Testing Image 1

p18_p19/p18/p18017335/s55311626.txt

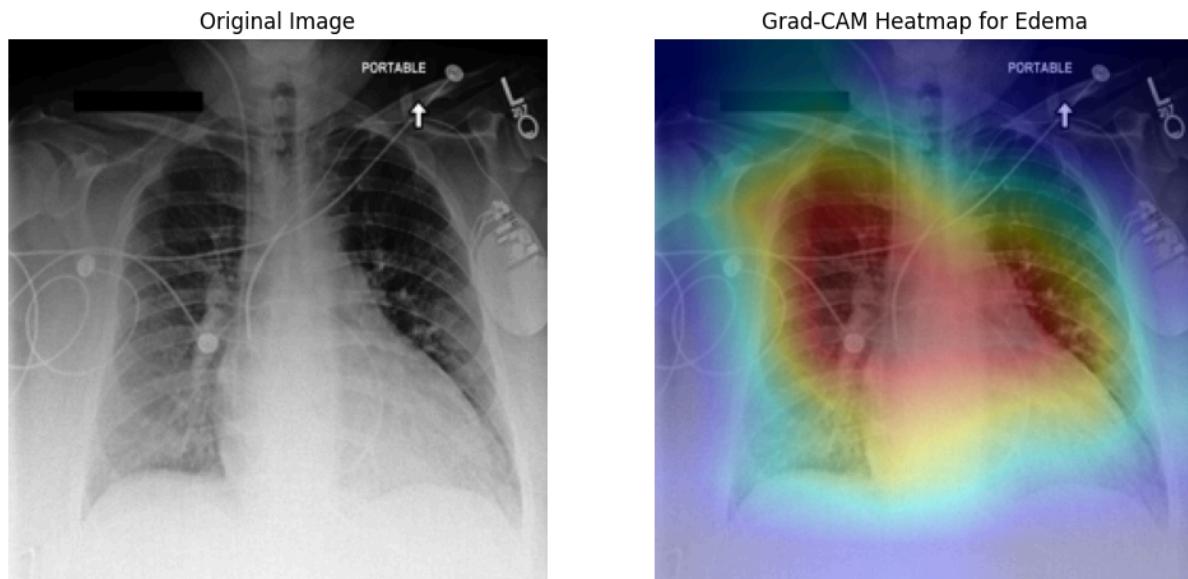
img = "dab94a45-c2c94579-3607bbaf-575b5829-26801e35"



Finding From Report: Mild edema in the right lung is improved.

Edema Testing Image 2
p18_p19/p18/p18031120/s55643091.txt

img = "0a23b5f3-77bca5ef-86f4365e-cae269b8-aa497f1a"

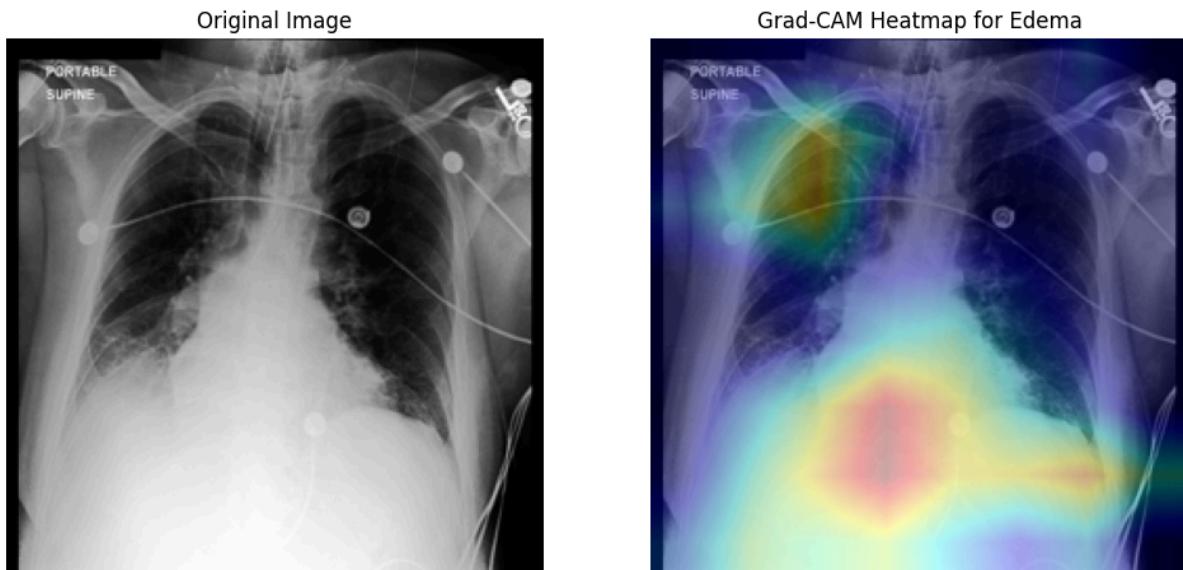


Finding From Report:

- Mild pulmonary edema with worsening in the right lower lobe.
- IMPRESSION: Mild pulmonary edema with worsening in the right lower lobe.

Edema Testing Image 3
p18_p19/p18/p18033645/s55530651.txt

img = "31faa2b6-3a8899c4-521ceae0-cefdb8f4-3b2f3680"

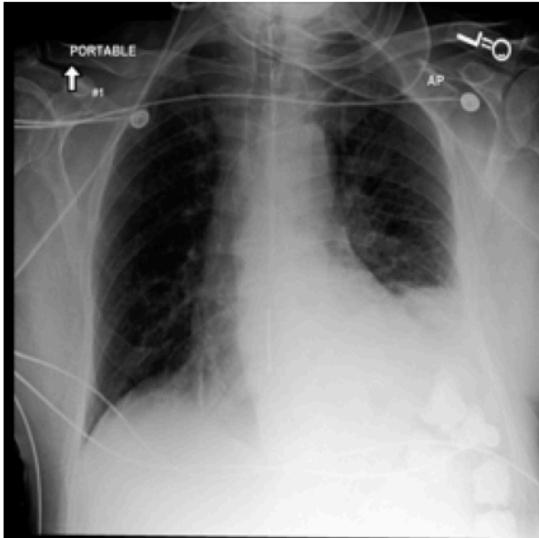


Finding From Report: Mild interstitial edema new at the base of the left lung lateral to what may be developing left lower lobe pneumonia.

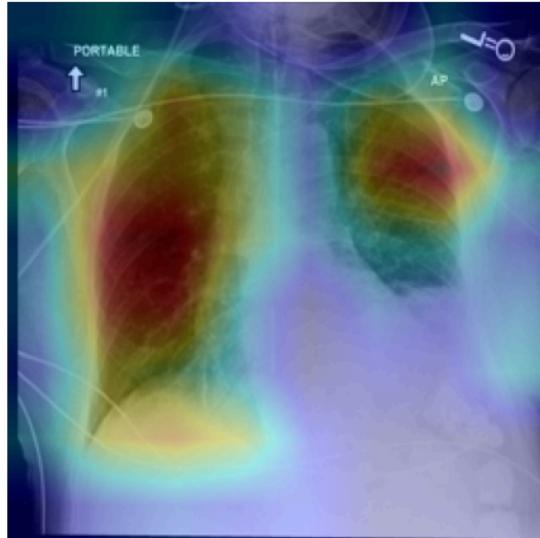
Edema Testing Image 4
p18_p19/p18/p18063505/s56916046.txt

img = "ab47c784-86aedf07-753e13a9-ce23e2c6-e997843b"

Original Image

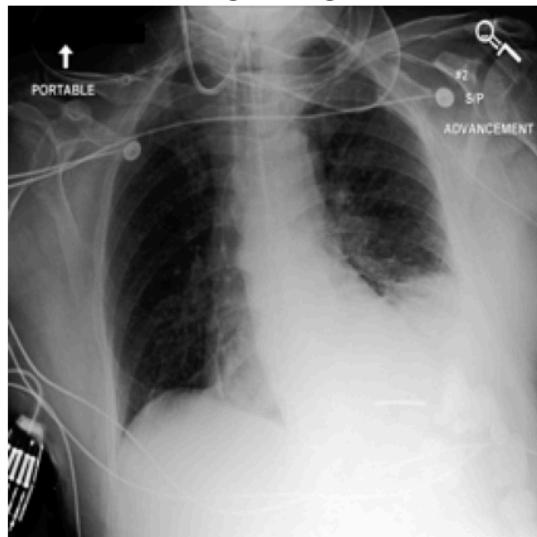


Grad-CAM Heatmap for Edema

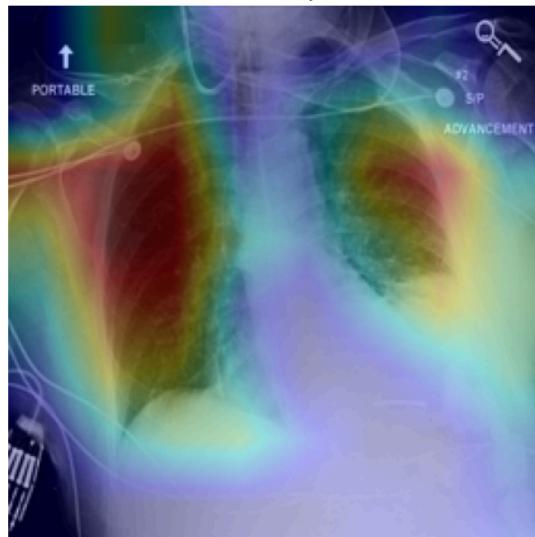


img = "28a689df-ddc46bd8-35daa760-5a5c309d-06758a2c"

Original Image



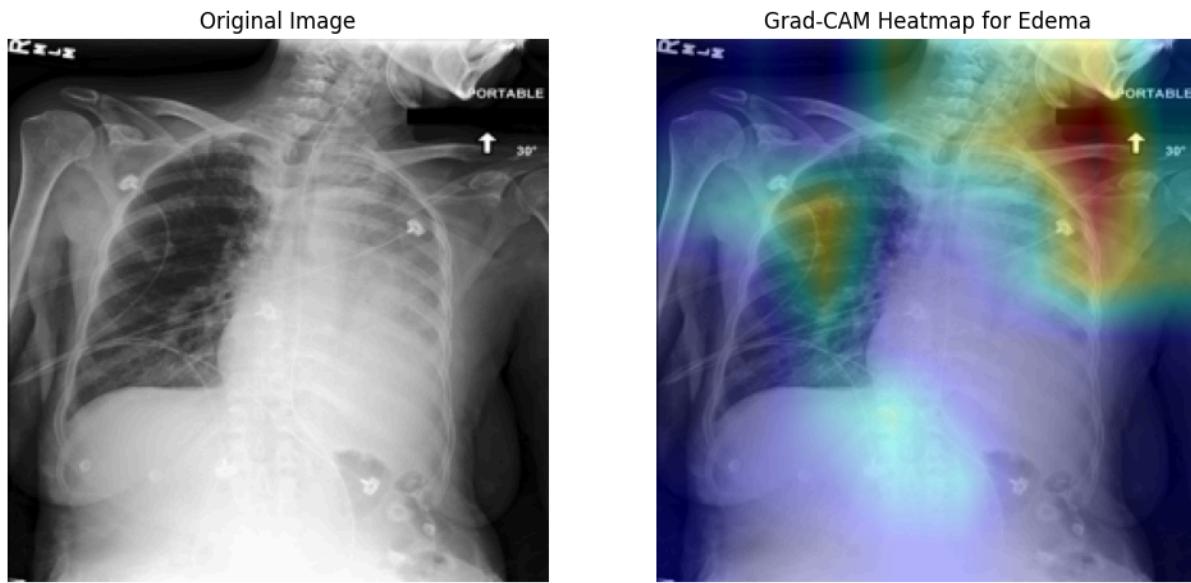
Grad-CAM Heatmap for Edema



Finding From Report: Previous edema more pronounced in the left lung, probably due to lymphatic obstruction by predominant left hilar adenopathy, has improved.

Edema Testing Image 5
p18_p19/p18/p18108905/s50619037.txt

img = "eeb6830f-5ab8e567-857fd6ed-5cec9b4a-177c157f"



Finding From Report: Minimal edema present in the right lung.

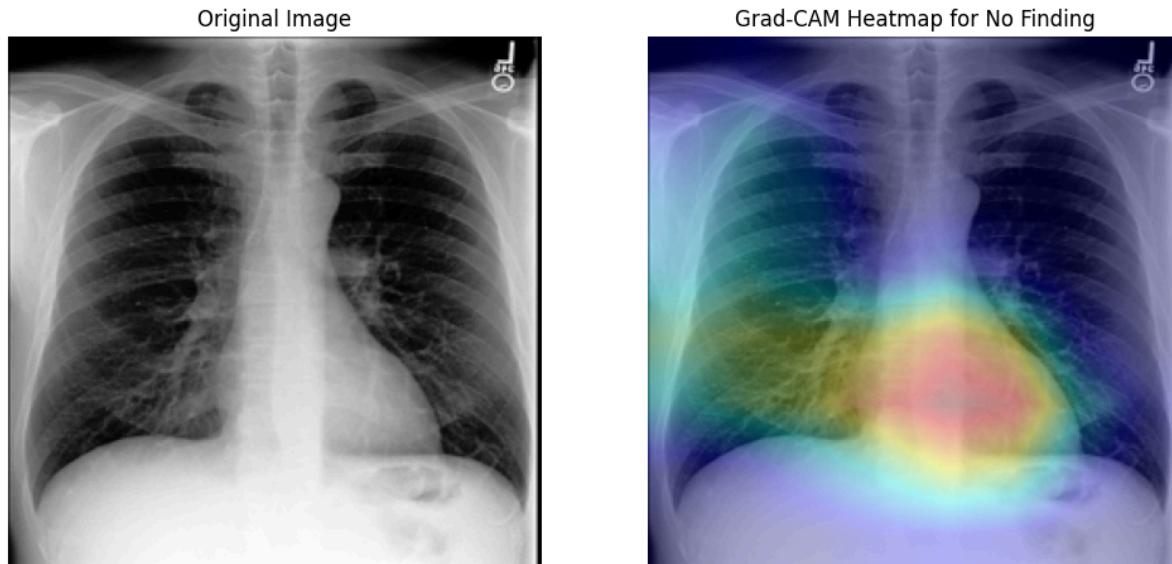
Analysis of Edema Highlighting

Overall the highlighting produced for Edema is generally in the correct area with image 5 being the only notable exception.

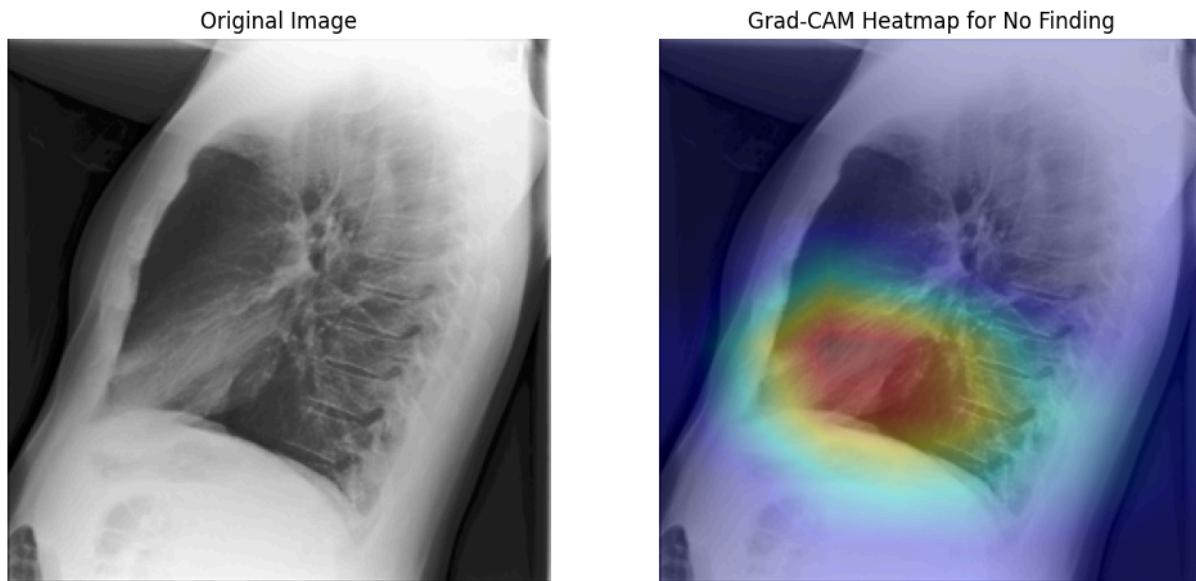
Testing Results for No Finding

No Finding Testing Image 1
p18_p19/p18/p18000291/s55388853.txt

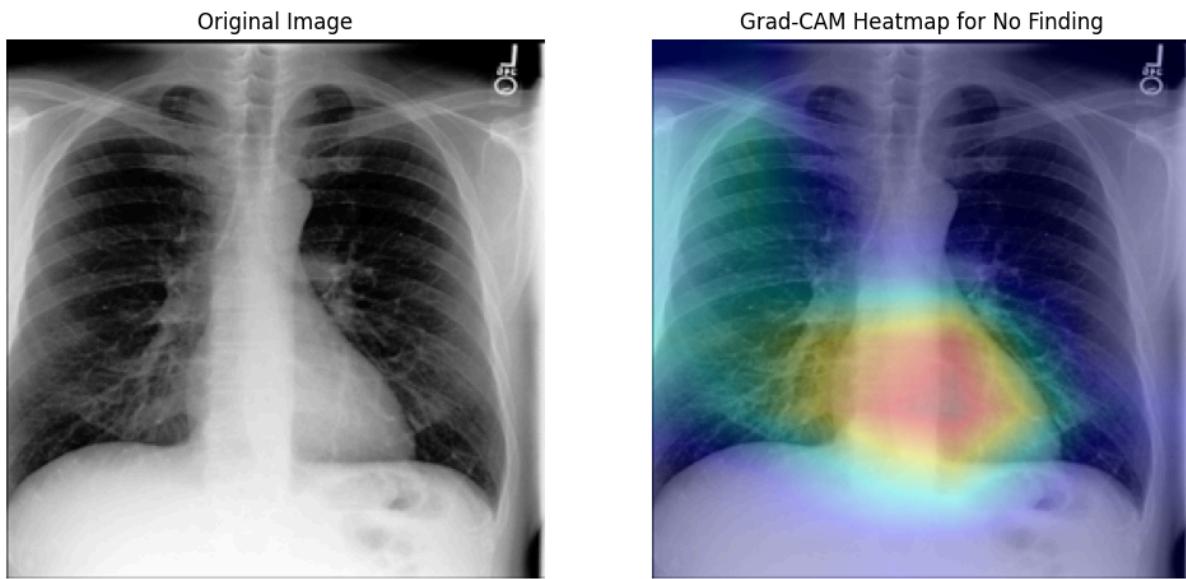
img = "04564240-d4e9e69c-1dd70a83-14b463cd-b7614743"



img = "a17c02b3-7c360f2c-a07734a0-76867697-36d59d3d"



img = "ed82a88d-499bed8a-de7539b2-417328ad-8aac20ee"

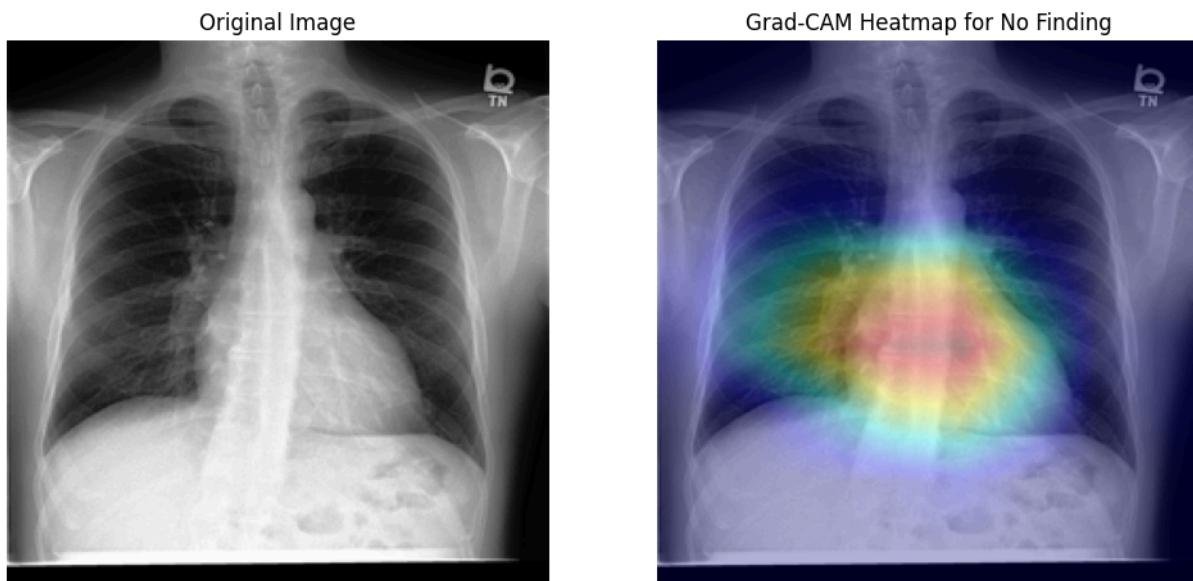


Finding From Report:

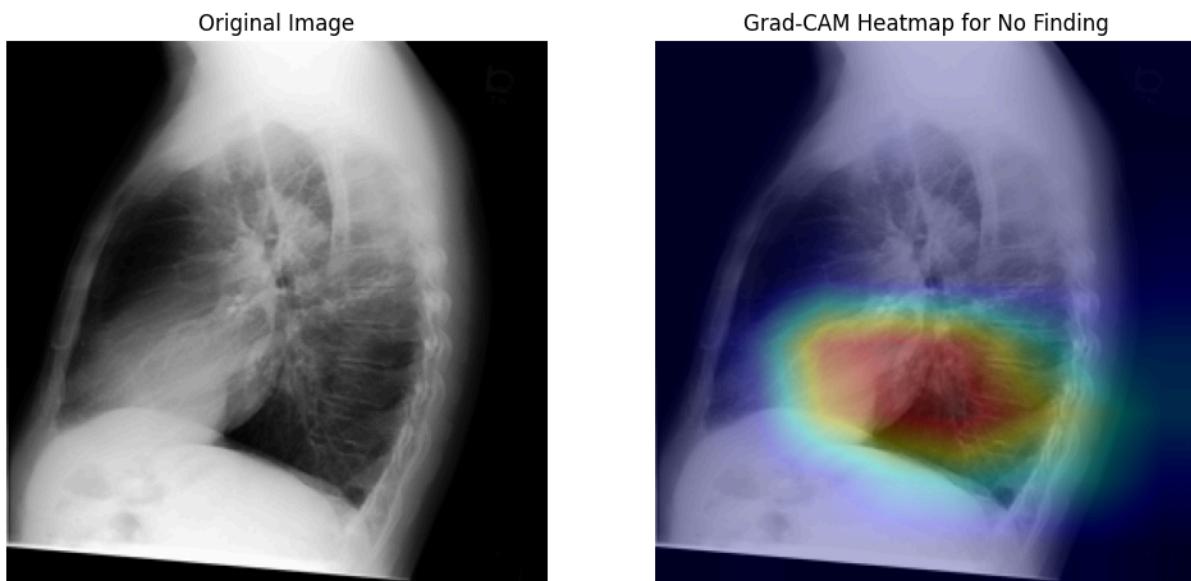
- The lungs are clear of focal consolidation, pleural effusion or pneumothorax. The heart size is normal. The mediastinal contours are normal.
- IMPRESSION: No acute cardiopulmonary process.

No Finding Testing Image 2
p18_p19/p18/p18000379/s55947854.txt

img = "6a9de243-1caf11ca-7f966bc6-0201c9e8-0dea45a3"



img = "964333fa-fd1c609d-9161adb3-03ef590f-45a8a95c"

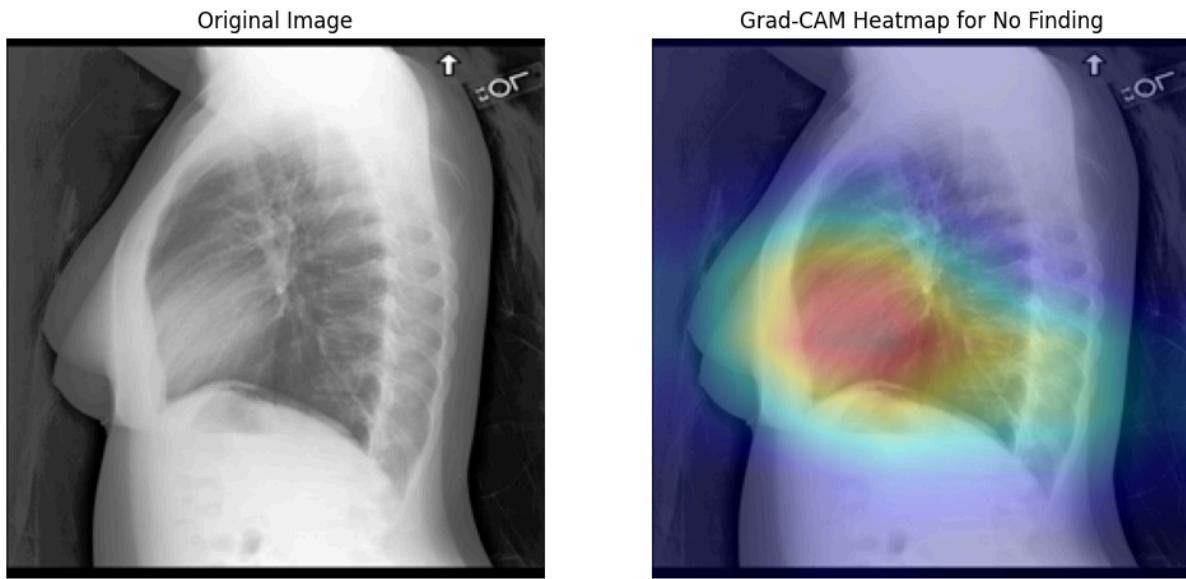


Finding From Report:

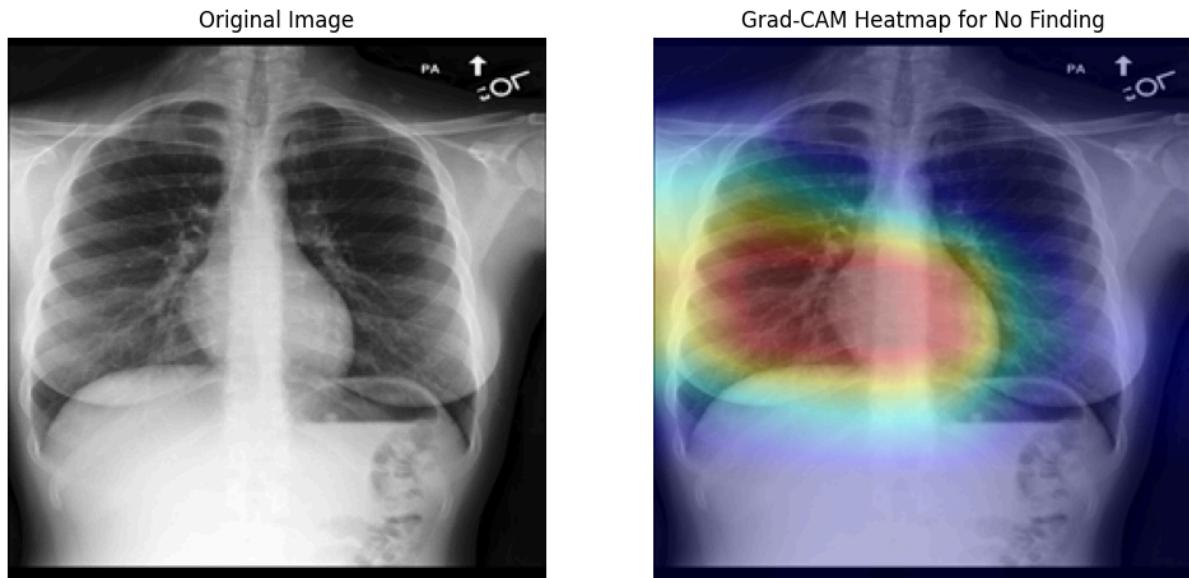
- The heart size is normal. The hilar and mediastinal contours are normal. The lungs are well expanded and clear. There is no pleural effusion or pneumothorax. The visualized osseous structures are normal.
- **IMPRESSION:** No acute abnormalities identified to explain the patient's history of fever and cough.

No Finding Testing Image 3
p18_p19/p18/p18000735/s50785186.txt

img = "e37d4ad7-dd9760f8-a2435064-618c3875-b85fad9c"



img = "f0d4b86f-aface71f-579cb776-b40850e6-948c4b8f"

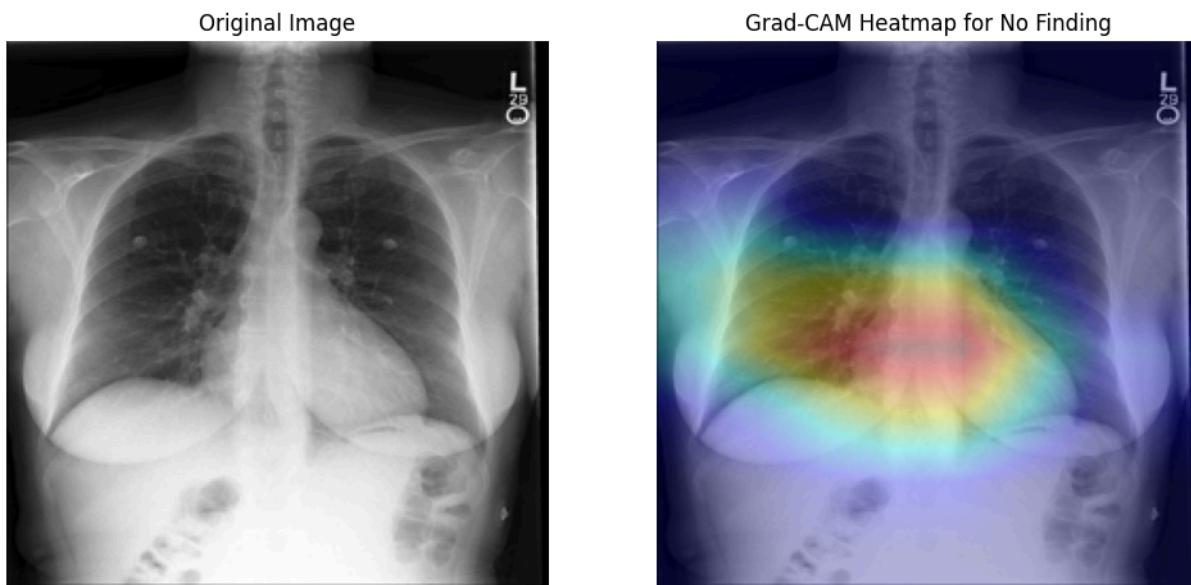


Finding From Report:

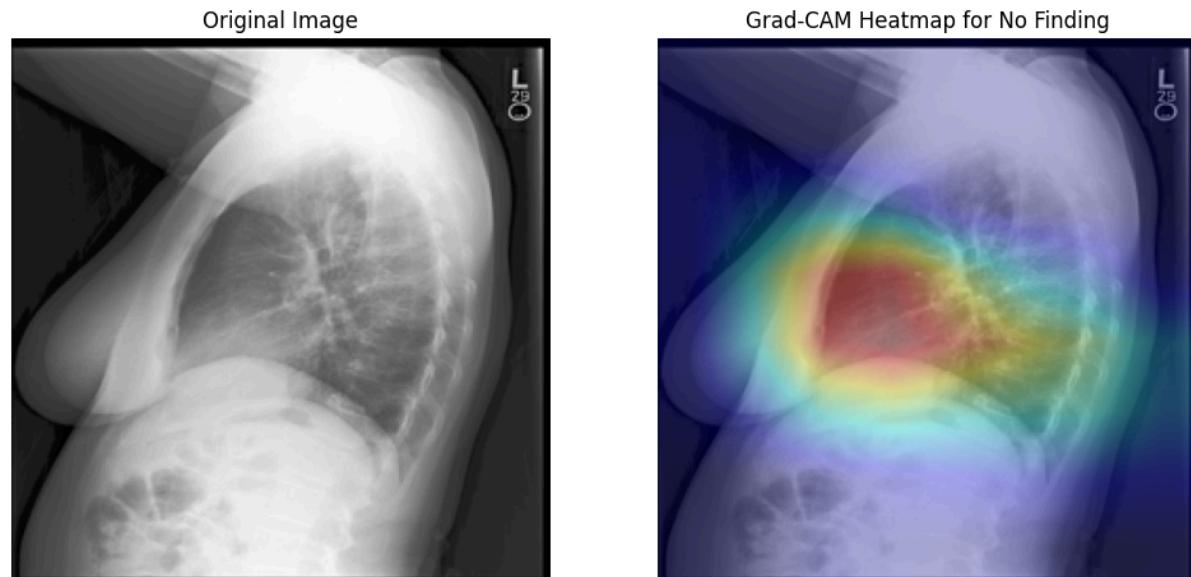
- The cardiomedastinal silhouette and pulmonary vasculature are normal. There is no pleural effusion or pneumothorax. No consolidation is identified.
- IMPRESSION: Unremarkable examination of the chest.

No Finding Testing Image 4
p18_p19/p18/p18000818/s59355587.txt

img = "1cf26fdd-54faeea9-6e331d75-f17c4cb9-11052ec0"



img = "4ed3eece-cf05be86-c4f43e0d-6fe4a79b-acad7a8c"



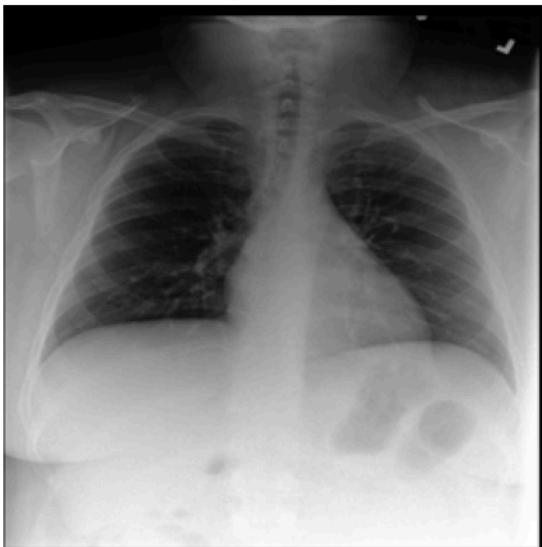
Finding From Report:

- No focal consolidation, pleural effusion, or evidence of pneumothorax is seen. The cardiac and mediastinal silhouettes are unremarkable.
- IMPRESSION: No acute cardiopulmonary process.

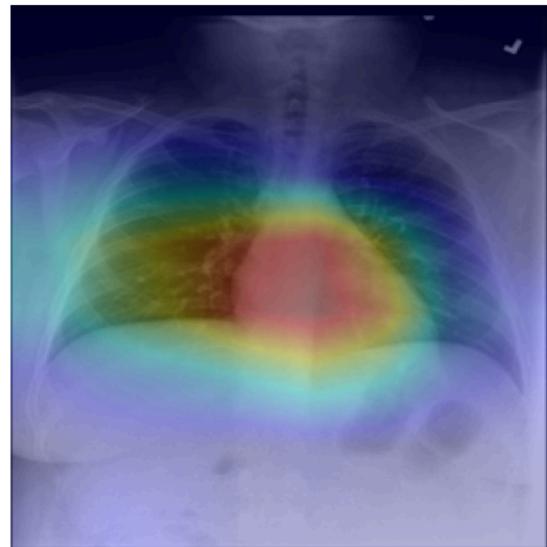
No Finding Testing Image 5
p18_p19/p18/p18001129/s51043598.txt

img = "11551ba5-8db1caaa-01677a1a-4b8297ef-10062dc7"

Original Image

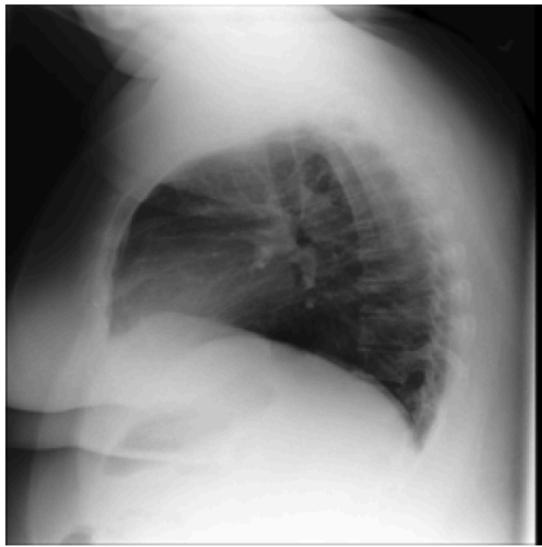


Grad-CAM Heatmap for No Finding

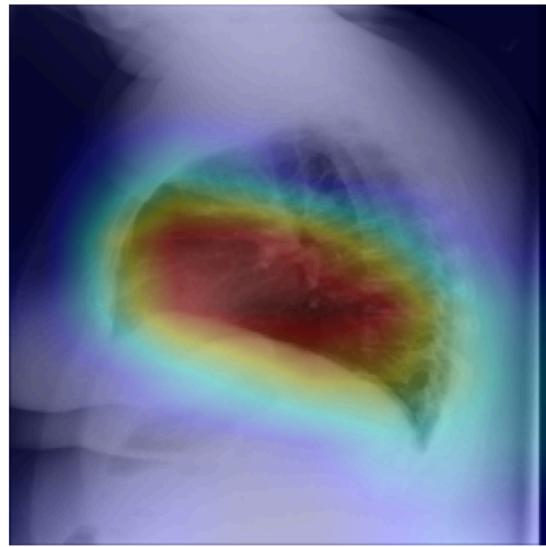


img = "973cc278-9b901ebc-9e0223d1-bee44581-e8b6452f"

Original Image



Grad-CAM Heatmap for No Finding



Finding From Report:

- The heart is normal in size. The mediastinal and hilar contours appear within normal limits. There is no pleural effusion or pneumothorax. The lungs appear clear. There has been no significant change.
- **IMPRESSION:** No evidence of acute cardiopulmonary disease.

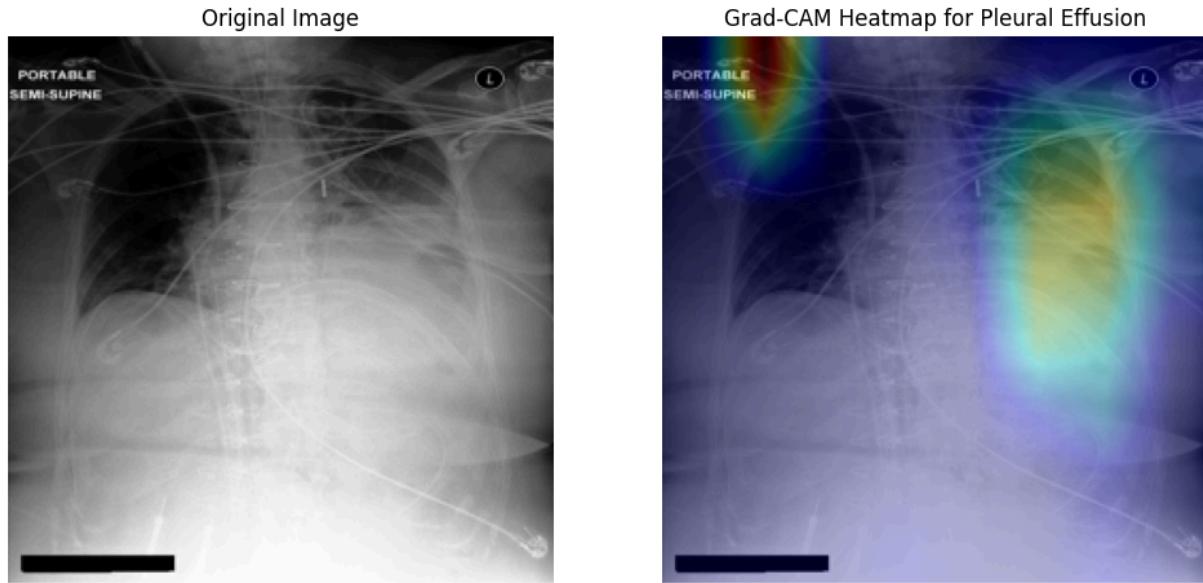
Analysis of No Finding Highlighting

The model is producing highlighting for no finding. This is expected since the highlighting represents attention mapping and even though there is no disease the model still generally focuses in the chest to determine if any diseases are present.

Testing Results for Pleural Effusion

Pleural Effusion Testing Image 1
p18_p19/p18/p18003894/s56902065.txt

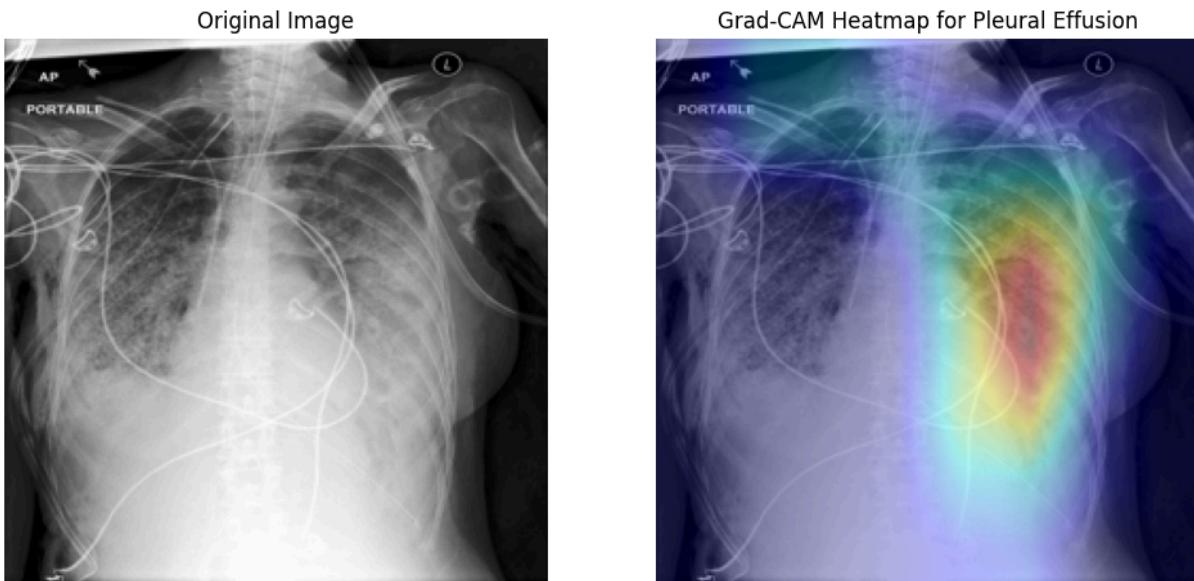
img = "98b2b140-a031b7d8-a5da775a-d9d0fb3d-79dfdc83"



Finding From Report: The left-sided pleural effusion and consolidation of the mid to lower lung fields has increased.

Pleural Effusion Testing Image 2
p18_p19/p18/p18019452/s57426164.txt

img = "363ec58a-5b27a2a2-d43ad3f9-f9f36538-f898b625"



Finding From Report:

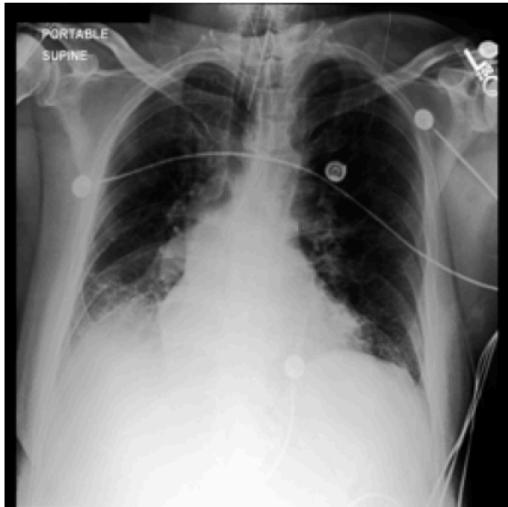
- Small bilateral pleural effusions, left greater than right, are thought likely.
- Increase in moderate left pleural effusion.
- No appreciable right pleural effusion.

Pleural Effusion Testing Image 3

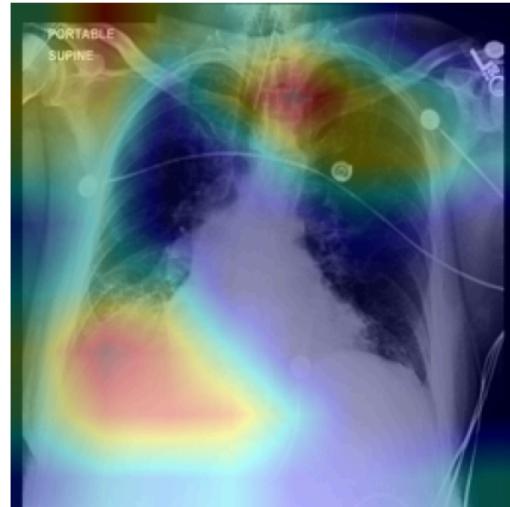
p18_p19/p18/p18033645/s55530651.txt

img = "31faa2b6-3a8899c4-521ceae0-cefdb8f4-3b2f3680"

Original Image



Grad-CAM Heatmap for Pleural Effusion



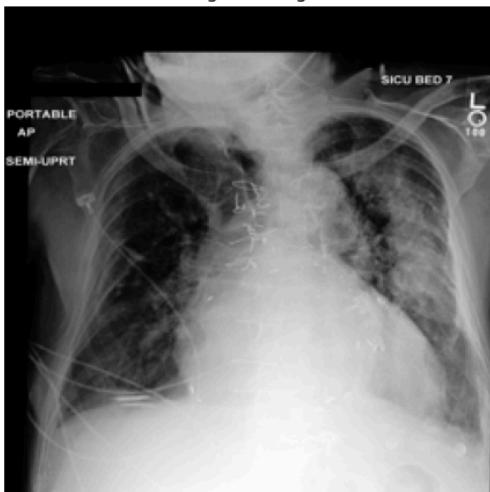
Finding From Report: Combination of right lower lobe collapse and right pleural effusion has not changed appreciably.

Pleural Effusion Testing Image 4

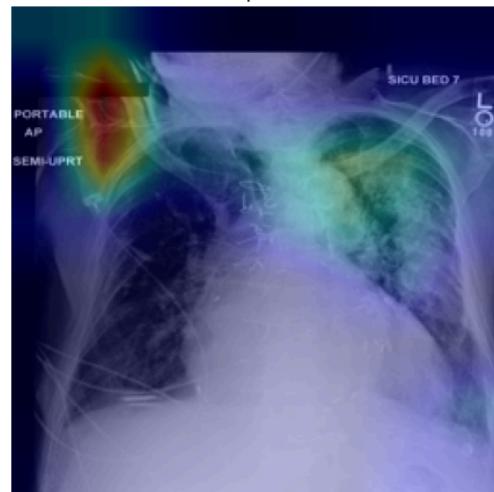
p18_p19/p18/p18042178/s52595621.txt

img = "9df9e68a-f1169710-43c90f3c-402cc9f6-39d6b824"

Original Image



Grad-CAM Heatmap for Pleural Effusion



Finding From Report:

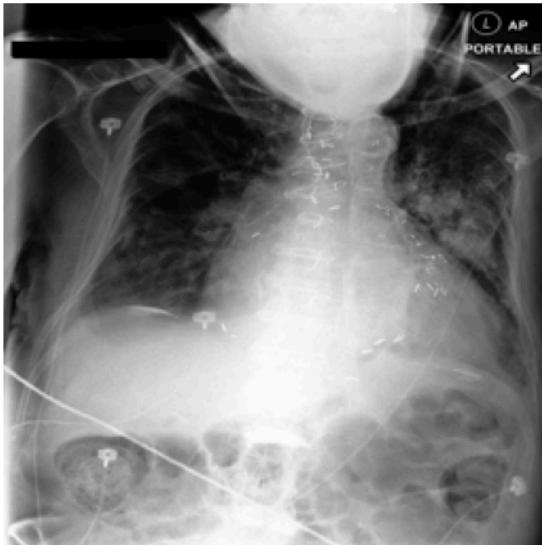
- In the left upper and lower lobes, there is an opacity with an unusual configuration, corresponding with the pleural effusion seen on the CT from the same date, and suggesting possible loculation.
- There may be right-sided pleural effusion as well, however correlation with a CT chest is done for confirmation.
- **IMPRESSION:** Left upper and lower lobe opacity, corresponding with the pleural effusion seen on the CT from the same date.

Pleural Effusion Testing Image 5

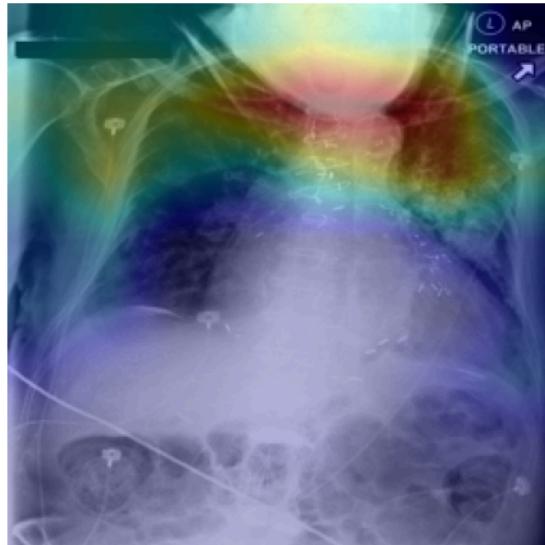
p18_p19/p18/p18042178/s55414081.txt

img = "a73f67cb-036cb1ae-34b24fc2-725883fd-b5961466"

Original Image



Grad-CAM Heatmap for Pleural Effusion



Finding From Report:

- The region of opacification of the left upper and lower lobes, corresponding to pleural effusion and atelectasis seen on CT from ___, has decreased in size over the interval.
- There are moderate-sized bilateral pleural effusions with adjacent atelectasis.
- **IMPRESSION:** Interval improvement in the degree of opacification of the left upper and lower lobes, corresponding to pleural effusion and atelectasis from CT dated ____.

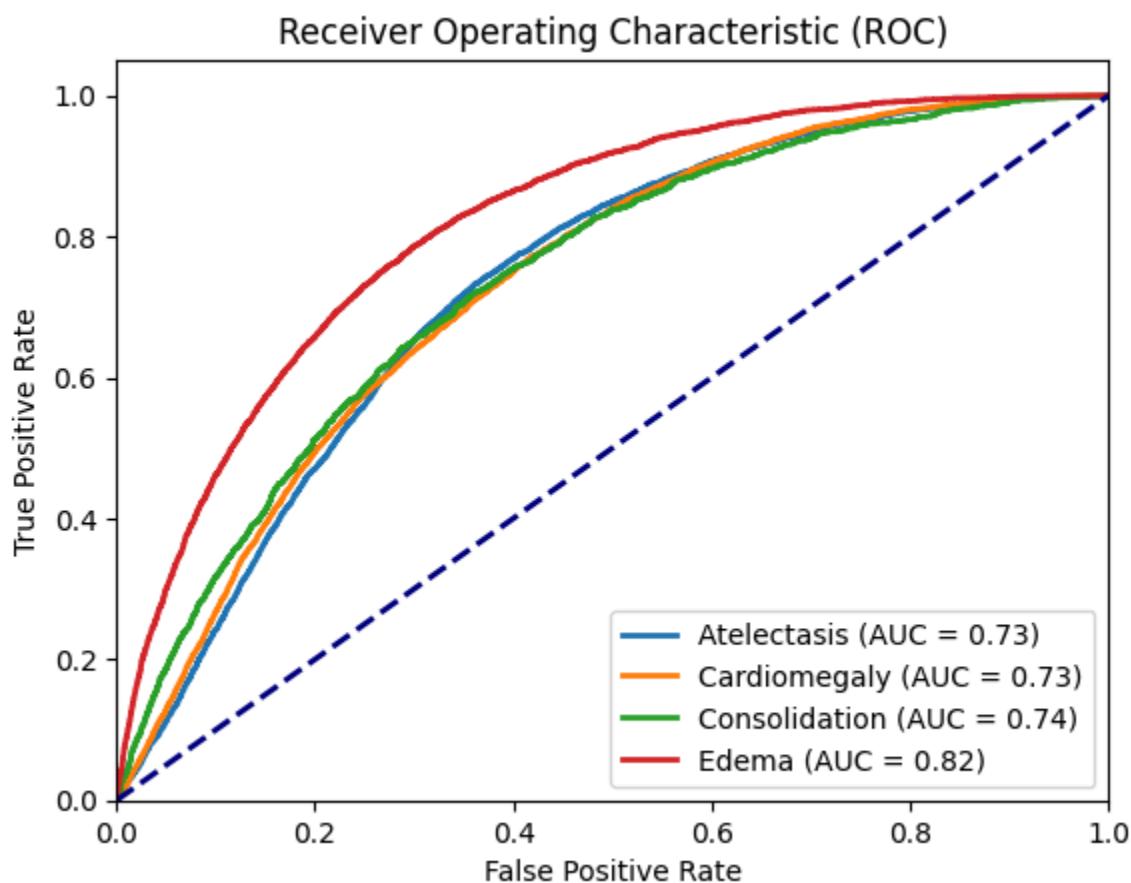
Analysis of Pleural Effusion Highlighting

For the first 3 images the highlighting covers the correct location but it also has some incorrect highlighting in other parts of the x-ray. The highlighting produced for images 4 and 5 is incorrect.

Single Image Testing Analysis

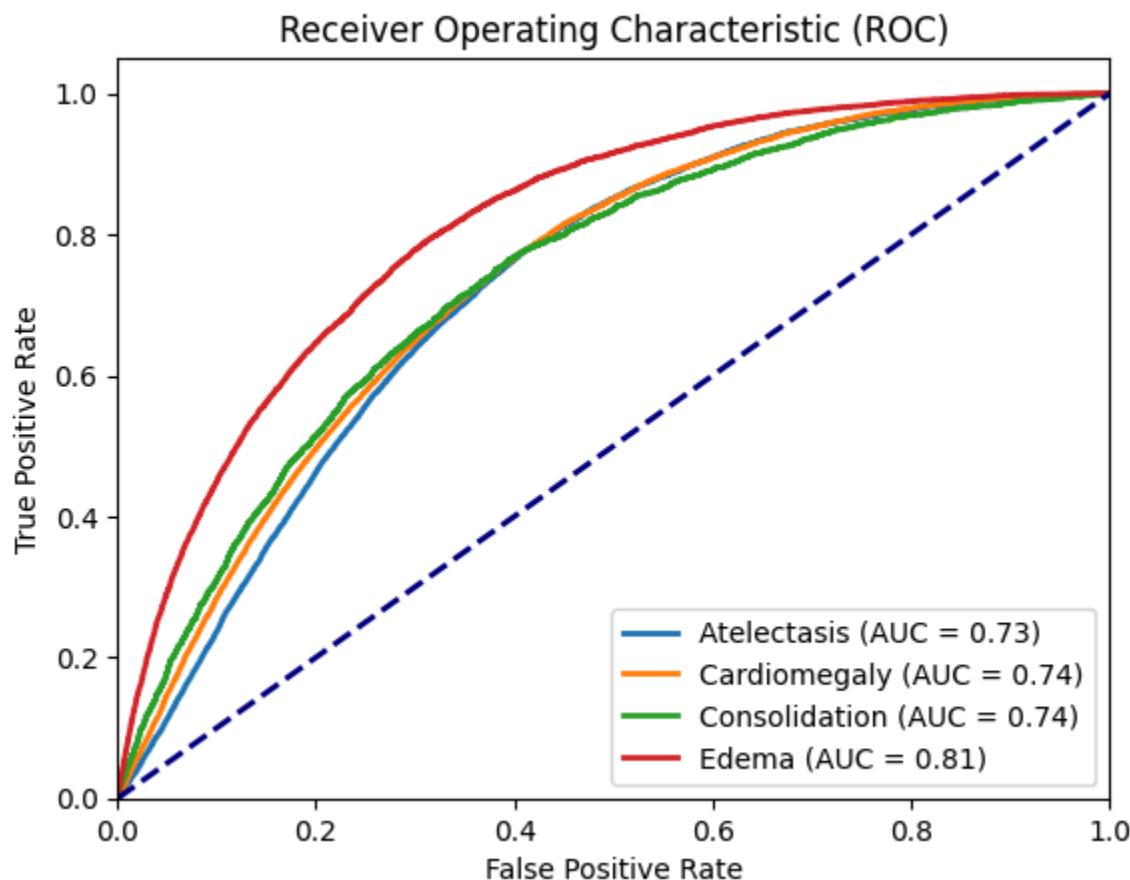
Overall, the highlighting produced by the model is generally covering correct regions but we do find that it does highlight incorrect parts of the x-rays too. Based on this we do not believe that the highlighting produced is accurate enough to assist radiologists in locating diseases.

Pre-Trained Model Validation Results



	Atelectasis	Cardiomegaly	Consolidation	Edema	No Finding	Pleural Effusion
AUC	0.7325	0.7339	0.7394	0.8185	N/A	N/A
Best Threshold	0.7022	0.6016	0.5189	0.5602	N/A	N/A
Precision	0.2958	0.2731	0.0744	0.2296	N/A	N/A
Recall	0.7326	0.7723	0.7150	0.7829	N/A	N/A
F1 Score	0.4214	0.4035	0.1348	0.3550	N/A	N/A

Pre-Trained Model Testing Results



	Atelectasis	Cardiomegaly	Consolidation	Edema	No Finding	Pleural Effusion
AUC	0.7300	0.7405	0.7410	0.8134	N/A	N/A
Best Threshold	0.6740	0.6079	0.5150	0.5579	N/A	N/A
Precision	0.2869	0.2818	0.0707	0.2120	N/A	N/A
Recall	0.7729	0.7748	0.7639	0.7898	N/A	N/A
F1 Score	0.4185	0.4133	0.1294	0.3342	N/A	N/A

The main focus of our capstone project was to train our own AI model to identify diseases in chest x-rays. Initially a pre-trained model was used to show our proof of concept and we transitioned into training our own model. In the end we decided to include a pre-trained model in our final code to see how well our model compared to a pre-trained model not tuned or further trained for our dataset. As expected our model out performed the pre-trained version because it was specifically trained and tuned for our dataset. Furthermore a value of "N/A" can be seen for the metrics of No Finding and Pleural Effusion since the pre-trained model did contain an output for these diseases.