

# Hierarchical Physiological & Radiological Model Analysis for Multimodal Respiratory Diagnostics

**Date:** January 29, 2026

**Project:** MedGemma 2026 Challenge (Healthcare Demo Pipeline)

**Target Models:** HeAR (Audio/CSI), MedSigLIP (Image/CXR)

## 1. Introduction & Objective

This document defines the **Hierarchical Physiological Model** designed to validate and explain the predictions of deep learning models (HeAR and MedSigLIP). The primary goal is to move beyond "black-box" classification by providing **interpretable, biological, and radiological evidence** rooted in clinical reasoning.

Instead of flat feature lists, we adopt a **Hierarchical Decision Tree** approach:

- Level 1 (Broad Pathophysiology):** Categorizes the general state (e.g., Infectious vs. Structural).
- Level 2 (differentiation Patterns):** Analyzes specific textures, distributions, or acoustic qualities.
- Level 3 (Disease-Specific Biomarkers):** Identifies unique signs (e.g., "Peripheral GGO" for COVID-19, "Fine Crackles" for Edema).

## 2. Part I: Audio & Scalogram Physiological Model (HeAR Target)

**Data Source:** CSI (Scalograms representing Cough/Breath sounds)

**Core Logic:** Analyzing spectral energy distribution, temporal patterns, and signal complexity.

### 2.1 The Audio Hierarchy

#### Cluster A: Infectious / Inflammatory (The "Noisy" Lungs)

*characterized by adventitious sounds (crackles, wheezes) due to secretions or inflammation.*

Disease	Level 2: Pattern (Mechanism)	Level 3: Specific Biomarker (Audio Signature)
COVID-19	Dry / Interstitial Pattern  Viral inflammation leads to non-productive coughs.	High-Harmonic "Dry" Signature <ul style="list-style-type: none"><li>Short explosive phase duration.</li><li>Energy concentrated in higher frequencies (&gt;1kHz).</li></ul>

		<ul style="list-style-type: none"> <li>• Absence of "wet" resonance (low spectral variance in tail).</li> </ul>
<b>Pneumonia</b>	<b>Wet / Alveolar Pattern</b>  Exudate accumulation leads to productive coughs.	<b>Low-Frequency "Wet" Resonance</b> <ul style="list-style-type: none"> <li>• Long explosive phase (mucus movement).</li> <li>• Strong energy &lt; 500Hz (fluid vibration).</li> <li>• <b>Coarse Crackles:</b> Discontinuous, loud popping sounds in early inspiration.</li> </ul>
<b>Tuberculosis</b>	<b>Chronic / Cavitory Pattern</b>  Long-term tissue damage and fibrosis.	<b>Paroxysmal &amp; Stridor Features</b> <ul style="list-style-type: none"> <li>• Repetitive bouts of coughing (high attack rate).</li> <li>• <b>Mixed Stridor/Wheeze:</b> Due to bronchial stenosis or obstruction.</li> <li>• <i>Caveat:</i> Can mimic Pneumonia but often shows chronic temporal irregularity.</li> </ul>

**Cluster B: Structural / Pleural (The "Silent" or "Diminished" Lungs)**

*Characterized by loss of aeration or lung collapse.*

Disease	Level 2: Pattern (Mechanism)	Level 3: Specific Biomarker (Audio Signature)
<b>Pneumothorax</b>	<b>Pleural Air Blockage</b>  Air in pleural space dampens sound transmission.	<b>"Silent Chest" / High Frequency Cutoff</b> <ul style="list-style-type: none"> <li>• <b>Significant RMS Energy Drop</b> compared to baseline.</li> <li>• Complete absence of breath sounds in affected channel (if stereo).</li> <li>• No crackles (pure silence vs. noise).</li> </ul>

<b>Atelectasis</b>	<b>Alveolar Collapse</b>  Collapsed lung tissue reduces airflow.	<b>Diminished + Re-expansion Sounds</b> <ul style="list-style-type: none"><li>• Reduced amplitude (similar to Pneumothorax but less severe).</li><li>• <b>Late Inspiratory Crackles:</b> Heard when collapsed alveoli pop open (unlike coarse pneumonia crackles).</li></ul>
--------------------	--	--

**Cluster C: Mass & Fluid (The "Resonant" Changes)**

*Characterized by solid tissue or fluid altering sound conduction.*

Disease	Level 2: Pattern (Mechanism)	Level 3: Specific Biomarker (Audio Signature)
<b>Edema</b>	<b>Fluid Overload</b>  Fluid in interstitium (not pus).	<b>"Velcro" Fine Crackles</b> <ul style="list-style-type: none"><li>• High-pitched, fine, continuous crackles.</li><li>• Predominant at the <b>end of inspiration</b>.</li><li>• Uniform pattern (unlike irregular pneumonia crackles).</li></ul>
<b>Lung Cancer</b>	<b>Obstruction / Mass Effect</b>  Tumor blocking airways.	<b>Localized Monophonic Wheeze</b> <ul style="list-style-type: none"><li>• Single-pitch whistling sound (due to fixed obstruction).</li><li>• Constant across breath cycles (doesn't clear with coughing).</li></ul>

**3. Part II: Radiological Physiological Model (MedSigLIP Target)**

**Data Source:** CXR (Chest X-Ray) / CT

**Core Logic:** Analyzing pixel intensity (Opacity), texture (GGO vs. Consolidation), and spatial distribution (Zonal Analysis).

**3.1 The Radiological Hierarchy**

## Cluster A: Increased Opacity (The "Whiter" Lungs)

*Differentiation based on Distribution (Central vs. Peripheral) and Texture (Hazy vs. Dense).*

Disease	Level 2: Distribution & Texture	Level 3: Specific Biomarker (Visual Sign)
COVID-19	<b>Peripheral &amp; GGO</b>  Bilateral, outer-lung dominance.	<b>Peripheral Zone Ratio &gt; 1.5</b> <ul style="list-style-type: none"><li>• High intensity in outer 1/3 of lung field.</li><li>• <b>Ground Glass Opacity (GGO):</b> Hazy whiteness (vessels visible) vs. dense consolidation.</li><li>• Bilateral involvement (Both lungs).</li></ul>
Pneumonia	<b>Focal &amp; Consolidation</b>  Lobar or segmental density.	<b>Lobar Segmentation + Air Bronchogram</b> <ul style="list-style-type: none"><li>• <b>Air Bronchogram Sign:</b> Dark tubular air streaks seen within dense white area.</li><li>• Sharp borders defined by lung fissures.</li><li>• Asymmetric (usually one side).</li></ul>
Edema	<b>Central &amp; Vascular</b>  Fluid backing up from heart.	<b>Bat-wing Appearance</b> <ul style="list-style-type: none"><li>• <b>Central/Hilar Dominance:</b> Opacity radiates from the center.</li><li>• <b>Cardiomegaly:</b> Cardiothoracic Ratio (CTR) &gt; 0.5.</li><li>• <b>Kerley B Lines:</b> Short horizontal lines at lung bases (lymphatic engorgement).</li></ul>
Tuberculosis	<b>Apical (Top) &amp; Cavitory</b>  Upper lobe preference.	<b>Apical Zone Dominance</b> <ul style="list-style-type: none"><li>• High intensity density in the <b>Apex (Top 1/3)</b>.</li><li>• <b>Cavitation:</b> "Donut-like" dark</li></ul>

		holes with white rims.  • Calcified nodules (high contrast spots).
<b>Lung Cancer</b>	<b>Nodular / Mass</b>  Distinct "blob" lesion.	<b>Spiculated Nodule</b>  • <b>Solitary Pulmonary Nodule:</b> Isolated high-intensity mass.  • <b>Irregular Margins:</b> Spiculated (spiky) borders indicating malignancy.  • No surrounding GGO (typically).

### Cluster B: Structural Changes (The "Shape Shifters")

*Differentiation based on lung volume and pleural boundaries.*

Disease	Level 2: Mechanism	Level 3: Specific Biomarker (Visual Sign)
<b>Pneumothorax</b>	<b>Pleural Air / Separation</b>	<b>Visceral Pleural Line</b>  • <b>Hyper-lucency:</b> Area appearing "too black" (no vessels).  • <b>Pleural Line:</b> Thin white line separating lung from air pocket.  • <b>Deep Sulcus Sign:</b> Abnormally deep costophrenic angle.
<b>Atelectasis</b>	<b>Volume Loss / Collapse</b>	<b>Mediastinal Shift (Pull)</b>  • Trachea/Heart shifts <b>TOWARDS</b> the opacity (unlike Pneumothorax/Fluid which push away).  • Elevated diaphragm on affected side.  • Triangular opacity (wedge shape).

## 4. Implementation Strategy: Evidence Extraction Pipeline

To implement this hierarchical model, the analysis/ module in the pipeline must perform specific feature extractions.

### 4.1 Audio Feature Extractor (Input: Scalogram Image)

Since raw audio might not be available, we approximate features from the Scalogram (Spectrogram) image:

1. **Frequency Recovery:** Map Y-axis pixels to Hz (0Hz - 8000Hz).
2. **Energy Profiling:** Calculate row-wise sum to get Power Spectral Density (PSD).
  - o *Low Band (0-500Hz):* Mucus/Fluid indicator.
  - o *High Band (>2000Hz):* Dry cough/Wheeze indicator.
3. **Temporal Variance:** Calculate column-wise variance to detect "Short bursts" (Cough) vs. "Continuous" (Wheeze).

### 4.2 Image Feature Extractor (Input: CXR)

Requires **Lung Segmentation** and **Zonal Grid Generation**.

1. **Segmentation:** Apply U-Net mask to isolate Left/Right Lungs.
2. **Zonal Grid:** Divide lung mask into 6 zones (Upper/Mid/Lower x Left/Right) + Central/Peripheral layers.
3. **Metrics Calculation:**
  - o *Mean Intensity per Zone:* To detect Apical vs. Basal patterns.
  - o *Texture Entropy:* To distinguish GGO (Low entropy) vs. Vessels (High entropy).
  - o *CTR Calculation:* Measure max cardiac width vs. max thoracic width.

### 4.3 Multi-Evidence Scoring Algorithm

For a given patient, we calculate a **Physiological Confidence Score**:

$$Score(Disease) = w_1 \cdot (L1_{match}) + w_2 \cdot (L2_{match}) + w_3 \cdot (L3_{match})$$

- **L1 (Category):** Does the global feature (e.g., Mean Opacity) match?
- **L2 (Pattern):** Does the texture/distribution match?
- **L3 (Biomarker):** Is the specific sign (e.g., Air Bronchogram) detected?

## 5. Summary of Deliverables

By integrating this model, the pipeline will output reports like:

**AI Prediction:** Pneumonia (Confidence: 92%)

**Physiological Evidence:**

- **[Audio]** Detected wet cough signature (High energy < 500Hz), consistent with alveolar fluid.
- **[CXR]** Detected lobar consolidation in Right Middle Lobe with Air-Bronchogram sign.
- **[Negative Evidence]** Absence of peripheral GGO (rules out COVID-19); Absence of Cardiomegaly (rules out Edema).