

KEPIL: KNOWLEDGE-ENHANCED PROMPT-IMAGE LEARNING FOR PROMPT-ROBUST DISEASE DETECTION - SUPPLEMENTARY MATERIALS

Anonymous authors

Paper under double-blind review

Medical Terms	merged_info
Pneumonia	The pneumonia is present in the Image. The Definition of pneumonia: An acute, acute and chronic, or chronic inflammation focally or diffusely affecting the lung parenchyma, due to infections (viruses, fungi, mycoplasma, or bacteria), treatment (e.g. radiation), or exposure (inhalation) to chemicals. Symptoms include cough, shortness of breath, fevers, chills, chest pain, headache, sweating, and weakness. The Radiographic Features of pneumonia: Airspace opacification; filling of the alveoli with infectious material and pus; initially patchy; becomes confluent as infection develops; air bronchograms; air-filled bronchi running through pus; filled alveoli; complications; pleural collection; cavitation.
Cardiomegaly	The cardiomegaly is present in the Image. The Definition of cardiomegaly: Increased size of the heart, clinically defined as an increased transverse diameter of the cardiac silhouette that is greater than or equal to 50% of the transverse diameter of the chest (increased cardiothoracic ratio) on a posterior-anterior projection of a chest radiograph or a computed tomography. The Radiographic Features of cardiomegaly: In most cases, merely 'eye-balling' a chest x-ray will be sufficient in detecting cardiomegaly (as the heart is either clearly normal in size or clearly abnormally enlarged). In equivocal cases, the cardiothoracic ratio (CTR) can be easily calculated on a PA chest x-ray. The CTR measures the width of the cardiac silhouette and the thoracic cavity; a ratio greater than 0.5 is an abnormal finding.
Pleural effusion	The pleural effusion is present in the Image. The Definition of pleural effusion: Presence of fluid in the pleural cavity resulting from excessive transudation or exudation from the pleural surfaces. It is a sign of disease and not a diagnosis in itself. The Radiographic Features of pleural effusion: Blunting of the costophrenic angle; blunting of the cardiophrenic angle fluid within the horizontal or oblique fissures; eventually, a meniscus will be seen, on frontal films seen laterally and gently sloping medially (note: if a hydropneumothorax is present, no such meniscus will be visible); with large volume effusions, mediastinal shift occurs away from the effusion (note: if coexistent collapse dominates then mediastinal shift may occur towards the effusion); Lateral films are able to identify a smaller amount of fluid as the costophrenic angles are deepest posteriorly. A subpulmonic effusion (a.k.a. intrapulmonary effusion) may be seen when there is previously established pulmonary disease, but can also be encountered in normal lungs. It can be difficult to identify on frontal radiographs. They are more common on the right, and usually unilateral. The following features are helpful: right: peak of the hemidiaphragm is shifted laterally left: increased distance between lower lobe air and gastric bubble.
Nodule	The nodule is present in the Image. The Definition of nodule: Focal rounded or ovoid opacity, not more than 3 cm in diameter. Pulmonary nodules are typically observed by chest radiography or computer tomography imaging. The Radiographic Features of nodule: They are generally homogeneous (without air bronchograms or alveolograms) and are well-defined since their margins are sharp and they are surrounded by normally aerated lung parenchyma.
Infiltrate	The infiltration is present in the Image. The Definition of infiltration: The process of the diffusion or accumulation in a tissue or cells of a substance not normal to it or in amounts above normal. The Radiographic Features of infiltration: Patchy or diffuse opacities on chest radiographs, often representing fluid, infection, or inflammatory cells.
Heart insufficiency	The heart insufficiency is present in the Image. The Definition of heart insufficiency: A condition where the heart cannot pump blood effectively to meet the body's needs. The Radiographic Features of heart insufficiency: Cardiomegaly, pulmonary edema, pleural effusions, and upper lobe venous diversion on chest radiographs.
Hyperinflated lung	The hyperinflate is present in the Image. The Definition of hyperinflate: Abnormal permanent enlargement of the lung air spaces distal to terminal bronchiole not resulted from wall destruction, e.g. due to loss of opposite lung. The Radiographic Features of hyperinflate: Observable features include: flattened hemidiaphragmatic contours considered one of the most sensitive indicators of hyperinflation and interobserver variability is small best seen on the lateral chest radiograph and consists of a loss of height of the convexity of the hemidiaphragm to measure, it is possible to draw a line connecting the sternophrenic angle and the posterior costophrenic angle this arch height should be greater than or equal to 2.5 cm it is considered clearly pathological when measures less than 1.5 cm minor measures correlate well with the functional importance of airflow obstruction retrosternal space measurement a horizontal line is drawn from a point allocated 3 cm below the junction between the manubrium and sternal body, on the posterior cortex of the sternum, and the ascending aorta point; in cases of increased retrosternal space, this distance is equal or greater than 2.5 cm air trapping: when comparing two radiographs acquired in maximal inspiration and maximal expiration, the vertical movement of the diaphragm is less than 3 cm appearance of the ribs: more than 6 anterior or 10 posterior ribs above the diaphragm level on the midclavicular line horizontalisation of ribs presence of air below the heart increased anteroposterior diameter of the chest, also called barrel chest hyperlucent lungs (i.e. less bronchovascular markings per unit area).

Table 1: PadChest dataset - Inference Prompt

1 DETAILED EXPERIMENTAL SETUP

1.1 PRETRAINING

Training utilized chest X-ray images, resized to a resolution of 224×224 pixels. The training process spanned 100 epochs, with a warmup period of 20 epochs, and employed a batch size of 48 per GPU for training and 6 for testing. Mixed precision training was enabled via a gradient scaler to enhance computational efficiency.

The optimizer was AdamW with a learning rate of 5×10^{-5} , weight decay of 0.02, and a cosine annealing scheduler that decayed the learning rate to a minimum of 1×10^{-6} . The warmup phase started with a learning rate of 1×10^{-6} . The image encoder was configurable as ViT-B, while the text encoder supported either a BERT-based CLP-clinical model or Llama3-OpenBioLLM-8B. In the LLM version setup, the text encoder was frozen, and a multilayer perceptron (MLP) projector mapped 4096-dimensional embeddings to 1024 dimensions, using three layers, a hidden dimension of 2048, and a dropout rate of 0.1. For the KEPIL configuration, an optional frozen BERT-based text encoder was paired with an MLP projector transforming 768-dimensional embeddings to 1024

dimensions, with two layers, a hidden dimension of 1536, and a dropout rate of 0.5. The λ_1 to λ_3 are set to 1 by default.

A contrastive loss with a temperature parameter of 0.07 was applied. The loss ratio between contrastive and semantic-awareness contrastive loss was set to 1. A momentum of 0.995 and an alpha of 0.4 were used for feature updates, with a queue size of 8192 for contrastive learning. The model was trained to predict 40 classes.

1.2 FINETUNING

Input images were resized to a resolution of 224×224 pixels, utilizing 100% of the available training data. Fine-tuning was conducted over 60 epochs, with a 5-epoch warmup period. The training batch size was set to 48 per GPU, while testing used a batch size of 256.

The AdamW optimizer was employed with a learning rate of 1×10^{-5} and a weight decay of 0.02. A cosine annealing scheduler maintained the learning rate at 1×10^{-5} throughout, with a warmup learning rate of 1×10^{-5} and no cooldown period.

1.3 PRE-TRAINING AND DOWNSTREAM DATASET

We utilize MIMIC-CXR (Johnson et al., 2019) as the pretraining dataset. This dataset comprises 377,110 chest X-ray images and their corresponding radiology reports from 65,379 patients. Following the preprocessing steps outlined in GLoRIA (Huang et al., 2021), we select 213,384 image-text pairs for pretraining.

1.4 DOWNSTREAM TASK DATASETS

We evaluate our approach on seven downstream task datasets:

CheXpert (Irvin et al., 2019): Contains 224,316 chest X-ray images labeled with 14 disease categories. We use the official validation set (234 images) for evaluation.

ChestXray-14 (Wang et al., 2017): Comprises 112,120 chest X-ray images annotated with 14 disease labels. The official test set (6,412 images) is used for evaluation.

PadChest (Bustos et al., 2020): Consists of 160,868 chest X-ray images and their corresponding radiology reports. We divide this dataset into three subsets: (i) **PadChest-seen**: Includes disease categories present in MIMIC-CXR; (ii) **PadChest-unseen**: Contains disease categories not present in MIMIC-CXR; (iii) **PadChest-rare**: Comprises rare disease categories within MIMIC-CXR. Each subset is evaluated using the official test set.

RSNA Pneumonia (Wu et al., 2024): Features 30,000 chest X-ray images for pneumonia detection and localization tasks. We utilize the official test set for evaluation.

SIIM-ACR (Anna Zawacki et al., 2019): Contains 12,089 chest X-ray images with corresponding pneumothorax segmentation masks. The official test set is used for evaluation.

COVIDx CXR-2 (Pavlova et al., 2022): Includes 29,986 chest X-ray images from 16,648 COVID-19 patients, accompanied by classification labels. Following the setup in (Wu et al., 2023), we split the dataset into training (70%), validation (20%), and test (10%) sets.

COVID Rural (Desai et al., 2020): Comprises over 200 chest X-ray images with corresponding COVID-19 segmentation masks. We divide this dataset into training (60%), validation (20%), and test (20%) sets.

1.5 EVALUATION METRICS

Consistent with (Wu et al., 2023; Boecking et al., 2022), we employ the following evaluation metrics:

- **Classification Tasks**: Area Under the Curve (AUC), F1 score, and accuracy.

- **Segmentation Tasks:** Dice score, Intersection over Union (IoU), and pixel-wise accuracy.

All metrics are reported as percentages.

Table 2: Ablation study of utilizing Llama3-OpenBioLLM-70B as text encoder on the ChestXRay14 dataset.

Zero-shot Inference	ChestXray14
KEPIL (LLM, without L_{sc})	0.7789
KEPIL (LLM)	0.7856

2 SELECTION OF CONTENT COMPLETION LLM

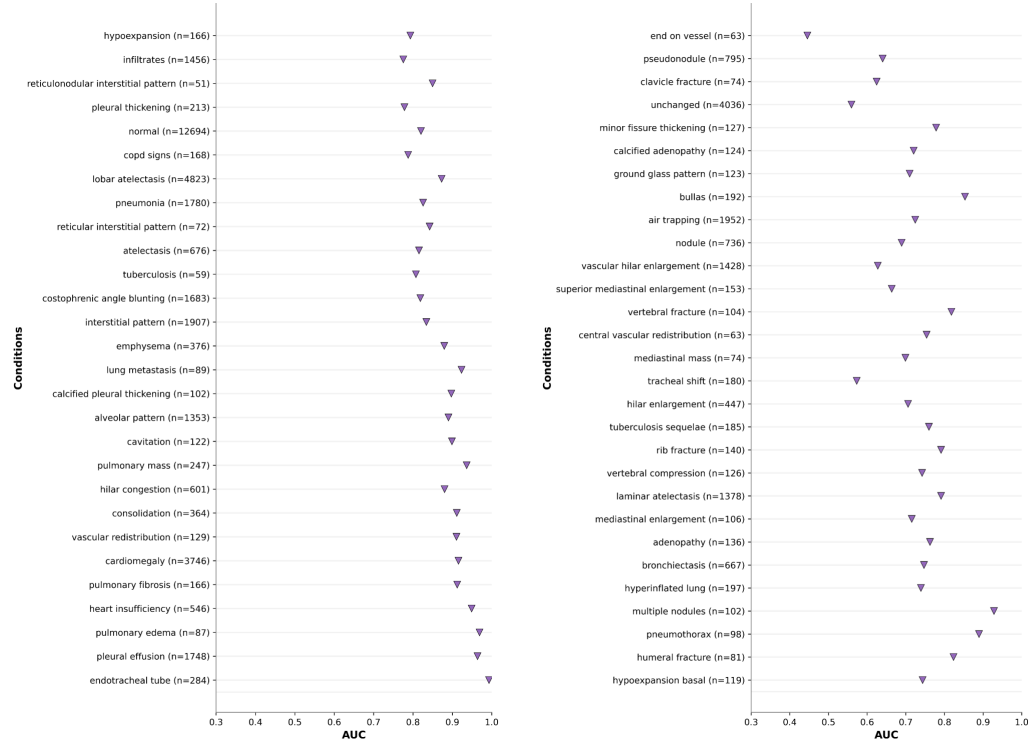


Figure 1: Model performance on the PadChest dataset. The evaluation was conducted on high-importance classes (≥ 50 samples) not used during training. KEPIL achieved $AUC \geq 0.900$ for 11 findings and ≥ 0.700 for 48 of 57. It outperformed CheXNet (Rajpurkar et al., 2017) on all 6 unseen diseases listed in (Rajpurkar et al., 2017).

In order to select the optimal LLM for content formatting and missing content completion, we tested the Open/Close LLM as shown in Table. 3.

3 INFERENCE PROMPT FOR PADCHEST DATASET

Table 1 shows eight examples of zero-shot inference for PadChest. Compared to simple Finding name or General definition, the prompts we generate are more comprehensive and accurate.

4 CLASSIFICATION UNDER REAL-WORLD SETTING.

Figure 1 illustrates the performance of our proposed method, KEPIL, on the PadChest dataset, focusing on high-impact classes with at least 50 samples. KEPIL achieves an $AUC \geq 0.900$ for 11

Table 3: A comparative study on the completion of missing visual descriptions: To select the best model, we evaluate the visual descriptions generated by Mixtral, Medorca, and PMC-llama across various metrics against the collected ground truth visual descriptions.

Metric	ChatGPT-4o	Mixtral-8×7B	Medorca-2x7B	PMC-llama-7B
BLEU-4	0.1180	0.1030	0.0018	0.0033
ROUGE-L	0.2695	0.2328	0.1461	0.1460
F1 RadGraph	0.2384	0.2046	0.0639	0.0979

conditions and ≥ 0.700 for 48 out of 57 conditions, demonstrating robust performance across a wide range of findings. Notably, KEPIL outperforms CheXNet (Finetuned) (Rajpurkar et al., 2017) on all 6 unseen diseases listed in (Rajpurkar et al., 2017), including conditions such as pleural effusion, pulmonary edema, and heart insufficiency. In addition, KEPIL achieves an average AUC of 0.7851 across 193 diseases in the full Padchest dataset.

5 PERFORMANCE OF LLM AS TEXT ENCODER

We also test the performance of replacing BERT encoder with Biomedical LLM. The performance is shown in Table 2.

6 THE PROMPT OF GENERATING KNOWLEDGE

The prompt of generating knowledge is shown as follows: *Please generate a structured medical description in the same style as the previous cases. Use the format: The Disease is present in the Image. The Definition of Disease: ... The Radiographic Features of Disease: ... The Definition should be referenced from UMLS, the Radiographic Features from Radiopaedia, and the writing style should follow the tone and structure of the previous case examples, with clear, concise, and professional medical English.*

7 THE USE OF LARGE LANGUAGE MODELS (LLMs)

We used ChatGPT-5 solely for checking typos and grammatical errors.

8 CLINICAL VARIATIONS

To better approximate how clinicians naturally vary their phrasing, we simulate four categories of prompt variations commonly observed in real clinical environments. These variants are illustrated in Fig. 2.

Medical Term	Abbrivation	Multilingual	Chinese_jargon	Medical_synonym
normal	The normal is present in the film. The Definition of nt. Indicates no acute pathology - reflects a The normal is present in the image. La definición de normal. Indica ausencia de patología aguda. se. The normal is present on this film. The discovery			The normal is present in the image. The Definition of normal. denotes an unremarkable study with no detectable disease. overall physiologic appearance. The finding
pleural effusion	The pleural effusion is present on the film. The Definition of PE. Accumulation of excess fluid. The pleural effusion is present in the image. La definición de efusión pleural. acumulación de líquido. The pleural effusion is present on this film. The			The pleural effusion is present in the image. The Definition of pleural effusion. accumulation of excess intrapleural fluid within the pleural compartment. The Radiographic
opacity	The opacity is present on the film. The Definition of an opac. any region on the radiograph. The opacity is present in the image. La definición de opacidez. zona en la radiografía que aparece más The opacity is present on this film. The discovery			The opacity is present in the image. The Definition of opacity. a region of increased radiographic attenuation that appears higher than expected, produced by fluid, or
pneumothorax	The pneumothorax is present on the film. The Definition of PTX. air within the pleural space. The pneumothorax is present in the image. La definición de pneumothorax. aire dentro del espacio. The pneumothorax is present on this film. The pneumothorax is present in the image. The Definition of pneumothorax. an intrapleural air collection resulting in partial or complete pulmonary collapse. The finding			The pneumothorax is present in the image. The Definition of pneumothorax. an intrapleural air collection resulting in partial or complete pulmonary collapse. The finding
adhesions	The adhesions is present on the film. The Definition of ETC. collapse of lung tissue leading. The adhesions is present in the image. La definición de adhesiones. colapso del tejido pulmonar con. The adhesions is present on this film. The discovery			The adhesions is present in the image. The Definition of adhesions. pleural adhesions involving the fissures or alveolar components, related to congenital or
tube	The tube is present on the film. The Definition of a tube. a medical device such as an ETT. The tube is present in the image. La definición de tubo. dispositivo médico como un tubo endotraqueal. The tube is present on this film. A tube has in The tube is present in the image. The Definition of tube. a radiopaque medical conduit such as an endotracheal or esophageal tube. The Radiographic Features of tub			The tube is present in the image. The Definition of tube. a radiopaque medical conduit such as an endotracheal or esophageal tube. The Radiographic Features of tub
consolidation	The consolidation is present on the film. The Definition of cons. Wedg. of pleural spaces with the consolidation is present in the image. La definición de consolidación. coagulación de los espacios. The consolidation is present on this film. Cons. The consolidation is present in the image. The Definition of consolidation. a parenchymal/fining process where alveolar spaces are occupied by fluid. inflammatory cell			The consolidation is present in the image. The Definition of consolidation. a parenchymal/fining process where alveolar spaces are occupied by fluid. inflammatory cell
enlarged cardiac	The enlarged cardiomegaly is present on the film. The Definition of enlarged CM. increased The enlarged cardiomegaly is present in the image. La definición de cardiomegalia aumentada. The enlarged cardiomegaly is present on this film. Present. The cardiomegaly is present in the image. Defined as increased dimension of the cardiac silhouette or mediastinal compartment. Radiographic features			The enlarged cardiomegaly is present in the image. Defined as increased dimension of the cardiac silhouette or mediastinal compartment. Radiographic features
tip	The tip is present on the film. The Definition of tip. the location point of a device such as a The tip is present in the image. La definición de tip. punto de localización de un dispositivo médico. The tip is present on this film. The tip refers to The tip is present in the image. The Definition of tip. the distal endpoint of a medical device such as a catheter or tube. The Radiographic Features include a radiolucent			The tip is present in the image. The Definition of tip. the distal endpoint of a medical device such as a catheter or tube. The Radiographic Features include a radiolucent
pneumonia	The pneumonia is present on the film. The Definition of PNA. infection-induced inflammation. The pneumonia is present in the image. La definición de neumonía. inflamación del parénquima pulmonar. The pneumonia is present on this film. Present. The pneumonia is present in the image. Defined as infection-related inflammation of pulmonary parenchyma. The Radiographic Features of pneumonia include patchy			The pneumonia is present in the image. Defined as infection-related inflammation of pulmonary parenchyma. The Radiographic Features of pneumonia include patchy
line	The line is present on the film. The Definition of a line. a cath or G-tube course or a linear rod. The line is present in the image. La definición de línea. catéter o tubo gástrico como una estructura lineal. The line is present on this film. lines. line. The line is present in the image. Defined as a catheter or gastrostomy appearing as a linear radiopaque structure. The Radiographic Features of the line include a slender			The line is present in the image. Defined as a catheter or gastrostomy appearing as a linear radiopaque structure. The Radiographic Features of the line include a slender
cardiomegaly	The cardiomegaly is present on the film. The Definition of CM. enlarged heart size usually. The cardiomegaly is present in the image. La definición de cardiomegalia. aumento del tamaño cardíaco. The cardiomegaly is present on this film. Card. The cardiomegaly is present in the image. Defined as an enlarged cardiac size based on an increased cardiothoracic ratio. The Radiographic Features include an abnor			The cardiomegaly is present in the image. Defined as an enlarged cardiac size based on an increased cardiothoracic ratio. The Radiographic Features include an abnor
fracture	The fracture is present on the film. The Definition of a fr. disruption of bone continuity. The fracture is present in the image. La definición de fractura. interrupción de la continuidad ósea. The fracture is present on this film. A fracture. The fracture is present in the image. Defined as a cortical discontinuity indicating a break or fissure. Radiographic features include a lucent line traversing the bony co			The fracture is present in the image. Defined as a cortical discontinuity indicating a break or fissure. Radiographic features include a lucent line traversing the bony co
calcification	The calcification is present on the film. The Definition of calc. deposition of calcium salts. The calcification is present in the image. La definición de calcificación. depósito de sales de calcio en. The calcification is present on this film. Calcifi. The calcification is present in the image. Defined as a calcium salt deposition within soft tissues or vascular walls. Radiographic features include high attenuation hard li			The calcification is present in the image. Defined as a calcium salt deposition within soft tissues or vascular walls. Radiographic features include high attenuation hard li
engorgement	The engorgement is present on the film. The Definition of engorg. vascular dilation from hyp. The engorgement is present in the image. La definición de congestión vascular. dilatación de los va. The engorgement is present on this film. Cong. The engorgement is present in the image. Defined as vascular engorgement due to increased transvascular blood volume. Radiographic features include pronounced or			The engorgement is present in the image. Defined as vascular engorgement due to increased transvascular blood volume. Radiographic features include pronounced or
nodules	The nodules is present on the film. The Definition of a nodule. a small rounded opacity. The nodules is present in the image. La definición de nódulo. opacidad redondeada de menor. The nodules is present on this film. A nodule. The nodules is present in the image. Defined as a small, rounded parenchymal opacity measuring less than 1 cm. Radiographic features include a well-defined or			The nodules is present in the image. Defined as a small, rounded parenchymal opacity measuring less than 1 cm. Radiographic features include a well-defined or
air	The air is present on the film. The Definition of air. a free radiopaque element. Typically, The air is present in the image. La definición de aire. elemento radiopaco débil. The air is present on this film. Free. The air is present in the image. Defined as a slender radiolucent element commonly associated with medical hardware. Radiographic features include a thin radiopac			The air is present in the image. Defined as a slender radiolucent element commonly associated with medical hardware. Radiographic features include a thin radiopac
pacemaker	The pacemaker is present on the film. The Definition of a PPM. an implanted device that reg. The pacemaker is present in the image. La definición de marcapasos. dispositivo implantado que reg. The pacemaker is present on this film. A pace. The pacemaker is present in the image. Defined as an implanted cardiac rhythm management device delivering electrical impulses. Radiographic features include a ri			The pacemaker is present in the image. Defined as an implanted cardiac rhythm management device delivering electrical impulses. Radiographic features include a ri
pleural thickening	The pleural thickening is present on the film. The Definition of pleural thick. fib. thickening. The pleural thickening is present in the image. La definición de engrosamiento pleural. engrosamiento. The pleural thickening is present on this film. The pleural thickening is present in the image. Defined as fibrous thickening of the pleural surfaces. Radiographic features include irregular pleural thickening or pleu			The pleural thickening is present in the image. Defined as fibrous thickening of the pleural surfaces. Radiographic features include irregular pleural thickening or pleu
masking	The masking is present on the film. The Definition of a masking. a radiographic marker or an. The masking is present in the image. La definición de mascar. señal radiográfica o anotación que. The masking is present on this film. A masking. The masking is present in the image. Defined as a radiographic annotation or orientation marker covering patient information. Radiographic features include radiolucent			The masking is present in the image. Defined as a radiographic annotation or orientation marker covering patient information. Radiographic features include radiolucent
scar	The scar is present on the film. The Definition of a scar. fib. tissue replacing normal lung pap. The scar is present in the image. La definición de cicatriz pulmonar. tejido fibrotico que reemplaza al. The scar is present on this film. A scar. The scar is present in the image. Defined as fibrotic replacement of normal lung parenchyma following prior injury. Radiographic features include linear fibrotic bands or			The scar is present in the image. Defined as fibrotic replacement of normal lung parenchyma following prior injury. Radiographic features include linear fibrotic bands or
hyperinflation	The hyperinflation is present on the film. The Definition of hyperinfl. overexpansion of lung. The hyperinflation is present in the image. La definición de hiperinflación. sobreexpansión de los. The hyperinflation is present on this film. Hyper. The hyperinflation is present in the image. Defined as overexpansion of the lung lobe due to impaired or			The hyperinflation is present in the image. Defined as overexpansion of the lung lobe due to impaired or
blurring	The blurring is present on the film. The Definition of CTR blur. loss of the scale contrast. The blurring is present in the image. La definición de borrosidad del escala. pérdida de. The blurring is present on this film. Blurring. The blurring is present in the image. Defined as loss of the normal sharp contrast/contrast angle due to fluid accumulation or pleural thickening. Radiographic features include			The blurring is present in the image. Defined as loss of the normal sharp contrast/contrast angle due to fluid accumulation or pleural thickening. Radiographic features include
collapse	The collapse is present in the image. The Definition of atelectasis. partial complete de alve. The collapse is present in the image. Colapso means the lung section has lost its air and structure. The collapse is present in the image. Defined as complete loss of aeration within a lung segment resulting in regional volume loss. Radiographic features include dense parenchymal opacification with volume reduc			The collapse is present in the image. Defined as complete loss of aeration within a lung segment resulting in regional volume loss. Radiographic features include dense parenchymal opacification with volume reduc
engorgement	The engorgement is present on the film. The Definition of engorg. disruption of alveolar wa. The engorgement is present in the image. La definición de engorgimiento. destrucción de los parénquimas. The engorgement is present on this film. Engorg. The engorgement is present in the image. Defined as alveolar wall destruction resulting in irreversible enlargement of distal alveoli. Radiographic features include di			The engorgement is present in the image. Defined as alveolar wall destruction resulting in irreversible enlargement of distal alveoli. Radiographic features include di
anastomosis	The anastomosis is present on the film. The Definition of an. presence of air within pulmonary. The anastomosis is present in the image. La definición de anastomosis. presencia de aire dentro de los vasos. The anastomosis is present on this film. Anastom. The anastomosis is present in the image. Defined as normal or contrast white pulmonary structures. Radiographic features include radiolucent long parenchymal walls with gres			The anastomosis is present in the image. Defined as normal or contrast white pulmonary structures. Radiographic features include radiolucent long parenchymal walls with gres
mass	The mass is present on the film. The Definition of a mass. focal opacity. The mass is present in the image. La definición de masa. opacidad focal mayor de 1 cm que equivo. The mass is present on this film. A mass here. The mass is present in the image. Defined as a focal parenchymal opacity exceeding 1 cm, converting for a suspicious or stable lesion. Radiographic features include			The mass is present in the image. Defined as a focal parenchymal opacity exceeding 1 cm, converting for a suspicious or stable lesion. Radiographic features include
radiation	The radiation is present on the film. The Definition of rad. diffuse or patchy accumulation. The radiation is present in the image. La definición de radiación. afectación difusa o parcheada de. The radiation is present on this film. Radiation. The radiation is present in the image. Defined as diffuse or patchy involvement of lung parenchyma by inflammatory cells or fluid. Radiographic features include patch			The radiation is present in the image. Defined as diffuse or patchy involvement of lung parenchyma by inflammatory cells or fluid. Radiographic features include patch
abnormal	The abnormal finding is present on the film. The Definition of abnor. inability to recognize anat. The abnormal finding is present in the image. La definición de anomalía. incapacidad para. The abnormal finding is present on this film. Abn. The abnormal finding is present in the image. Defined as impaired visualization of normal anatomical structures. Radiographic features include loss of normal soft-tissue d			The abnormal finding is present in the image. Defined as impaired visualization of normal anatomical structures. Radiographic features include loss of normal soft-tissue d
defect	The defect is present on the film. The Definition of defect. abnormal normal image of struc. The defect is present in the image. La definición de defecto. alteración de la forma normal de la. The defect is present on this film. A defect. The defect is present in the image. Defined as alteration of the normal configuration of a line or soft tissue. Radiographic features include abnormal normal image			The defect is present in the image. Defined as alteration of the normal configuration of a line or soft tissue. Radiographic features include abnormal normal image
hernia	The hernia is present on the film. The Definition of a hern. protrusion of an organ through th. The hernia is present in the image. La definición de hernia. protrusión de un órgano a través de la pa. The hernia is present on this film. A hernia. The hernia is present in the image. Defined as protrusion of an organ through the body wall or a fluid-containing body's			The hernia is present in the image. Defined as protrusion of an organ through the body wall or a fluid-containing body's
drainage	The drainage is present on the film. The Definition of a drain. placement of a tube or cath. The drainage is present in the image. La definición de drenaje. colocación de un tubo o catéter. The drainage is present on this film. Drainage. The drainage is present in the image. Defined as placement of a tube or catheter intended to evacuate fluid or air. Radiographic features include a radiopaque tube wit			The drainage is present in the image. Defined as placement of a tube or catheter intended to evacuate fluid or air. Radiographic features include a radiopaque tube wit
distortion	The distortion is present on the film. The Definition of distort. expansion or dist. of a tube. The distortion is present in the image. La definición de distorsión. expansión o distorsión de una est. The distortion is present on this film. Distorted. The distortion is present in the image. Defined as expansion or distortion of a tubular anatomical structure. Radiographic features include distortion of the tubule or a non			The distortion is present in the image. Defined as expansion or distortion of a tubular anatomical structure. Radiographic features include distortion of the tubule or a non
displacement	The displacement is present on the film. The Definition of a shift. displacement of mediastinal or other area. The shift is present in the image. La definición de desplazamiento. movimiento del mediastino o otras. The shift is present on this film. A shift here. The shift is present in the image. Defined as displacement of mediastinal or adjacent anatomical structures. Radiographic features include deviation of the trachea or a			The shift is present in the image. Defined as displacement of mediastinal or adjacent anatomical structures. Radiographic features include deviation of the trachea or a
stent	The stent is present on the film. The Definition of a st. an expandable metal mesh. inserted in. The stent is present in the image. La definición de stent. dispositivo de malla expandible insertado para. The stent is present on this film. A stent here. The stent is present in the image. Defined as an expandable metallic mesh device designed to maintain luminal patency of vessel or airways. Radiographic features in			The stent is present in the image. Defined as an expandable metallic mesh device designed to maintain luminal patency of vessel or airways. Radiographic features in
lesion	The lesion is present on the film. The Definition of a les. an area of abnormal tissue that no. The lesion is present in the image. La definición de lesión. zona de tejido anormal que no es. The lesion is present on this film. A lesion here. The lesion is present in the image. Defined as a region of abnormal tissue that may represent benign or malignant pathology. Radiographic features include an area of			The lesion is present in the image. Defined as a region of abnormal tissue that may represent benign or malignant pathology. Radiographic features include an area of
hardware	The hardware is present on the film. The Definition of hardware. surgical impl. such as plate. The hardware is present in the image. La definición de hardware quirúrgico. implante como placa. The hardware is present on this film. Hardware. The hardware is present in the image. Defined as surgical metallic implants including plates, screws, or rods. Radiographic features include high-density components			The hardware is present in the image. Defined as surgical metallic implants including plates, screws, or rods. Radiographic features include high-density components
dilation	The dilation is present on the film. The Definition of dilap. widening of a tubular structure. The dilation is present in the image. La definición de dilatación. ensanchamiento de una estructura. The dilation is present on this film. Dilation here. The dilation is present in the image. Defined as enlargement of the lumen of a tubular structure such as a vessel or duct. Radiographic features include an abnormally			The dilation is present in the image. Defined as enlargement of the lumen of a tubular structure such as a vessel or duct. Radiographic features include an abnormally
apoptosis	The apoptosis is present on the film. The Definition of apop. relaxation of tissue material into. The apoptosis is present in the image. La definición de apoptosis. relajación de material estructural. The apoptosis is present on this film. Apoptosis. The apoptosis is present in the image. Defined as relaxation of tissue material into the alveoli or pulmonary parenchyma. Radiographic features include patchy con			The apoptosis is present in the image. Defined as relaxation of tissue material into the alveoli or pulmonary parenchyma. Radiographic features include patchy con

Figure 2: The generated 4 types of clinical prompt variations.

REFERENCES

- Carol Wu Anna Zawacki, Julia Elliott George Shih, ParasLakhani Mikhail Fomitchev, Mohanad Hussain, and Shunxing Bao Phil Culliton. Siim-acr pneumothorax segmentation. <https://kaggle.com/competitions/siim-acr-pneumothorax-segmentation>, 2019.
- Benedikt Boecking, Naoto Usuyama, Shruthi Bannur, Daniel C Castro, Anton Schwaighofer, Stephanie Hyland, Maria Wetscherek, Tristan Naumann, Aditya Nori, Javier Alvarez-Valle, et al. Making the most of text semantics to improve biomedical vision–language processing. In *European conference on computer vision*, pp. 1–21. Springer, 2022.
- Aurelia Bustos, Antonio Pertusa, Jose-Maria Salinas, and Maria De La Iglesia-Vaya. Padchest: A large chest x-ray image dataset with multi-label annotated reports. *Medical image analysis*, 66: 101797, 2020.
- Shivang Desai, Ahmad Baghal, Thidathip Wongsurawat, Piroon Jenjaroenpun, Thomas Powell, Shaymaa Al-Shukri, Kim Gates, Phillip Farmer, Michael Rutherford, Geri Blake, et al. Chest imaging representing a covid-19 positive rural us population. *Scientific data*, 7(1):414, 2020.
- Shih-Cheng Huang, Liyue Shen, Matthew P Lungren, and Serena Yeung. Gloria: A multimodal global-local representation learning framework for label-efficient medical image recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3942–3951, 2021.
- Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silvana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pp. 590–597, 2019.
- Alistair EW Johnson, Tom J Pollard, Nathaniel R Greenbaum, Matthew P Lungren, Chih-ying Deng, Yifan Peng, Zhiyong Lu, Roger G Mark, Seth J Berkowitz, and Steven Horng. Mimic-cxr-jpg, a large publicly available database of labeled chest radiographs. *arXiv preprint arXiv:1901.07042*, 2019.
- Maya Pavlova, Naomi Terhijan, Audrey G Chung, Andy Zhao, Siddharth Surana, Hossein Aboutaleb, Hayden Gunraj, Ali Sabri, Amer Alaref, and Alexander Wong. Covid-net cxr-2: An enhanced deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Frontiers in Medicine*, 9:861680, 2022.
- Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.
- Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2097–2106, 2017.
- Chaoyi Wu, Xiaoman Zhang, Ya Zhang, Yanfeng Wang, and Weidi Xie. Medklip: Medical knowledge enhanced language-image pre-training for x-ray diagnosis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 21372–21383, 2023.
- Linghua Wu, Jing Zhang, Yilin Wang, Rong Ding, Yueqin Cao, Guiqin Liu, Changsheng Liufu, Baowei Xie, Shanping Kang, Rui Liu, et al. Pneumonia detection based on rsna dataset and anchor-free deep learning detector. *Scientific Reports*, 14(1):1929, 2024.