**Experimental Analysis Steps[[1]](#endnote-1)**

**Summary of Experimental Flow:**

1. **Pretrained CNN**: Start with a CNN pre-trained on a large dataset.
2. **Fine-tuning**: Use an initial small, labeled set (~1% of your training data) to adapt this pre-trained model to your X-ray pneumonia detection task.
3. **Active Learning Phase**:
   * **Iteration 1**: Train the fine-tuned model on a small, labeled set.
   * **Subsequent Iterations**: Use acquisition functions to select new, informative images from the unlabeled pool, add them to the labeled set, and train the model further.
   * **Repeat**: Continue iterating through the active learning process, updating the model with each batch of new labeled samples.
4. **Performance Comparison:**The active learning approach is compared to a baseline model trained with random sampling. Improvements in performance are tracked across iterations, with the goal of achieving high accuracy using only ~10% of the labeled dataset.
5. **Final Evaluation**:  
   Once the target performance is reached, the model is evaluated on the test set to report overall performance metrics such as accuracy, sensitivity, and specificity, with a focus on pneumonia detection.

**Fully Detailed Experimental Flow:**

1. **Data Preparation and Preprocessing[[2]](#endnote-2):**
   * **Dataset Structure**: The dataset consists of 5,863 chest X-ray images organized into three sets: train, validation, and test. The images are categorized into two labels: "Normal" and "Pneumonia". The training set is imbalanced, with 1,341 normal and 3,875 pneumonia images, while the validation and test sets are more balanced.
   * **Preprocessing**: Apply common image processing techniques such as normalization, resizing, and possibly data augmentation (random flips, rotations) to standardize input images. These methods enhance image quality and create augmented versions for more robust training, especially given the class imbalance.
   * **Split Utilization**:
     + **Training**: 100% of the training set is used for initial representation learning.
     + **Validation**: The validation set (8 normal, 8 pneumonia) is reserved for model fine-tuning and hyperparameter selection.
     + **Test**: The test set (234 normal, 390 pneumonia) is used strictly for final model evaluation.
2. **Representation Learning with a Convolutional Neural Network (CNN):**
   * **Unsupervised Pretraining**: Leverage unsupervised learning with a CNN backbone (e.g., ResNet) to generate feature representations from the training and pool data. No labels are used in this step.
     + **Data Usage**: The full training set is utilized, including both labeled and unlabeled data, for generating robust feature representations.
   * **Representation Size**: The learned feature vectors (e.g., 2,048 dimensions) will be used as inputs for the classifier in subsequent steps.
   * **Fine-tuning with Labels**: Fine-tune the CNN model using the labeled data, starting from the small subset of labeled images. The initial labeled set consists of approximately 1% of the total dataset, selected randomly (or via acquisition function).
3. **Active Learning Framework:**
   * **Initial Labeled Set**: Begin active learning by selecting ~1% of the labeled training data (around 55-60 images) from the training set, split between the two classes.
   * **Active Learning Iterations**: In each iteration, use acquisition functions (entropy, uncertainty, or a hybrid of both) to select the most informative samples from the pool of unlabeled images. Labeling is done incrementally to reduce annotation cost while improving model performance.
     + **Unlabeled Pool**: The remaining 99% of the training set forms the pool of unlabeled samples for active selection.
     + **Training Set Growth**: At each step, the labeled set grows by a small percentage (~1% of the full dataset), incorporating new labeled samples from the pool based on the acquisition function.
4. **Classifier Training with a Hybrid CNN-GP Model:**
   * **GP Classifier**: Use the learned feature vectors from the CNN to train a Gaussian Process (GP) classifier, which is well-suited for imbalanced data due to its ability to model uncertainty in predictions.
     + **Data Usage**: The initial labeled set is used to train the GP model, and with each active learning iteration, more samples are added to refine the classifier.
   * **Evaluation and Performance Metrics**: After each iteration, evaluate the model on the test set. Track key metrics such as accuracy, sensitivity, specificity, and confusion matrix to assess the classifier’s performance on both the normal and pneumonia classes.
     + **Focus on Class Imbalance**: Special attention is given to how well the model performs on the pneumonia class, as this class is more prevalent in the dataset. The GP classifier should handle the class imbalance better than traditional classifiers like fully connected neural networks.
5. **Performance Comparison and Iterative Refinement:**
   * **Baseline Comparison**: Compare the performance of the active learning model against a baseline trained with random sampling. Highlight the benefits of using acquisition functions to select informative samples for labeling.
   * **Iteration Monitoring**: As active learning progresses, monitor improvements in the model’s performance with each batch of newly labeled samples. The goal is to achieve high accuracy with significantly fewer labeled samples (e.g., ~10% of the full dataset).
   * **Final Evaluation**: Once the target performance is reached (e.g., a certain accuracy threshold), conduct a final evaluation on the test set to report overall performance metrics.

1. Suggested experimental setup is based on and influenced from the paper "[Highly Efficient Representation and Active Learning Framework for Imbalanced Data and its Application to COVID-19 X-Ray Classification](https://europepmc.org/article/ppr/ppr336925#full-text-links)".   
   Hao, H., Didari, S., Woo, J. O., Moon, H., & Bangert, P. (2021). Highly efficient representation and active learning framework for imbalanced data and its application to covid-19 x-ray classification. [↑](#endnote-ref-1)
2. <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia/data> [↑](#endnote-ref-2)