# VinBigData Project Challenge

# Introduction to the VinBigData Challenge and data

nature › scientific data › data descriptors › article

Data Descriptor | Open access | Published: 20 July 2022

## VinDr-CXR: An open dataset of chest X-rays with radiologist's annotations

Ha Q. Nguyen, Khanh Lam, Linh T. Le, Hieu H. Pham ✉, Dat Q. Tran, Dung B. Nguyen, Dung D. Le, Chi M. Pham, Hang T. T. Tong, Diep H. Dinh, Cuong D. Do, Luu T. Doan, Cuong N. Nguyen, Binh T. Nguyen, Que V. Nguyen, Au D. Hoang, Hien N. Phan, Anh T. Nguyen, Phuong H. Ho, Dat T. Ngo, Nghia T. Nguyen, Nhan T. Nguyen, Minh Dao & Van Vu

*Scientific Data* **9**, Article number: 429 (2022) | Cite this article

**21k** Accesses | **85** Citations | **2** Altmetric | Metrics

VINGROUP BIG DATA INSTITUTE · FEATURED PREDICTION COMPETITION · 3 YEARS AGO

## VinBigData Chest X-ray Abnormalities Detection

Automatically localize and classify thoracic abnormalities from chest radiographs

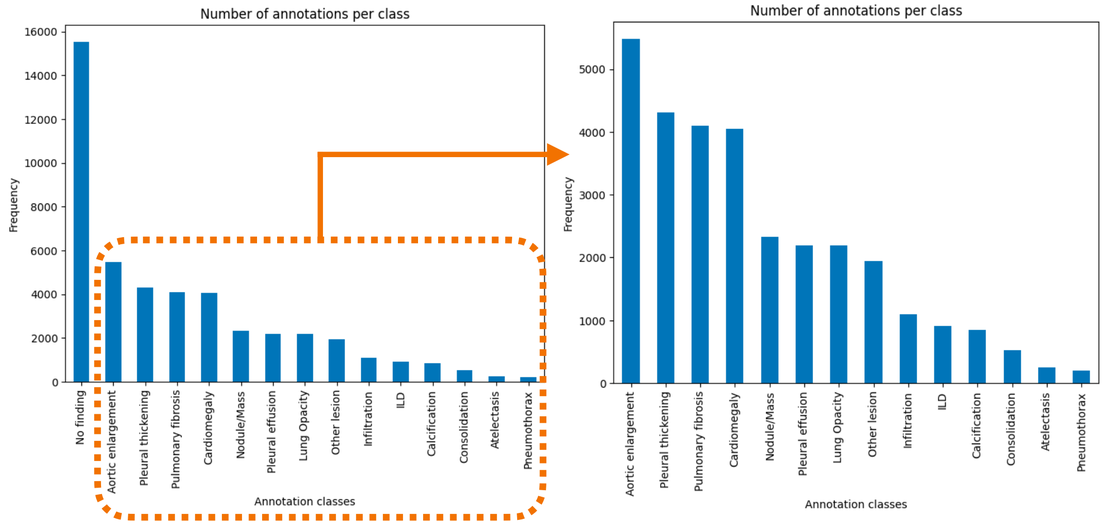
**Scope:** **Improve detection of pulmonary diseases 🫁 (including outliers such objects) through AI**

**Dataset size:** **15k 🩻 (57.2% train - 42.8% test)**

*Total of 45925 annotations in training (multiple annotations per image) and 21989 in test*

**Dataset labels:** **14 abnormalities Labels + No Finding**

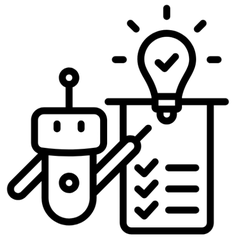
**Clearly unbalanced data!**



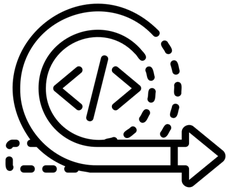
# Strategy I

## The simplest the best

Start from the ground and build from that



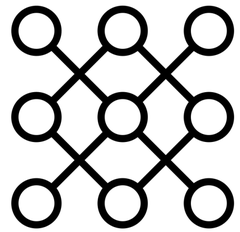
## Explainability first



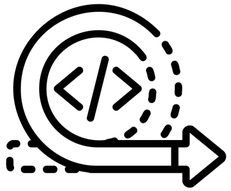
## Agile methodology

# Strategy II

Final Model

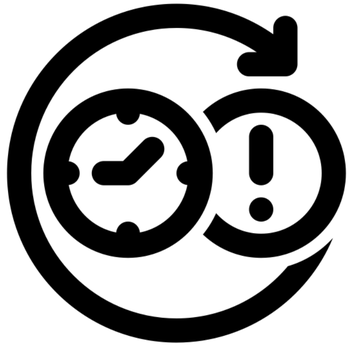




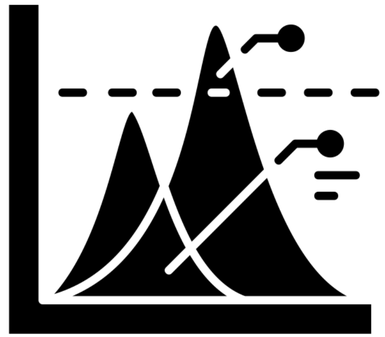



**2 models to achieve the better performance**

# Challenges and Limitations

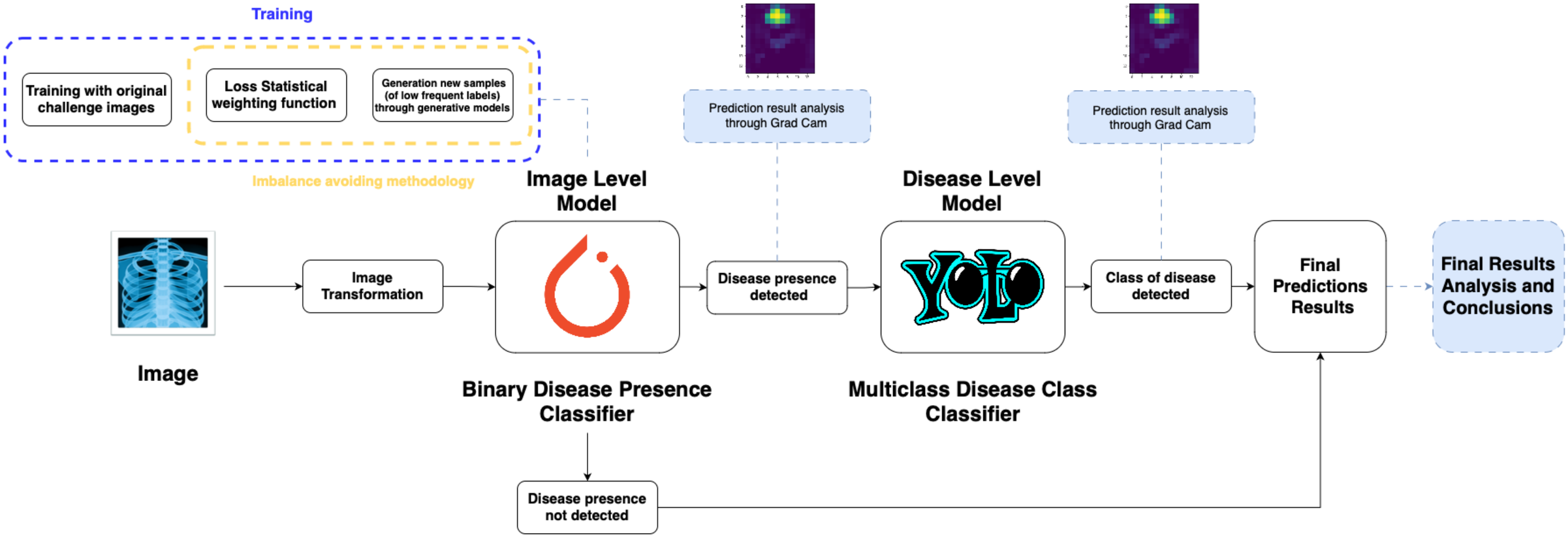


**GPU Time Limit**



**Sophisticated metrics**

# Upcoming Steps



# VinBigData Challenge 🫁

# Thank you!