Przygotowujemy z Weroniką i Kubą papier na workshop

Beyond Fairness: Towards a Just, Equitable, and Accountable Computer Vision CVPR 2021 Workshop

https://sites.google.com/view/beyond-fairness-cv/call-for-papers

- ***Computer vision in practice: Who is benefitting and who is being harmed?*** We welcome submissions examining how computer vision technology intersect with, and amplifies, structural inequality and how the research practices and incentive structures within the academic field are implicated in the harms to marginalized groups. We welcome submissions that critically examine how one's own research and their roles within the community can contribute to a strengthening or dismantling of existing systems of power.

- ***Cross-disciplinary research methods and methodologies.*** Computer vision is simultaneously a social and technical endeavor. Yet, current computer vision education and publication incentives tend to valorize the technical and devalue the social. This knowledge hierarchy is directly implicated in the harms being perpetuated by the field. We encourage submissions that introduce methods and methodologies that have been developed in other fields and by communities experiencing marginalization.

- ***Accountability and transparency.*** The field of computer vision is currently facing a crisis of accountability. Computer vision systems are being developed and deployed at a rapid pace, often in highly socially consequential domains. Yet, computer vision models and datasets are frequently developed with little transparency into the design and development process, and few mechanisms of accountability, contestability, or recourse for individuals impacted by the systems. **We encourage submissions that audit models or datasets, that examine model or dataset development processes, and that introduce frameworks that promote transparent and accountable model and dataset development** (e.g. Model Cards, Datasheets). We also welcome submissions that explore ethical obligations of researchers and discuss mechanisms of ethical oversight within academic research.

- ***Activism and collective organizing*****.** We believe computer vision experts have an important role to play in shifting public discourse and public policy regarding the use of computer vision technologies, and shifting the research culture, norms, and incentive structures in a manner that will ultimately promote more responsible research practices. We hope to empower researchers to take an active role in all these realms. We welcome submissions that examine the institutional barriers within computer vision that are contributing to the extreme concentration of power within the field, in an effort to better diagnose the current condition of the field, and provide insights into how to best effect change.

- ***Historical perspectives.*** Computer vision methods have advanced rapidly in recent years. However, many of the practical applications of computer vision methods have long histories that predate the field. These histories can provide valuable insight into the latent assumptions and ideologies underlying modern computer vision tools as well as the harms these tools can cause to marginalized groups. We encourage submissions that offer historical perspectives on the field of computer vision, including, but not limited to, histories of computer vision datasets and the trajectory of various applications across time.

Ongoing Kaggle challenge on CXR

- Aktualny – kończy się za pare dni
- Zawiera labelki z lokalizają. Istnieje wiele zbiorów, nawet liczniejszych które jednak nie są olableowane z lokalizacją (tylko opisy)

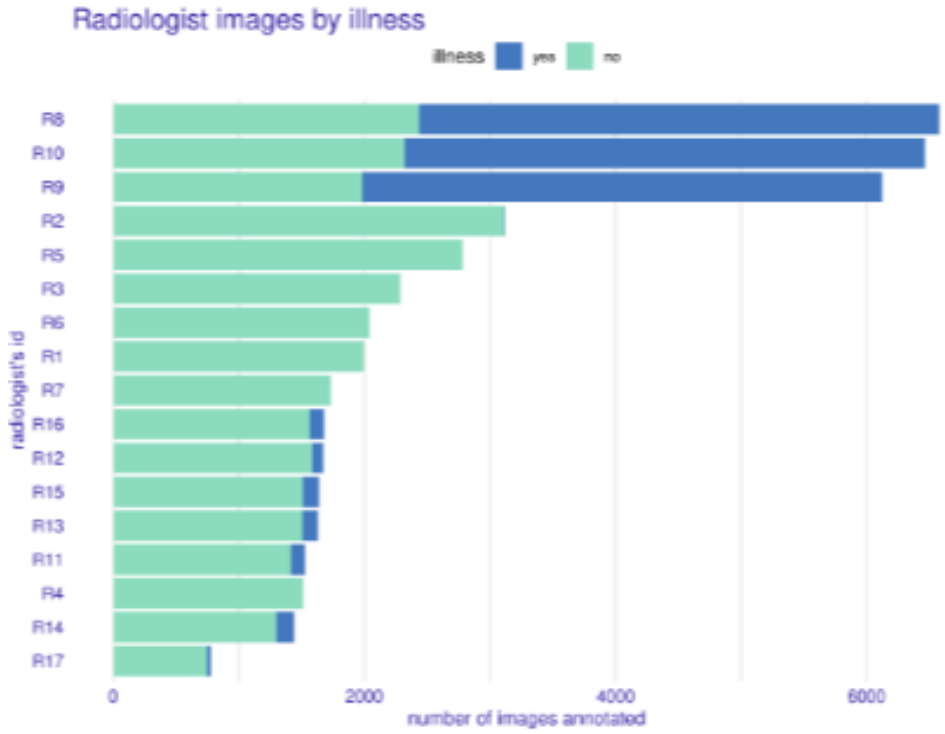
CXR datast

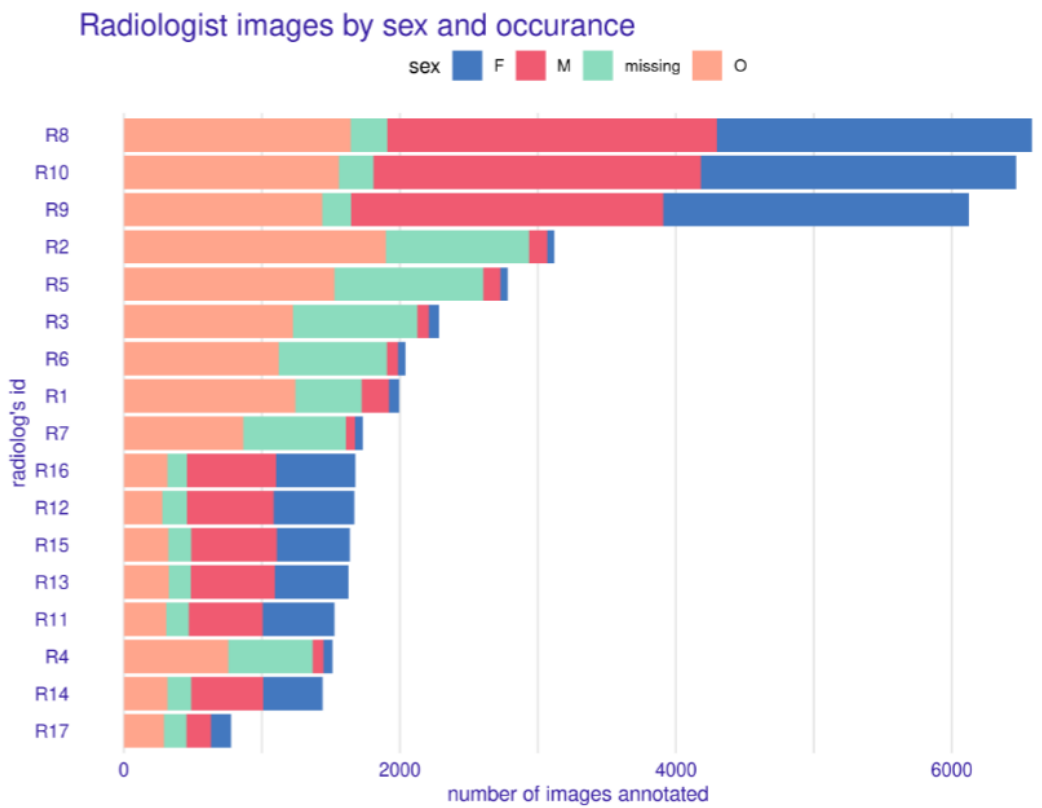
- 17 radiologist, 8 years experience at least
- The training set contains 15,000 lung images in DICOM format with annotations. Each image was annotated by three radiologists.
- There are fourteen labels for lesions and one additional label for images of healthy lungs
- Test set has 3,000 dicom files with no annotations as the challenge is ongoing.

Missing information in DICOM

- 68% of the observations do not have information about age, and 17% about sex
- dataset is fairly bal-anced (M: 26%, F: 23%)
- This leaves us with only 25%of images with valid ages between 1-99. This leaves uswithout some valid information as age and sex might becorrelated with certain diseases

Unequal division of annotation work between radiologists

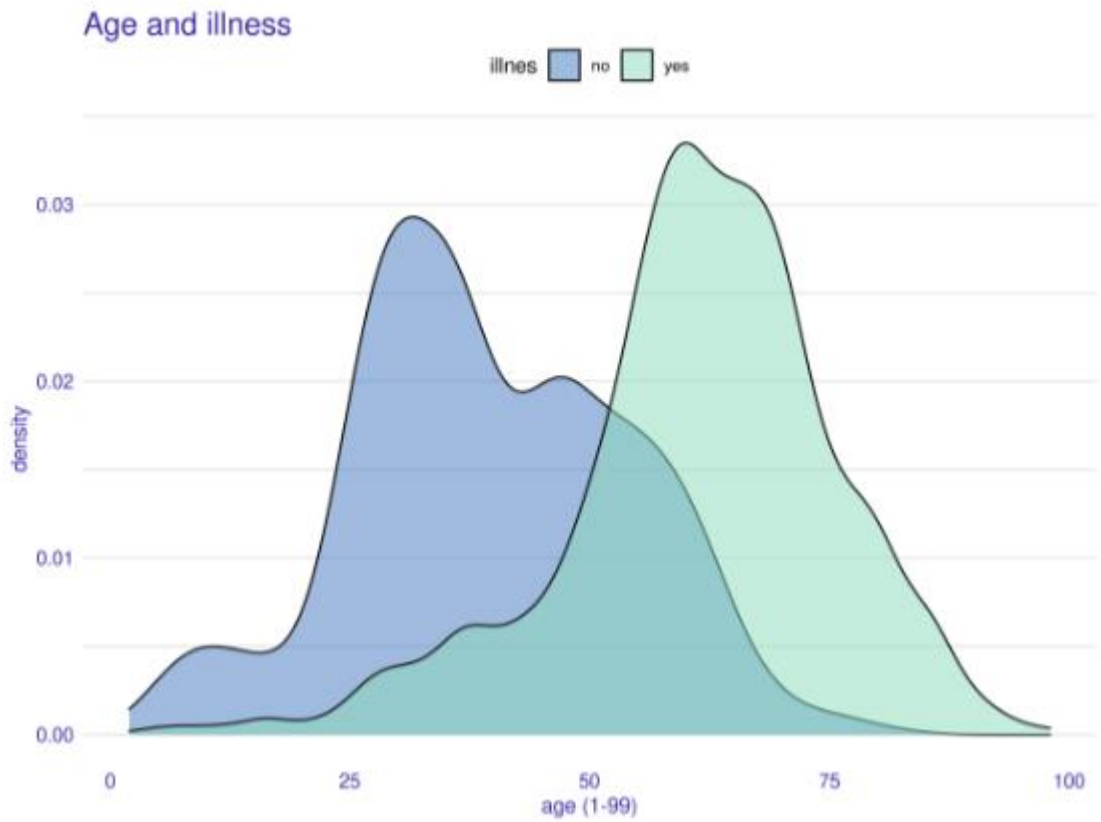




Low consistency in annotations between radiologists

Agreement	R1-R7	R8-R10	R11-R17
Agreed with at least one colleague on all classes	100%	69%	96%
Agreed with both colleagues on all classes	100%	46%	94%
When proposing classes for record at least one colleague agrees	100%	77%	97%
When proposing classes for record both colleagues agree	100%	61%	95%

Age/illness shift



Different procedure of preparing train and test set

- In the whole dataset, the annotations were made independently by three radiologists for each record.
  - Problem: considerable differences between radiologists.
- in the test set, the labels were verified one more time and a consensus between radiologists was reached (by two additional radiologists).
  - Problem: two radiologists perform consensus

Lesions present on chests with ”no findings” label

- There are typical findings for older patients – many times some radiologists did not mark such.
- wielkości boxow, tj czy zaznaczany jeden duzy box, czy dwa mniejsze (np jeden duzy box obejmuje całą klatke piersiową, albo dwa mniejsze osobno na oba pluca)

Different label for the same pathology

- Nasi radiolodzy zauważyli, że 3 radiologów ze zbioru oznacza te same patologie inną labelką
- Labelki mają niejednoznaczne definicje (definicje zmian pokrywają się bazując na artykule opisującym dataset.
  - Lung opacity, pokrywa 6 innych klas
  - ILD i Pulmonary fibrosis – pokrywają się
  - Consolidation i Infiltration – pokrywają się

Parts of clothes present in the X-rays.

Figure 4. Example of clothes artifacts. From the left there are: ZIP fastener, bra, buttons, necklace

Lesions localization imbalance

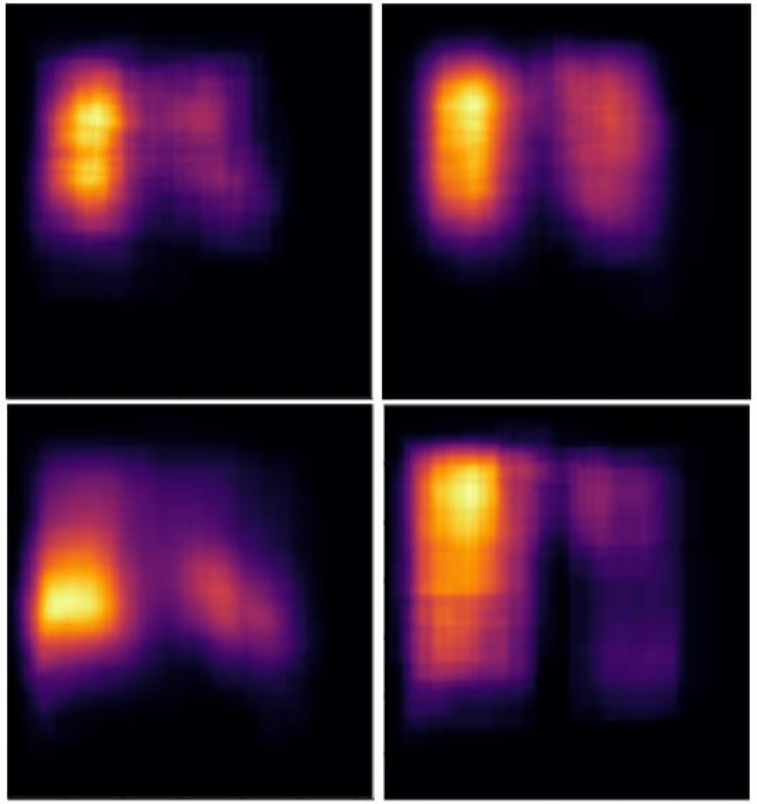


Figure 3. Example of anomalies which should be symmetric. From the top left there are: Consolidation, Infiltration, Pleural Effusion, Pneumothrox

ongoing

- How annotations quality can influence model usage

- Analiza typowego modelu pod kątem jakich zależności się uczy.
  - Object detector pretrenowany na normalnych obrazkach i dotrenowany do wybranych klas medycznych na zbiorze treningowym

- baseline model do vinbigdata-kaggle-challenge

-> input image size 256x256, augmentacje (HorizontalFlip, ShiftScaleRotate, RandomBrightnessContrast), 10ksteps (3h na GTX 1050 Ti) -> 0.19 mAP-40 na testset, 0.117 mAP-40 na kaggle-test-set czyli miejsce 793

-> input image size 1024x1024, augmentacje (HorizontalFlip, ShiftScaleRotate, RandomBrightnessContrast), 10ksteps (8h na GTX 1050 Ti) -> 0.21 mAP-40 na testset, 0.128 mAP-40 na kaggle-test-set czyli miejsce 770

-> aktualne problemy:
+ jest duża różnica między testset a kaggle-test-set -> planowane przejscie na dane tylko 3 super radiologów
+ dodanie innych kanałów na wejście czyli 8-high-bits i 8-low-bits as additiona input channels
+ retrening backbone-a od zera na danych medycznych w prostym zadaniu klasyfikacyjnym (bez potrzeby labelek), dodatkowe
+ augmentacje (CoarseDropout, RandomResizeCrop, ElasticDeform)
+ sprawdzenie podejścia dwu-stopniowego: 1-klasyfikacja no-finding/findings 2-lokalizacja i klasyfikcaja findings
+ filtracja danych wejściowych uzględniając wiek/warunki akwizycji/ułozenie pacjenta PA/AP.