Past radiological exams alongside expert annotations improve **long-range dependency** and **generalisation** of image-report generation model

Automated Enriched Medical Concept Generation for Chest X-ray Images

Aydan Gasimova

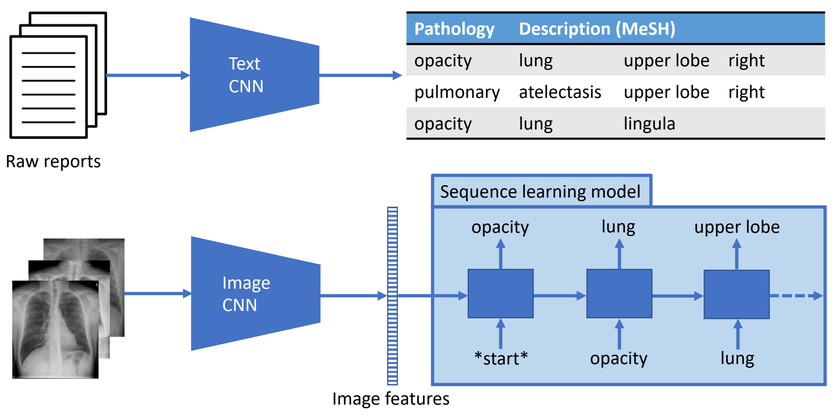
INTRO

- Past radiological exams have the potential to be used as training data for automated radiology report generation
- Motivation: large quantities of data, expert annotations, contain context to pathology such as **anatomical location** and **severity**
- Challenges: raw reports are noisy, unstructured, contain non visually-significant information

PROPOSED SOLUTION

1. Train model to extract vocab-controlled key visual concepts from raw reports (Medical Subject Headings, MeSH)
2. Use model predictions as augmented dataset for structured report generation
3. Train sequence model conditioned on image features for automated key medical concept generation

METHODS



Pathology	Description (MeSH)		
opacity	lung	upper lobe	right
pulmonary	atelectasis	upper lobe	right
opacity	lung	lingula	

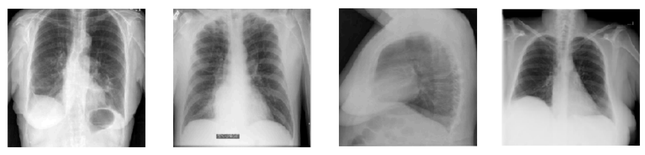
RESULTS

	BLEU-1 Train/Val/Test	BLEU-2 Train/Val/Test	BLEU-3 Train/Val/Test	BLEU-4 Train/Val/Test
Learning to Read Chest X-rays[1]	**97.2**/68.1/**79.3**	**67.1**/30.1/9.1	14.9/5.2/0.0	2.8/1.1/0.0
RNN1 + resnet50, 1,000 MeSH	92.6/24.3/31.6	55.6/13.2/15.2	**37.2**/7.0/7.2	**24.0**/4.7/3.5
RNN1 + resnet50, 1,000 MeSH + TextCNN predictions	73.6/41.5/41.6	50.0/29.7/28.2	30.9/15.9/13.2	17.8/7.2/8.1
RNN1 + resnet50, all MeSH	83.3/**68.5**/70.1	47.5/**52.1**/**49.5**	30.0/**29.9**/**27.2**	19.2/**16.9**/**16.8**

DISCUSSION

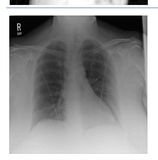
- Achieves higher BLEU for larger *n*-grams -> maintained better visual correspondence when generating longer sequences of words
- Higher BLEU on val/test than [1] as the model is trained end-to-end (no error propagation caused by cascaded training) -> improved generalisation
- Higher BLEU on val/test than only training on sub-sampled GS annotations -> reduced overfitting

SAMPLE CONCEPT GENERATION

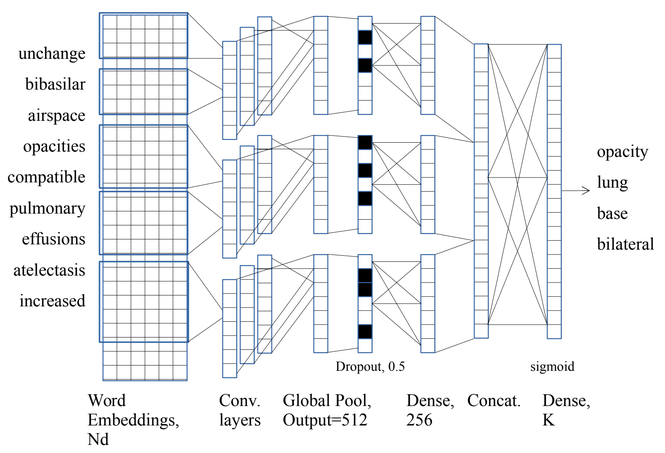


Gold-standard (GS) MeSH	**cicatrix, lung, lower lobe, left**	**opacity, lung, lower lobe, right, mild**	**atherosclerosis, aorta**	**density, costophrenic angle, right**
TextCNN predictions	lower lobe, lung, left	opacity, lung, right, mild, multiple	atherosclerosis, aorta	pulmonary atelectasis, right
RNN1 + resnet50, 1,000 MeSH + textCNN predictions	cicatrix, lung, lower lobe, left	markings, bronchovascular	granulomatous disease	density, costophrenic angle, right
RNN1 + resnet50, all MeSH	cicatrix, lung, lower lobe, left	opacity, lung, base, bilateral	normal	density, costophrenic angle, right

CHEST X-RAY DATA OPENI[2]

X-ray image	Report	MeSH
	The heart size and mediastinal contours appear within normal limits. There is blunting of the right lateral costophrenic sulcus which could be secondary to a small effusion versus scarring. No focal airspace consolidation or pneumothorax. No acute bony abnormalities.	Costophrenic Angle, right, blunted
	The cardiomediastinal silhouette and pulmonary vasculature are within normal limits. There is no pneumothorax or pleural effusion. There are no focal areas of consolidation. There are calcifications projecting of the left midlung, unchanged from prior, this is is XXXX sequela of prior granulomatous disease. There are small T-spine osteophytes.	Calcinosis, lung, lingual Granulomatous Disease Osteophyte, thoracic vertebrae, multiple, small
	Normal cardiomediastinal silhouette. Interval improvement in lung volumes bilaterally. Improved aeration of the right and left lung bases. Bilateral small pleural effusions and left base atelectatic change, with interval improvement. Visualized XXXX of the chest XXXX are within normal limits.	Pleural Effusion, bilateral, small Pulmonary Atelectasis, base, left

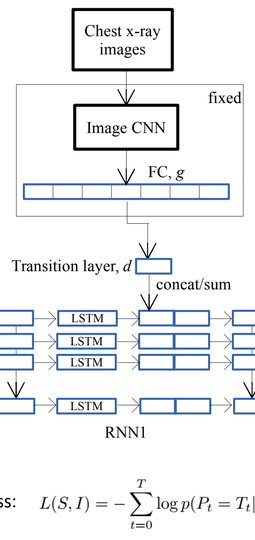
TRAINING AND MODEL DETAILS



Architecture of Multi-Instance Multi-Label Text Classification for MeSH Extraction

Loss:

$$\widehat{SCE}_i = -\lambda_1 \sum_{j=1}^{K} (y_j log(f(s_{ij})) + (1 - y_j log(1 - f(s_j))) - \lambda_2 \sum_{j=1}^{K} y_j f(s_j) / \sum_{j=1}^{K} (y_j f(s_j) + y_j (1 - f(s_j))) - \lambda_3 \sum_{j=1}^{K} (1 - y_j)(1 - f(s_j)) / \sum_{j=1}^{K} ((1 - y_j)(1 - f(s_j)) + (1 - y_j)(f(s_j)))$$



Architecture for Sequence Model for MeSH Generation

LSTM hidden and memory states at time step *t*:

$$h_t = f_t \odot h_{t-1} + i_t \odot \tanh(W^{(hx)} x_t + W^{(hm)} m_{t-1})$$
$$m_t = o_t \odot \tanh(h_t)$$

Image embedding CNN(I) extracted from final spatial avg. pooling layer of pre-trained CNN, passed through Transition layer and combined with output of recurrent layer:

$$\boldsymbol{dec}_t = \text{relu}(W^{(z)}(o_t * \text{relu}(W^{(dg)}\text{CNN}(I)))$$

Loss: $L(S, I) = -\sum_{t=0}^{T} \log p(P_t = T_t | \text{CNN(I)}, P_0 \ldots P_{t-1})$

[1] Shin, H. C., Roberts, K., Lu, L., Demner-Fushman, D., Yao, J., & Summers, R. M. (2016). Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2497-2506).
[2] https://openi.nlm.nih.gov

Imperial College London

BioMedIA

EPSRC Pioneering research and skills

paper

code