Knowledge-Guided Multiview Deep Curriculum Learning for Elbow Fracture Classification

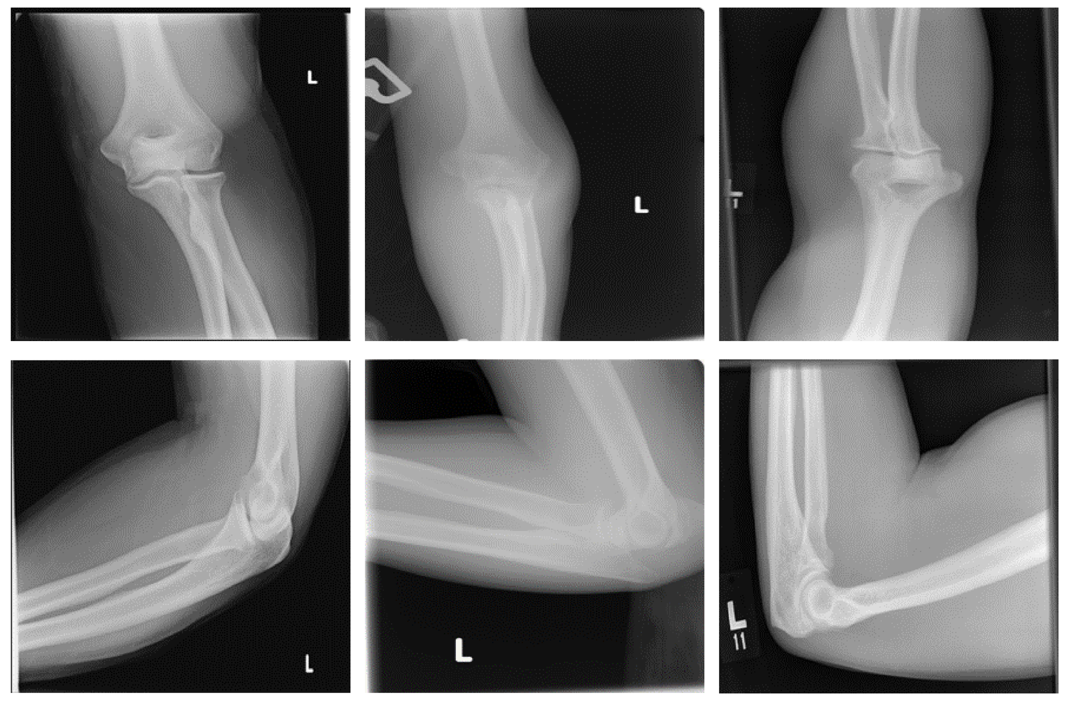
Jun Luo, Gene Kitamura, Dooman Arefan, Emine Doganay, Ashok Panigrahy, Shandong Wu
University of Pittsburgh, Pittsburgh, PA, USA (jul117@pitt.edu)
University of Pittsburgh Medical Center, Pittsburgh, PA, USA (wus3@upmc.edu)

MLMI @ MICCAI2021

Introduction

- Potential elbow fracture patients are often required to take both frontal view and lateral X-rays. In practice, having a single view is also common.

Normal | Ulnar fracture | Radial fracture
Frontal view
Lateral view

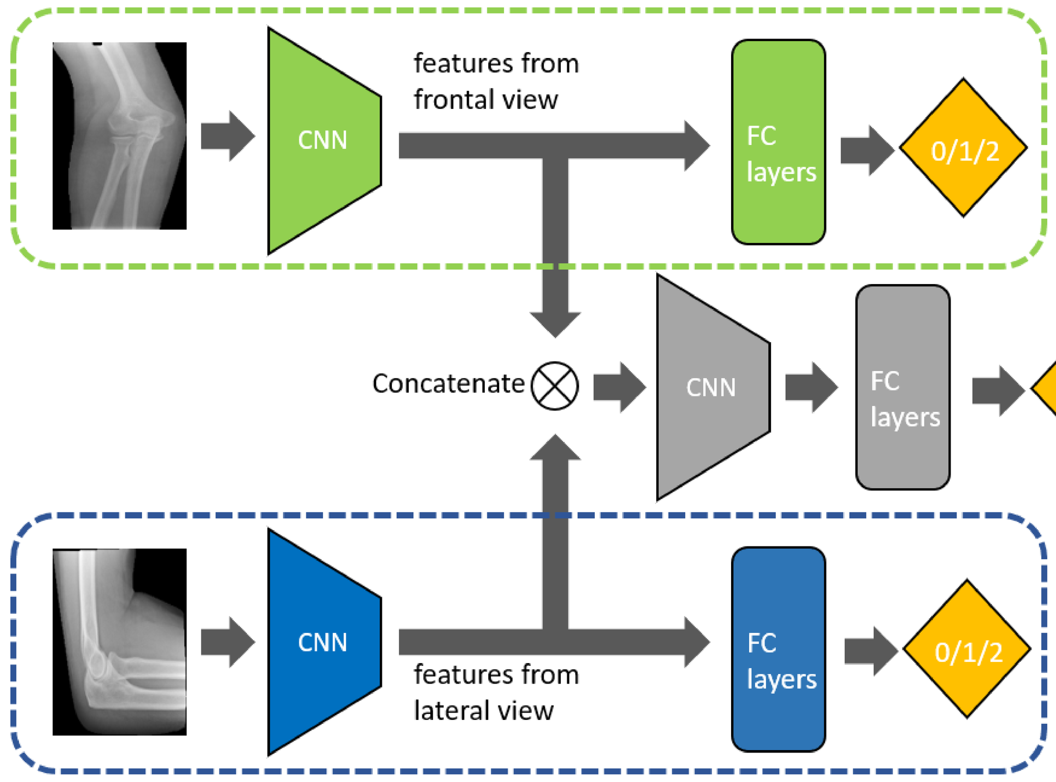


- Deep learning methods facilitate automation of elbow fracture diagnosis. Few existing methods leverage multiview information.
- We propose a multiview deep learning network architecture for elbow fracture subtype classification that takes frontal and lateral view elbow radiographs.
  - ✓ Dual-view architecture, flexible inference (infer from images from either one view or two views)
  - ✓ Homogeneous transfer learning from single view models
  - ✓ Curriculum learning guided by quantified medical knowledge
- Evaluation of our method:
  - ✓ Conduct experiments on a classification task of three classes of elbow fractures:
    - Normal
    - Ulnar fracture
    - Radial fracture

Method

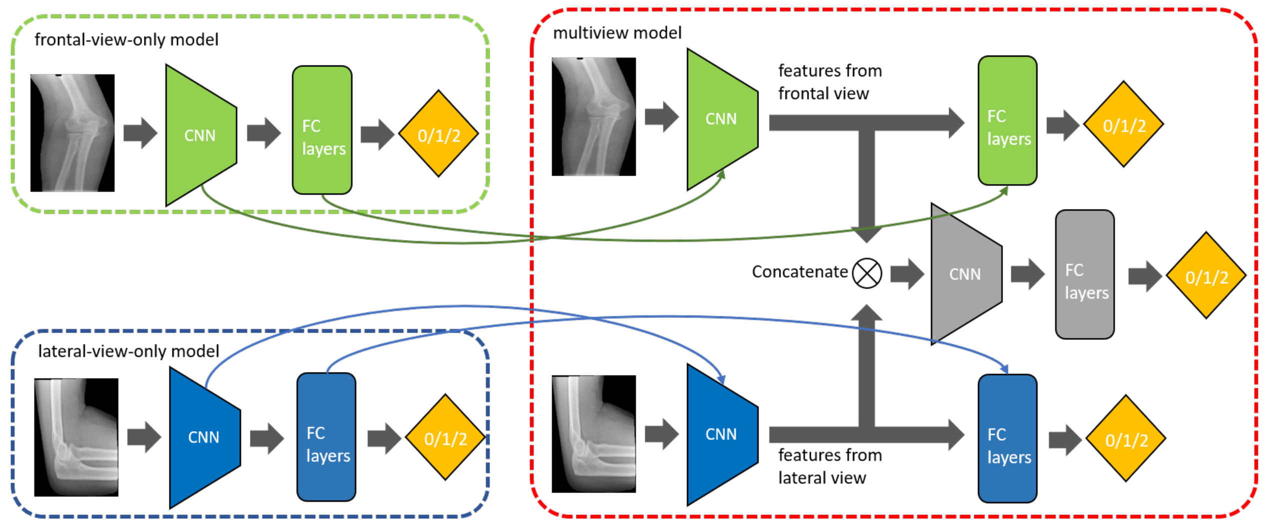
- Multiview model architecture
  - ✓ Model consists of three modules:
    - $\mathcal{F}$, frontal view module (green dotted line box)
    - $\mathcal{L}$, lateral view module (blue dotted line box)
    - $\mathcal{M}$, merge module (middle branch)
  - ✓ During training, a data sample triplet $\mathcal{D}_i$ with frontal image $x_i^{(F)}$, lateral image $x_i^{(L)}$ and label y_i generates loss:

$$J_\theta(\mathcal{D}_i) = J_{\theta_\mathcal{F}}\left(x_i^{(F)}, y_i\right) + J_{\theta_\mathcal{L}}\left(x_i^{(L)}, y_i\right) + J_{\theta_\mathcal{M}}\left(x_i^{(F)}, x_i^{(L)}, y_i\right)$$



- ✓ During testing
  - if both frontal and lateral view images are presented, the predicted label comes from $\mathcal{M}$.
  - Otherwise, the predicted label comes from the corresponding module of the input ($\mathcal{F}$ or $\mathcal{L}$).

- Homogeneous transfer learning
  - ✓ Train two single-view models
    - A frontal view model
    - A lateral view model
  - ✓ Transfer the trained weights to corresponding layers of multiview model (links in graph)
    - Convolutional and FC layers' weights of single-view model to corresponding module of multiview model.



- Knowledge-guided curriculum learning
  - ✓ Quantified medical knowledge into scores representing classification difficulty of certain fracture subtype

	Normal	Ulnar	Radial
Frontal view only	30	30	30
Lateral view only	35	60	45
Both views	45	65	55

(scores given by radiologist, 1=hardest; 100=easiest)

Method (cont'd)

- ✓ Permute training set at the beginning of every epoch
  - Permutation by sampling without replacement
  - Sampling probability at epoch e of sample i with score s_i is computed by

$$p_i^{(e)} = \begin{cases} \dfrac{s_i}{\sum_k s_k} & e = 1 \\ p_i^{(e-1)} \cdot \sqrt[E']{\dfrac{1/N}{p_i^{(1)}}} & 2 \le e \le E' \\ 1/N & E' < e \le E \end{cases}$$

Experiments

- Dataset
  - ✓ 982 subjects, each with a frontal and a lateral view X-ray image, 1,964 images in total
    - 500 non-fracture (normal) cases
    - 98 ulnar fractures cases
    - 384 radial fracture cases
  - ✓ 8-fold cross validation
- Metrics
  - ✓ Accuracy (denoted as Acc.) & AUC
  - ✓ Balanced accuracy (mean of # true positive / # samples of each class, denoted as Bal'd acc.)
  - ✓ Binary task accuracy (normal vs. fracture)
  - ✓ Binary task AUC (normal vs. fracture)
- Compared methods
  - ✓ Single-view model
  - ✓ [1] Jiménez-Sánchez et al., 2020
  - ✓ [2] Luo et al., 2021
  - ✓ Multiview with standard training
  - ✓ Multiview with different combination of proposed learning strategies

  (denote proposed transfer learning and curriculum learning strategy as TL and CL respectively)
- Compared with other methods (see results in next column),
  - ✓ with dual-view input, our method achieves the highest AUC and balanced accuracy with a margin of up to 0.118
  - ✓ with frontal view as only input, our method reaches highest performance per each metric
  - ✓ with lateral view as only input, our method has competitive performance.

Experiments (cont'd)

- Dual-view input and single-view input results

Model	Acc.	AUC	Bal'd acc.	Binary task acc.	Binary task AUC
Dual-view input					
Single-view-frontal	0.683	0.807	0.570	0.732	0.813
Single-view-lateral	0.856	0.954	0.807	0.895	0.959
Multiview	0.854	0.958	0.796	0.884	0.964
Multiview + TL	**0.891**	0.966	0.847	**0.916**	0.973
Multiview + [1]	0.818	0.939	0.746	0.864	0.952
Multiview + [1] + TL	0.870	0.961	0.811	0.898	0.973
Multiview + CL	0.889	0.970	0.847	0.908	**0.978**
Multiview + CL + TL	0.889	**0.974**	**0.864**	0.910	0.976
One-view input (frontal)					
Single-view-frontal	0.720	0.828	0.593	0.761	0.844
Single-view + CL [2]	0.683	0.807	0.570	0.732	0.813
Multiview	0.658	0.749	0.514	0.702	0.766
Multiview + TL	0.738	0.827	0.617	0.774	0.829
Multiview + [1]	0.566	0.675	0.396	0.575	0.648
Multiview + [1] + TL	0.737	0.815	0.605	0.773	0.831
Multiview + CL	0.723	0.814	0.602	0.761	0.823
Multiview + CL + TL	**0.756**	**0.829**	**0.636**	**0.786**	**0.846**
One-view input (lateral)					
Single-view-lateral	0.856	0.954	0.807	**0.895**	0.959
Single-view + CL [2]	0.840	0.946	0.809	0.872	0.948
Multiview	0.844	0.951	0.800	0.870	0.956
Multiview + TL	0.848	0.954	0.804	0.876	0.961
Multiview + [1]	0.837	0.945	0.779	0.870	0.949
Multiview + [1] + TL	**0.857**	**0.960**	**0.819**	0.885	**0.969**
Multiview + CL	0.838	0.956	0.807	0.867	0.956
Multiview + CL + TL	0.840	0.955	0.794	0.874	0.960

Conclusion

- Our method leverages multiview information for elbow fracture & incorporates medical knowledge.
- Our method outperforms the compared methods, and inference functions seamlessly on multiview input and single-view input.

Acknowledgements

- NIH/NCI 1R01CA193603, 1R01CA218405
- Pittsburgh Center for AI Innovation in Medical Imaging (Pitt Momentum Funds)
- Amazon Machine Learning Research Award.