24TH INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING & COMPUTER ASSISTED INTERVENTION
September 27 - October 1, 2021 • Strasbourg, FRANCE
MICCAI 2021 VIRTUAL

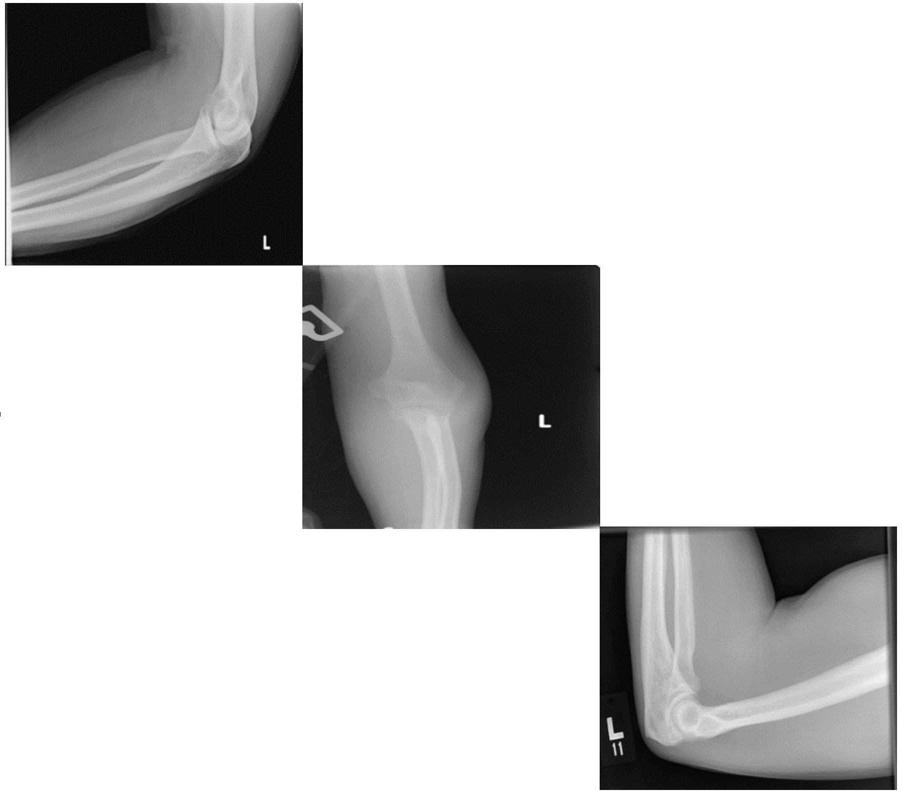
Knowledge-Guided Multiview Deep Curriculum Learning for Elbow Fracture Classification

Jun Luo[1], Gene Kitamura[2], Dooman Arefan[2], Emine Doganay[2], Ashok Panigrahy[2,3], and Shandong Wu[1,2,4]

[1]Intelligent Systems Program / [2]Dept. of Radiology / [4]Dept. of Biomedical Informatics and Dept. of Bioengineering, University of Pittsburgh, Pittsburgh, PA USA
[3]University of Pittsburgh Medical Center Children's Hospital of Pittsburgh, Pittsburgh, PA USA

Background

- Potential elbow fracture patients are often required to take both frontal view and lateral X-rays.
- Frontal view: distal humerus, proximal ulna, radius
- Lateral view: coronoid process, olecranon process
- In practice, it is also common only a single view is available/acquired.



- Deep learning has been shown effective in bone fracture detection and diagnose.
- Few has leveraged multiview information of elbow for deep learning.
- Clinical knowledge on qualitative imaging interpretation exists but usually is ignored in data-driven learning

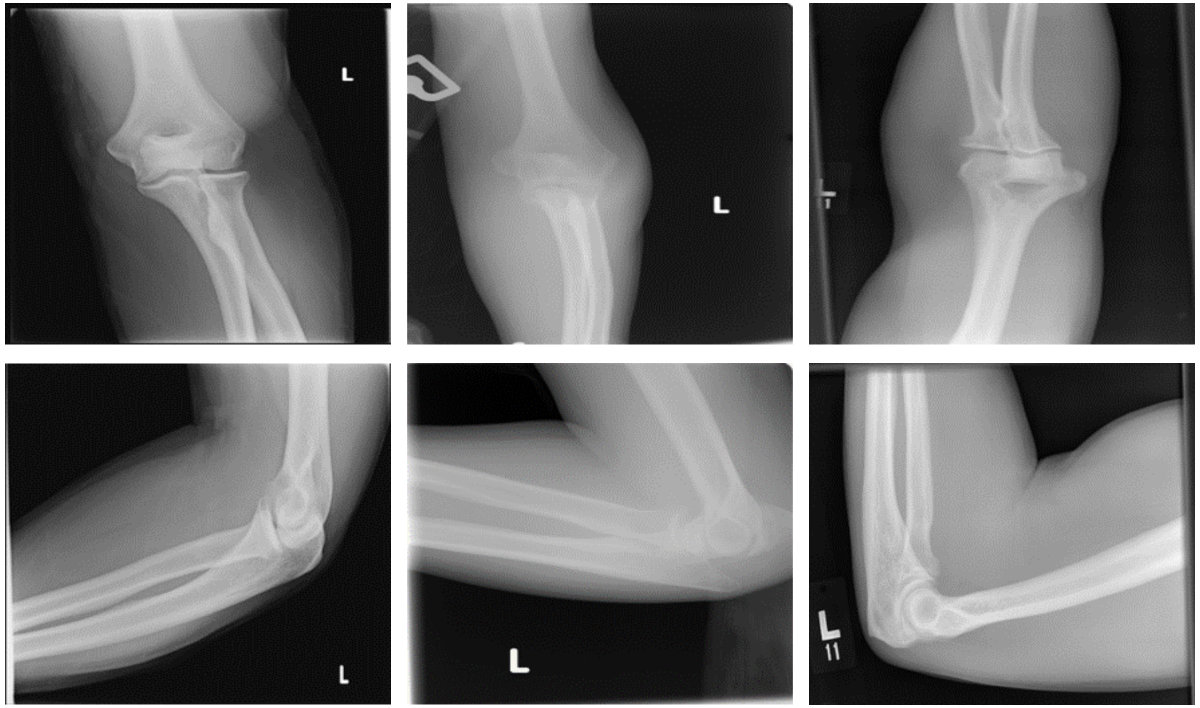
Contribution

- Investigate into a multiview deep learning architecture for elbow fracture diagnosis
- Develop a training strategy via homogeneous transfer learning and curriculum learning that leverages medical knowledge from radiologists.
- Evaluate our method on a three-class classification (Normal vs. Ulnar fracture vs. Radial fracture) task of elbow fracture.

Normal | Ulnar fracture | Radial fracture

Frontal view

Lateral view



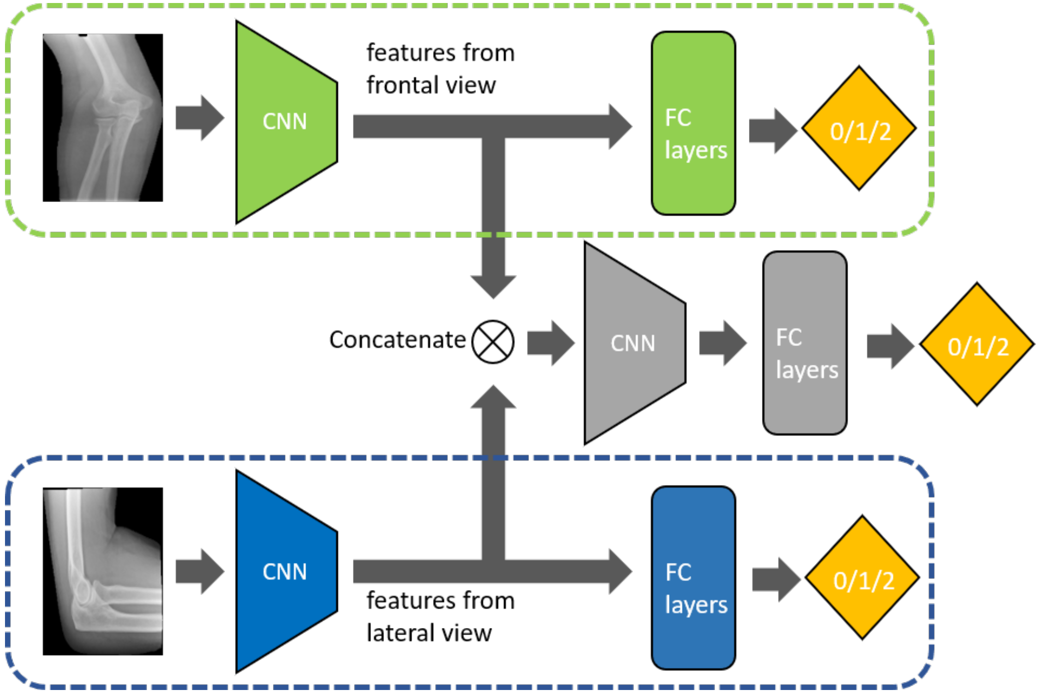
Multiview Model Architecture

- 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:

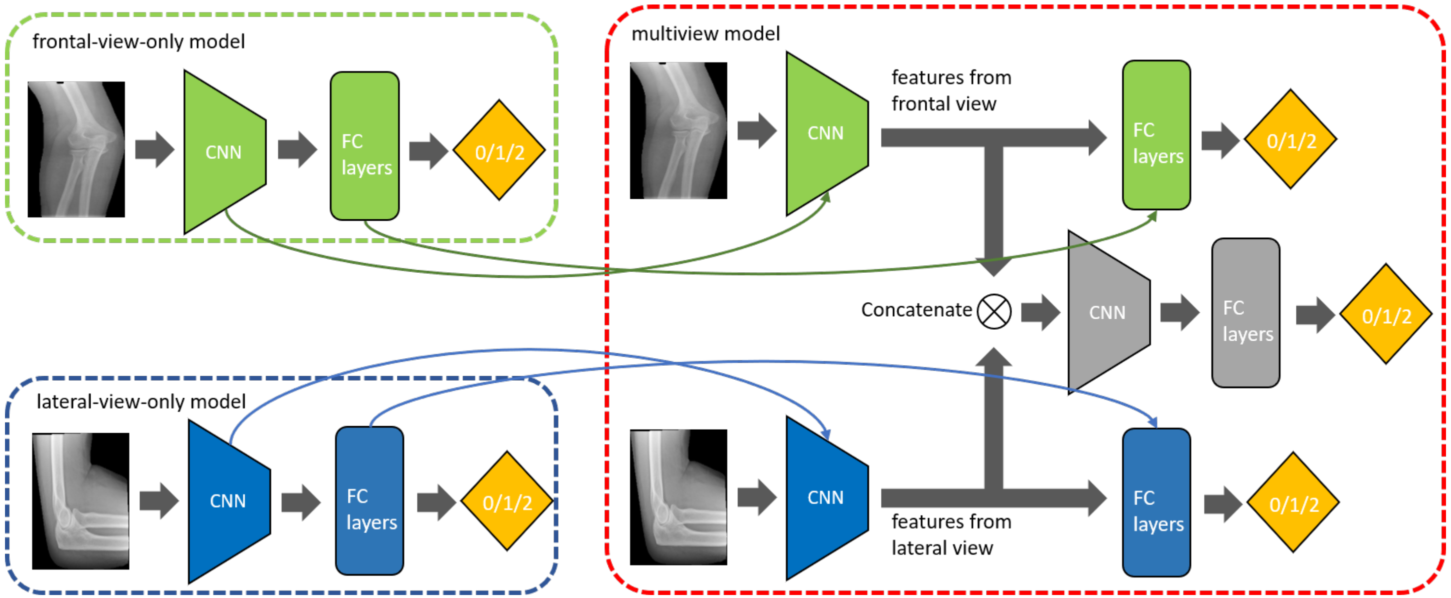
$$J_\theta(\mathcal{D}_i) = J_{\theta_\mathcal{F}}\left(x_i^{(F)}, y_i\right) + J_{\theta_\mathcal{L}}\left(x_i^{(L)}, y_i\right) + J_{\theta_\mathcal{M}}\left(x_i^{(F)}, x_i^{(L)}, y_i\right)$$

- 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
- Scores given by radiologist, 1=hardest; 100=easiest

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

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

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
  - AUC
  - Balanced accuracy (mean of # true positive / # samples of each class)
  - 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 combinations of proposed learning strategies

Model
Single-view
Single-view + CL [2]
Multiview
Multiview + TL
Multiview + [1]
Multiview + [1] + TL
Multiview + CL
Multiview + CL + TL

TL= proposed transfer learning strategy
CL= proposed curriculum learning strategy

Experiments (cont'd)

- Dual-view input experiment results

Model	Accuracy	AUC	Balanced accuracy	Binary task accuracy	Binary task AUC
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

- with dual-view input, our method achieves the highest AUC and balanced accuracy with a margin of up to 0.118

TL= proposed transfer learning strategy
CL= proposed curriculum learning strategy

Experiments (cont'd)

- Single-view input experiment results
  - Frontal view as the only input
- Lateral view as the only input

Model	Accuracy	AUC	Balanced accuracy	Binary task accuracy	Binary task AUC	Accuracy	AUC	Balanced accuracy	Binary task accuracy	Binary task AUC
Single-view	0.720	0.828	0.593	0.761	0.844	0.856	0.954	0.807	**0.895**	0.959
Single-view + CL [2]	0.683	0.807	0.570	0.732	0.813	0.840	0.946	0.809	0.872	0.948
Multiview	0.658	0.749	0.514	0.702	0.766	0.844	0.951	0.800	0.870	0.956
Multiview + TL	0.738	0.827	0.617	0.774	0.829	0.848	0.954	0.804	0.876	0.961
Multiview + [1]	0.566	0.675	0.396	0.575	0.648	0.837	0.945	0.779	0.870	0.949
Multiview + [1] + TL	0.737	0.815	0.605	0.773	0.831	**0.857**	**0.960**	**0.819**	0.885	**0.969**
Multiview + CL	0.723	0.814	0.602	0.761	0.823	0.838	0.956	0.807	0.867	0.956
Multiview + CL + TL	**0.756**	**0.829**	**0.636**	**0.786**	**0.846**	0.840	0.955	0.794	0.874	0.960

- 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.

TL= proposed transfer learning strategy
CL= proposed curriculum learning strategy

Conclusion

- We proposed a novel method that leverages multiview information and incorporates qualitative medical knowledge into deep curriculum learning for elbow fracture detection.
- Our method demonstrated superior performance with both dual-view (frontal and lateral) input and single-view input.
- Clinical knowledge-guided deep learning is an important direction for future AI research in medical applications.

Acknowledgements

Intelligent Computing for Clinical Imaging (ICCI) Lab, University of Pittsburgh

- NIH/NCI R01 (#CA193603)
- NIH/NCI R01 (#EB032896)
- NIH/NCI R01 (#CA218405)
- NSF (OAC #2115082)
- PHDA / UPMC Enterprise
- Amazon Machine Learning Award
- RSNA Research Scholar Grant (#RSCH1530)
- Pittsburgh Center for AI Innovation in Medical Imaging (Pitt Momentum Funds)
- UPCI-IPM Pilot Award (#MR2014-77613)
- Pitt CTSI Biomedical Modeling Pilot Award

Thank you!

Questions?

jul117@pitt.edu