Application of Computer Vision in Intracranial Hemorrhage (ICH) Detection

Amit Bhatia | amit0911@stanford.edu | YouTube video link: https://youtu.be/UZVD9WLmiZ4

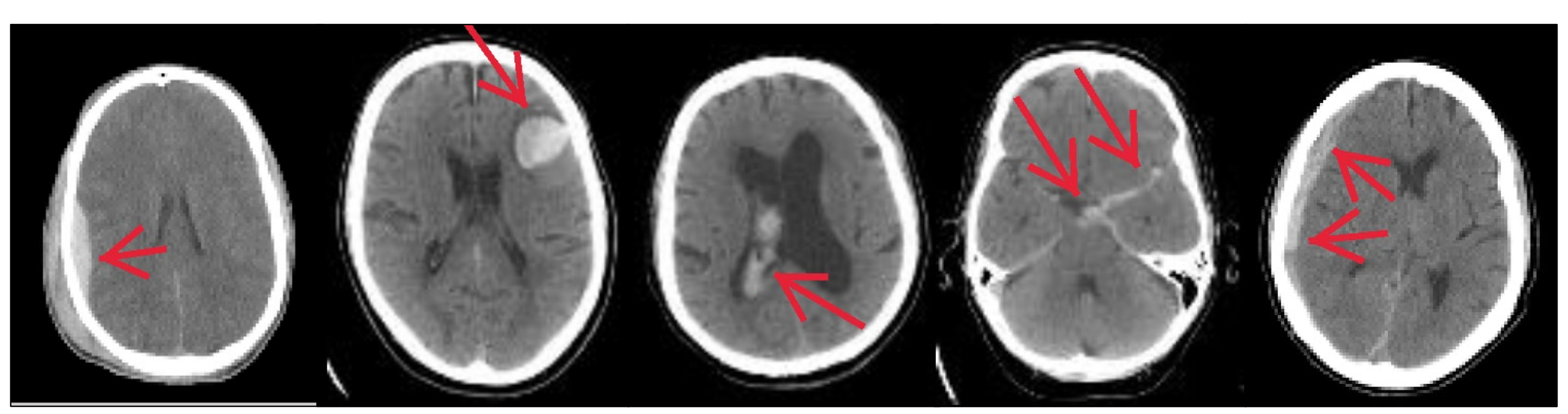
1. Motivation

- Intracranial hemorrhage (ICH), bleeding that occurs inside the cranium, is a serious health problem
- Computed tomography (CT) is the most commonly used medical imaging technique
- Visual inspection by radiologists & manual quantitative estimation of the type & size of hematoma
- Procedure is time-consuming & requires the availability of trained radiologists at every moment
- Automated hemorrhage detection tools capable of providing robust inference can be life saving

2. Data

- Dataset of CT scan (DICOM) images provided by Radiological Society of North America for the Kaggle competition - RSNA Intracranial Hemorrhage Detection
- Labelled training dataset has 753K images with each image having following attributes:
  - Raw pixel array (512, 512) of Hounsfield Units (HU)
  - Meta-data related to Windows (centre, width, rescale intercept / slope), CT Volumes (Study / Sequence IDs, Position, Orientation), etc.
- Y vector of dimension 6 corresponding to the following labels: 1. Epidural (ED), 2. Intraparenchymal (IP), 3. Intraventricular (IV), 4. Subarachnoid (SA), 5. Subdural (SD), and 6. Any (of the 5 sub-types)
- An image can have 0 – 5 ICH sub-type labels

	ED	IP	IV	SA	SD	Any
#	3,145	36,118	26,205	35,675	47,166	107,933
%	0.42%	4.80%	3.48%	4.74%	6.27%	14.34%



3. Approach

- Literature review: CV & medical image analysis
- Used a random sample of 100K images for rapid prototyping – ensured similar class distribution
- Explored multiple approaches along:
  - Pre-processing: Raw HU values, Linear & Sigmoid Windowed images (Brain, Sub-dural & Bone)
  - Network architecture: 2-D ConvNet, 3-D Convnet, 2-D ConvNet + Bidirectional LSTM
  - Transfer learning: ConvNets pre-trained on Imagenet
  - Training strategy: Weighted/Unweighted Binary cross-entropy, Focal loss
- Used weighted multi-label logarithmic loss as the evaluation measure
- Implementation done using Keras. Training done on single Tesla P100 GPU

4. Results

Data	Pre-proc.	Network	Opt. Loss	Train	Test
SampA	Raw HU	7 *(Conv-BN-Relu-MaxPool) + 2*Dense	Wtd BCE	1.744	1.722
SampA	Raw HU	7 *(Conv-BN-Relu-MaxPool) + 2*Dense	Un-Wtd BCE	1.147	1.131
SampA	Raw HU	Pre-tr IncResNetV2 w/o Top + 2*Dense	Un-Wtd BCE	1.198	1.200
SampA	Raw HU	Pre-tr EfficientNetB7 w/o Top + 2 Dense	Un-Wtd BCE	1.213	1.305
SampA	Raw HU	Pre-tr DenseNet w/o Top + 2 Dense	Un-Wtd BCE	1.058	1.082
SampA	Raw HU	Pre-tr DenseNet w/o Top + 2 Dense	Focal Loss	1.250	1.270
SampA	Lin. BSB	Pre-tr DenseNet w/o Top + 2 Dense	Un-Wtd BCE	1.264	1.240
SampA	Sig. BSB	Pre-tr DenseNet w/o Top + 2 Dense	Un-Wtd BCE	0.942	1.127
SampA	3D adj slcs	Pre-tr DenseNet w/o Top + 2 Dense	Un-Wtd BCE	0.995	1.088
SampB	Raw HU	Pre-tr DenseNet w/o Top + 4 Dense	Un-Wtd BCE	1.175	1.476
SampB	Raw HU	Pre-tr DenseNet w/o Top + Bi-LSTM	Un-Wtd BCE	1.084	1.534
SampB	Raw HU	Pre-tr DenseNet w/o Top + Bi-LSTM+Attn	Un-Wtd BCE	1.308	1.571
FullD	**Raw HU**	**Pre-tr DenseNet w/o Top + 4 Dense**	**Un-Wtd BCE**	**0.743**	**0.874**
FullD	3D adj slcs	Pre-tr DenseNet w/o Top + 4 Dense	Un-Wtd BCE	0.553	0.886

	Any ICH		ED		IP	
	Train	Test	Train	Test	Train	Test
AUC	0.9843	**0.9329**	0.9581	0.9004	0.9678	0.9237
P@85pR	61.87%	**51.37%**	0.42%	0.40%	23.33%	19.99%
ACC@85pR	90.93%	**86.41%**	0.42%	0.40%	84.66%	82.85%
R@85pR	95.66%	**84.16%**	100.00%	100.00%	95.88%	86.20%

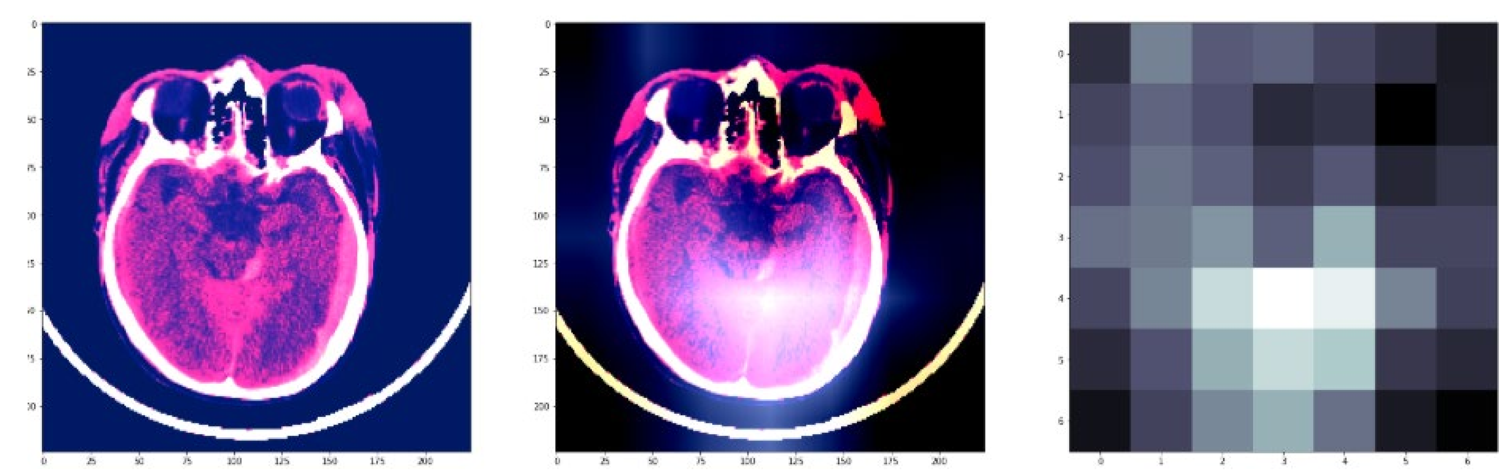
	IV		SA		SD	
	Train	Test	Train	Test	Train	Test
AUC	0.9882	0.9567	0.9625	0.9053	0.9735	0.9238
P@85pR	37.32%	31.96%	21.34%	17.61%	26.67%	23.09%
ACC@85pR	94.28%	93.44%	82.92%	80.71%	83.22%	81.05%
R@85pR	95.27%	81.12%	96.35%	84.69%	96.62%	86.86%

5. Discussion

- Efficacy of Transfer learning (based on DenseNet [1])
- Simplicity outshined - Raw HUs, 2-D ConvNet, Un-weighted BCEL
- No incremental gains – Windowing, Hybrid 3-D ConvNet, CNN + Bidirectional LSTM, Focal Loss
- Test AUC of 0.9329 for ICH detection. Recall of 84.2% with Precision of 51.4%. Accuracy of 86.4%
- Scope for improvement in ICH sub-type detection. Precision ranges from 18-32% for Recall of 80-85%
- Results are sub-optimal for epidural ICH which is a rare class with event rate of 0.42%

6. Visualization

- Used Class Activation Maps [2] to highlight regions in the image that are influencing the network's decision



7. Future

- Multi-GPU training on PyTorch / MXNet
- Deeper transfer learning (fine tuning Conv layers / entire architecture)
- Hierarchical decision system: 2 class (Detector) + 5 class (Labeler) with Focal loss for class epidural

8. References

[1] G. Huang, Z. Liu, K. Q. Weinberger, and L. Maaten. Densely connected convolutional networks. In CVPR, 2017. 2, 6
[2] B. Zhou, A. Khosla, L. A., A. Oliva, and A. Torralba. Learning Deep Features for Discriminative Localization. In CVPR, 2016. 2, 3, 4, 5