nature medicine

SUPPLEMENTARY INFORMATION
https://doi.org/10.1038/s41591-018-0147-y

In the format provided by the authors and unedited.

Automated deep-neural-network surveillance of cranial images for acute neurologic events

Joseph J. Titano[1,5], Marcus Badgeley[2,5], Javin Schefflein[1], Margaret Pain[2], Andres Su[1], Michael Cai[1], Nathaniel Swinburne[1], John Zech[1], Jun Kim[3], Joshua Bederson[2], J. Mocco[2], Burton Drayer[1], Joseph Lehar[4], Samuel Cho[2,3], Anthony Costa[2] and Eric K. Oermann[2]*

[1]Department of Radiology, Icahn School of Medicine, New York, NY, USA. [2]Department of Neurological Surgery, Icahn School of Medicine, New York, NY, USA. [3]Department of Orthopedic Surgery, Icahn School of Medicine, New York, NY, USA. [4]Bioengineering and Bioinformatics, Boston University, Boston, MA, USA. [5]These authors contributed equally: Joseph J. Titano, Marcus Badgeley. *e-mail: eric.oermann@mountsinai.org

Supplemental Figures

Timepoints Recorded

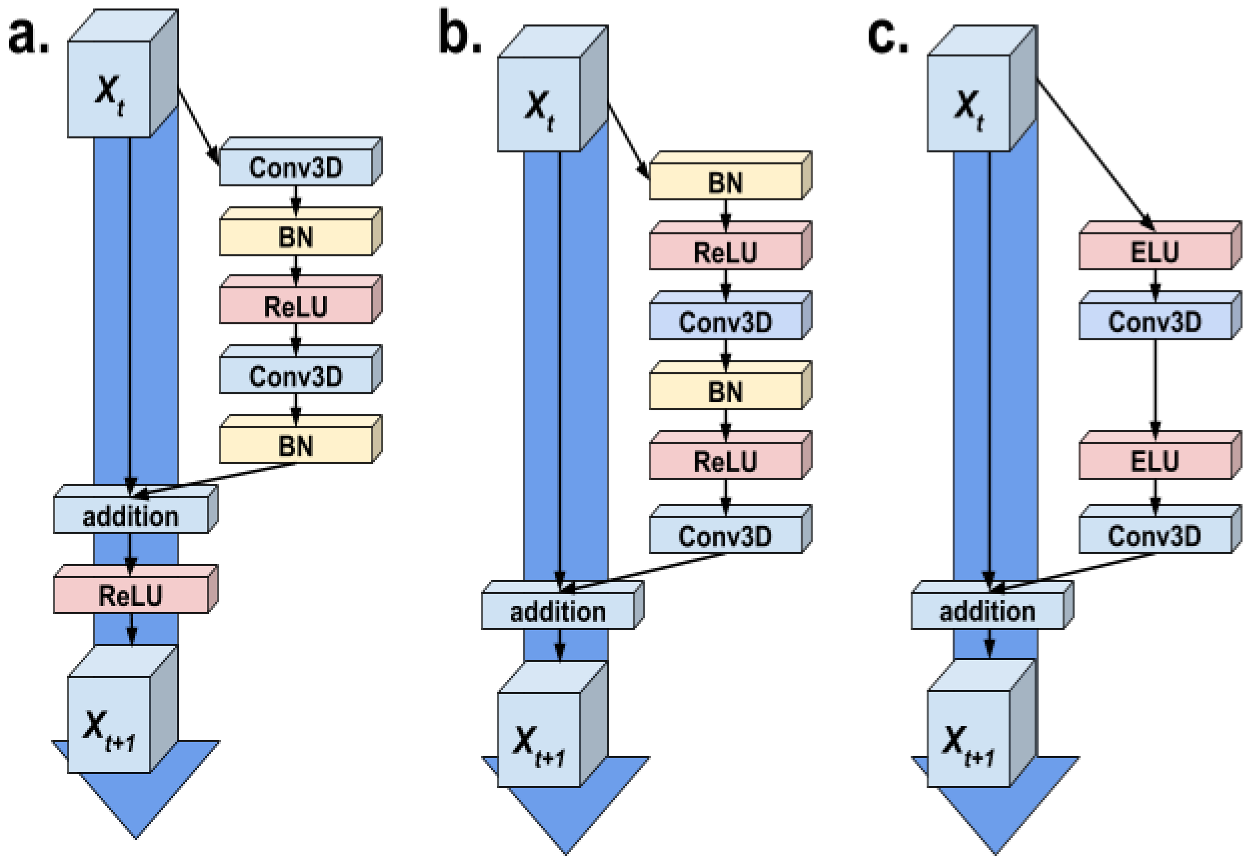
1) Opening study
2) Notify a clinician (if necessary)
3) Finishing a dictation

Workflow Instructions

1) Ensure at least 10 studies are enqueued on PACS before starting.
2) Dictate in accordance with usual practice; use the addendum function of the dictation system without permanently recording your dictation.
3) Only access prior studies/reports that came before the images of interest.
4) Set timer to 30 min.
5) While recording timestamps, reproduce the time displayed on the times. Do not spend time performing any mathematics to obtain the time since you first began. That is, a time that is recorded three minutes into the simulation should be recorded as 27 minutes.
6) Do not begin dictation until the timer has started.
7) Do not stop the timer each time you finish a case.
8) This is simulating real experience, therefore accuracy matters - speed much less so (but not irrelevant).
9) Use clinical history only from sidebar as this represents the history provided to the radiologist at the original time of interpretation.
10) Keep in mind - as in all imaging study interpretation satisfaction of search is a real challenge here.

Supplemental figure S1. Prospective Time Trial Instructions

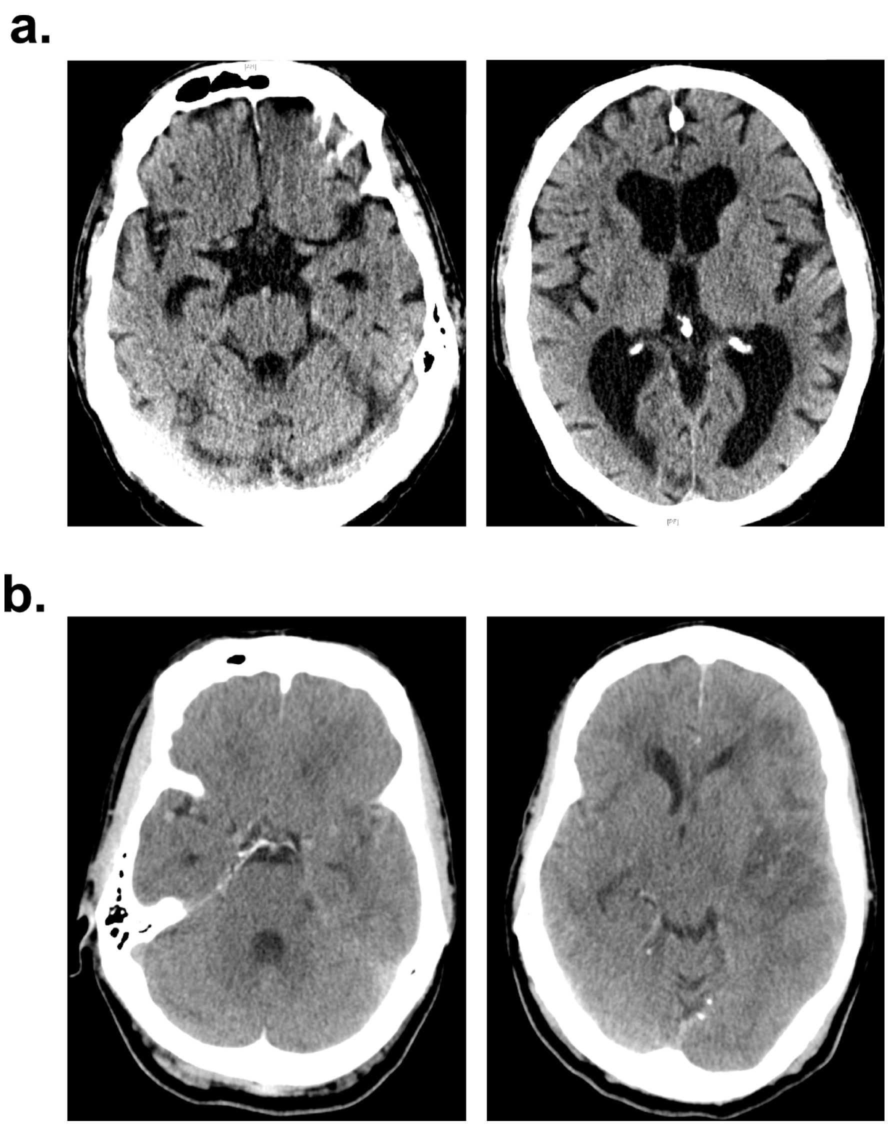
To try and assess the actual impact of performing imaging triage and enqueuing utilizing a deep learning based system we performed a simulated prospective trial of a radiology workflow with and without the use of our system. The above handout was provided to radiologists prior to initiating the trial which was conducted as described above.



Supplemental figure S2. Three Ways to Build a Volumetric Residual Module

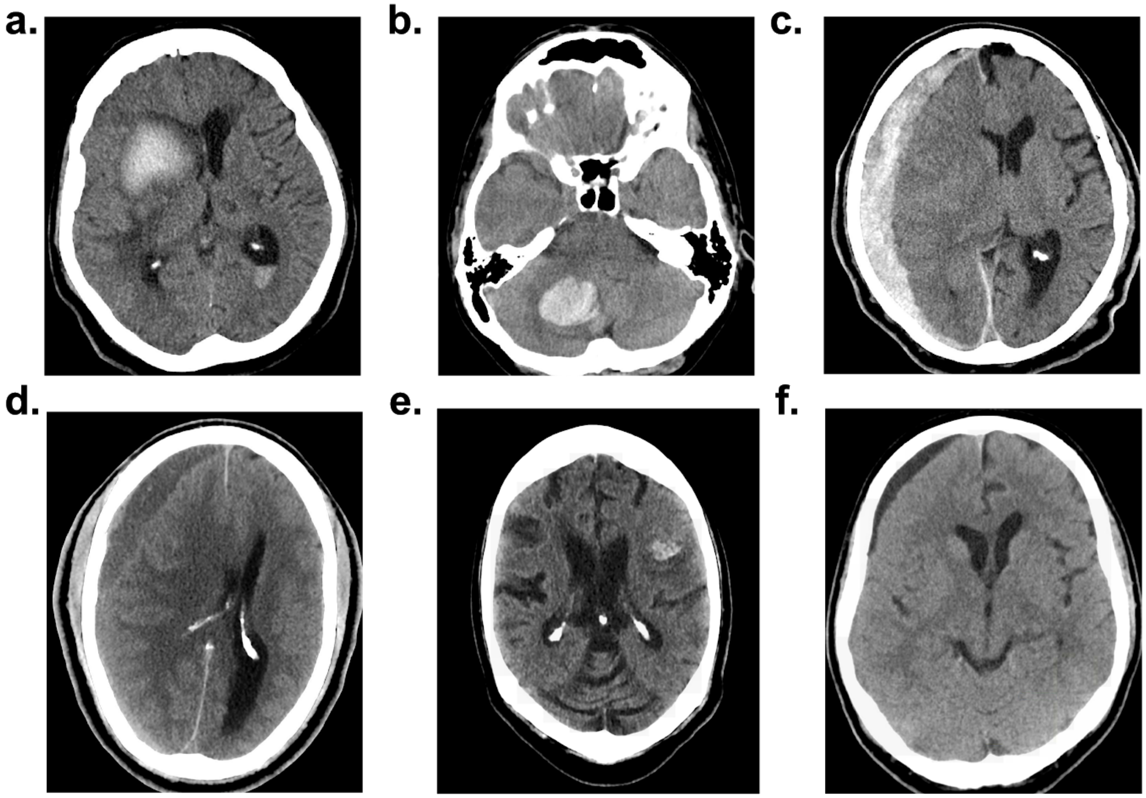
We experimented with three different ways of constructing volumetric residual modules for VolResNet-50. Models were trained in triplicate without an appreciable impact on accuracy. **(a)** Starting in the manner of the original work by He et al. (He, Zhang, Ren, & Sun, 2015), we proceeded to follow the development course established by the original ResNet-50 authors by reshuffling the ordering of operations within the module. **(b-c)** As a further attempt to streamline the model, we also experimented with eliminating the batch normalization (BN) and rectified linear units (ReLU) in favor of the exponential linear unit (ELU). The ELU unit utilizes an exponential function that may improve performance without suffering from the internal covariate shift that can occur with ReLU without batch normalization.

He, K., Zhang, X., Ren, S., & Research, J. S. M. (n.d.). Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015/papers/He_Delving_Deep_into_ICCV_2015_paper.pdf



Supplemental figure S3. Two Examples of Failed Cases

An advantage of focusing on triage rather than diagnosis is that the model is imperfect when it comes to diagnostic performance. **(a)** A notable false positive for hydrocephalus is demonstrated here. The patient was an 82 year old male presenting with ataxia. While his CT demonstrates clear ventriculomegaly, the imaging study also shows significant volume loss and a prior study from a year ago demonstrating similar findings. The patient was later diagnosed with pleuritis, and improved after treatment with azithromycin. While this patient clearly does have ventriculomegaly, which is successfully captured by the CNN, in the context of the significant volume loss (even without the benefit of the prior studies or clinical history) it is unlikely to be hydrocephalus. **(b)** A notable false negative is a case of acute ischemic stroke. The patient, a 61 year old male inpatient, acutely became lethargic and aphasic. A head CT demonstrated a large left middle cerebral artery infarction. Despite that, the algorithm failed to recognize this specific case. Although it is somewhat speculative, acute stroke *is* difficult to identify on non-contrast head CT and the parenchymal changes here and ventricular effacement are not especially pronounced given the large territory that is involved in the infarct. Furthermore, it is worth noting, although this fell under the margin for classification, it just barely did so, and in a queuing system would still be prioritized over other “negative” scans.



Supplemental figure S4. Diversity of Pathology Appearance on Computed Tomography: Subtle and Salient Findings While the appearance of hemorrhage subtypes have consistent, reproducible “classic” descriptions, the entirety of hemorrhage appearances is quite expansive. Representative images are provided above. **(a)** Acute Right Basal Ganglia Intraparenchymal Hemorrhage. Intraparenchymal hemorrhage occurs within the brain tissue itself and may occur in any region of the brain. Also, note the presence of layering left-sided intraventricular hemorrhage. **(b)** Acute Right Cerebellum Intraparenchymal Hemorrhage. **(c)** Acute Right Subdural Hemorrhage. Given its location between the brain tissue and the overlying dura, subdural hemorrhage takes on a “crescent” shape in most cases. **(d)** Subacute Right Frontal Subdural Hemorrhage. Depending upon the time elapsed from the onset of hemorrhage to its discovery with CT, the density of the hemorrhage – as measured by Hounsfield units – may decrease to the point of appearing similar to adjacent brain tissue (as shown here) or perhaps to a lower density than that of brain tissue (not displayed in this image). In some cases, the pathologies detected by the CNN can be subtle. **(e)** Small, Acute Left Frontal Hemorrhage. **(f)** Chronic Right Frontal Convexity Subdural Hematoma.

Supplemental Tables

Supplemental Table S1: UMLS Diagnosis-Criticality Abstraction.

The attached taxonomy was utilized for mapping specific Unified Medical Language System Concept Unique Identifiers (UMLS CUI) to more general clinical categories and, ultimately, to our target labels of critical or non-critical.

See attached file

Supplemental Table S2: Critical Case Frequency in Train and Test Datasets.

The frequency of critical cases within the training and validation datasets identified utilizing *silver standard labels* from the NLP tool varied depending upon the tool's sensitivity and were compared against a ground truth frequency in the test dataset.

Partition	**NLP_4**	**NLP_8**	**NLP_10**	**Gold labels**
Train-val (N=37,326)	7197 (7.56%)	812 (0.85%)	696 (0.73%)	--
Test (N=180)	90 (50%)	2 (1.11%)	2 (1.11%)	56 (31%)

Supplemental Table S3: Diagnoses in Test Cases.

The patients in our test dataset had a variety of clinical diagnoses that resulted in their receiving a head CT. Many were acute neurologic events, however many patients were ultimately diagnosed with *non-neurologic* diseases.

DIAGNOSIS	**N**	**%**
Trauma	35	19.44%
Hydrocephalus	15	8.33%
Idiopathic	13	7.22%
Post-surgical	11	6.11%
Syncope	10	5.56%
Hemorrhage	6	3.33%
Stroke - Subacute	6	3.33%
Tumor	6	3.33%
Substance abuse	5	2.78%
Adverse drug reaction	4	2.22%
Epilepsy	4	2.22%
Sepsis	4	2.22%
SDH - Subacute	3	1.67%
Stroke - Acute	3	1.67%
Stroke - Chronic	3	1.67%
IVH	3	1.67%
Migraine	3	1.67%
Research	3	1.67%
UTI	3	1.67%
Psychiatric	3	1.67%

SDH - Chronic	3	1.67%
Sinusitis	3	1.67%
Failure to thrive	2	1.11%
Pneumonia	2	1.11%
Decompensated heart failure	2	1.11%
Calcification	2	1.11%
TIA	2	1.11%
Dementia	2	1.11%
Cardiac Arrest	1	0.56%
Pleuritis	1	0.56%
Foreign Body	1	0.56%
SAH	1	0.56%
COPD Exacerbation with HCAP	1	0.56%
SDH - Subacute + Midline shift	1	0.56%
Cardiogenic	1	0.56%
Viral meningitis	1	0.56%
Diabetes	1	0.56%
COPD Exacerbation	1	0.56%
BPPV	1	0.56%
Pleural Effusion	1	0.56%
UGIB	1	0.56%
Meningocele	1	0.56%
Meningitis	1	0.56%
Retinal artery occlusion	1	0.56%
DVT	1	0.56%
Cervical Myelopathy	1	0.56%

Facial fracture	1	0.56%

Supplemental Table S4: Presenting Symptoms in Test Cases.

For the patients in the test dataset, their presenting symptoms at the clinical encounter resulting in the ordering of a head CT (emergency department, outpatient clinic, or inpatient unit) were recorded.

SYMPTOMS	**N**	**%**
Headache	49	27.22%
AMS	30	16.67%
Surveillance	28	15.56%
Ataxia	11	6.11%
Loss of consciousness	10	5.56%
Cognitive changes	7	3.89%
Focal neurological deficit	7	3.89%
Dizziness	6	3.33%
Seizures	5	2.78%
Intubated, AMS	5	2.78%
Sensory changes	4	2.22%
Fatigue	2	1.11%
Hemiplegia	2	1.11%
Vision loss - permanent	2	1.11%
Vision loss - transient	2	1.11%
Headache, nausea, vomiting	2	1.11%
Vertigo	2	1.11%
Lethargic	1	0.56%
Rhinorrhea	1	0.56%

Headache, ataxia, urinary incontinence	1	0.56%
Vision loss - transient, weakness	1	0.56%
Facial swelling	1	0.56%
Lethargic, aphasia	1	0.56%

Supplemental Table S5: Radiographic Findings in Test Cases.

For the patients in the test dataset, the frequency of documented radiographic findings from the imaging reports were recorded.

RADIOGRAPHIC FINDING/DIAGNOSIS	**N**	**%**
Hydrocephalus	15	8.33
Noncommunicating	7	3.89
Communicating	8	4.44
Collapsed Ventricles	2	1.11
Fracture	5	2.78
Linear	4	2.22
Depressed	1	0.56
Post-operative Skull	27	15
Subdural Collection - Nonhemorrhagic	8	4.44
Isodense	3	1.67
Hypodense	1	0.56
Mixed Density	4	2.22
Subarachnoid Collection	1	0.56
Isodense	1	0.56
Parenchyma Hypodensity - Likely Acute Ischemia	10	5.56
Territorial	7	3.89
Lacunar	3	1.67
Parenchyma Hypodensity - Likely Chronic Stroke	36	20
Territorial	21	11.67
Lacunar	15	8.33
Chronic Small Vessel Ischemia	55	30.56
Mass Lesion	15	8.33
Sinus Opacification	35	19.44
Foreign Body	35	19.44
Therapeutic	34	18.89
Unexpected	1	0.56
Parenchymal Volume Loss	59	32.78
Global	30	16.67
Focal	29	16.11
Mass Effect	22	12.22
Herniation	3	1.67
Hemorrhage	18	10
Subdural	1	0.56
Epidural	0	0
Subarachnoid	2	1.11
Parenchymal	10	5.56

Intraventricular	5	2.78
Soft Tissue Abnormality	16	8.89
Parenchymal Calcification	8	4.44
Mastoid Opacification	4	2.22
Normal	31	17.22