Stroke Classification

UmedMI

Agenda

1. Data Set (Normal-Abnormal)
2. Papers (Normal-Abnormal)
3. DataSet (Hemorrhagic and Ischemic Stroke)
4. Data Set (Other)
5. Papers (Hemorrhagic and Ischemic)

Data Set (Normal-Abnormal)

The Anatomical Tracings of Lesions After Stroke (ATLAS) Dataset.

- ❖ The dataset contained stroke anatomical brain images and manually lesion segmentation.
- ❖ The data was in Nifti format and contained around 229 T1-weighted MRI scans with manual segmentation and metadata in .csv format.
- ❖ Normal patient dataset contained around 400 MRI scans.

Papers (Normal-Abnormal)

Brain Stroke Detection Using Convolutional Neural Network and Deep Learning Models

- **Date**

  2019

- **Used Machine Learning Algorithms**

  LeNet (convolutional neural network) model

❖ The Dataset

- **(ATLAS) Dataset**
- The dataset contained 229 T1-weighted MRI images suffered from stroke. 210 abnormal images and 210 normal images.
- Then labelled normal patient images as **0** and abnormal patient images as **1** for classification.

Fig. 5. .Abnormal Axial Images

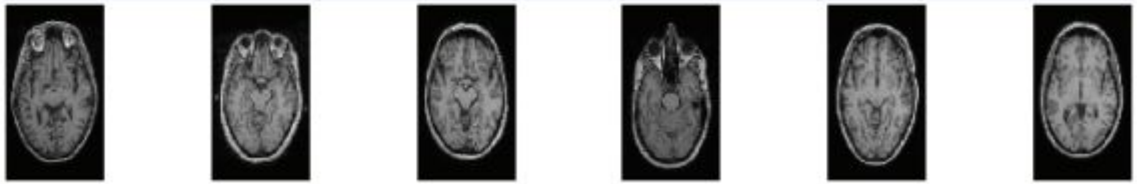


Fig. 6. Normal Axial Images

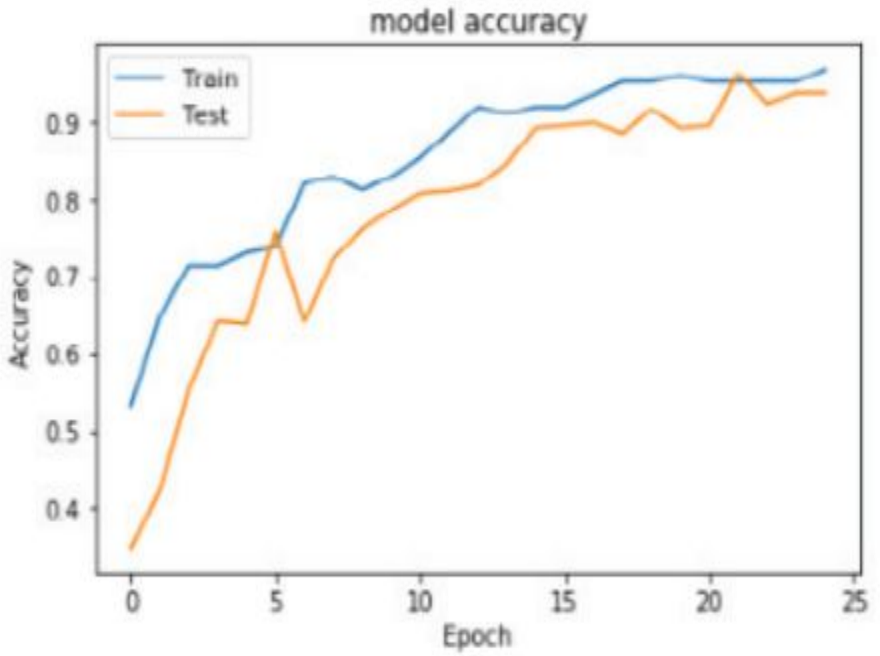
❖ Accuracy

TABLE IV. CONFUSION MATRIX OF PROPOSED METHOD

Testing		Predicted value	
		Abnormal	Normal
Actual value	Abnormal	True Positive (TP)=67	False Positive (FP)=4
	Normal	False negative (FN)=0	True Negative(TN)=51

TABLE V. CLASSIFICATION PERFORMANCE OF PROPOSED METHOD

	Accuracy	Precision	Recall	F1-Score
LeNet	0.9894	0.97	0.97	0.97



❖ Conclusion

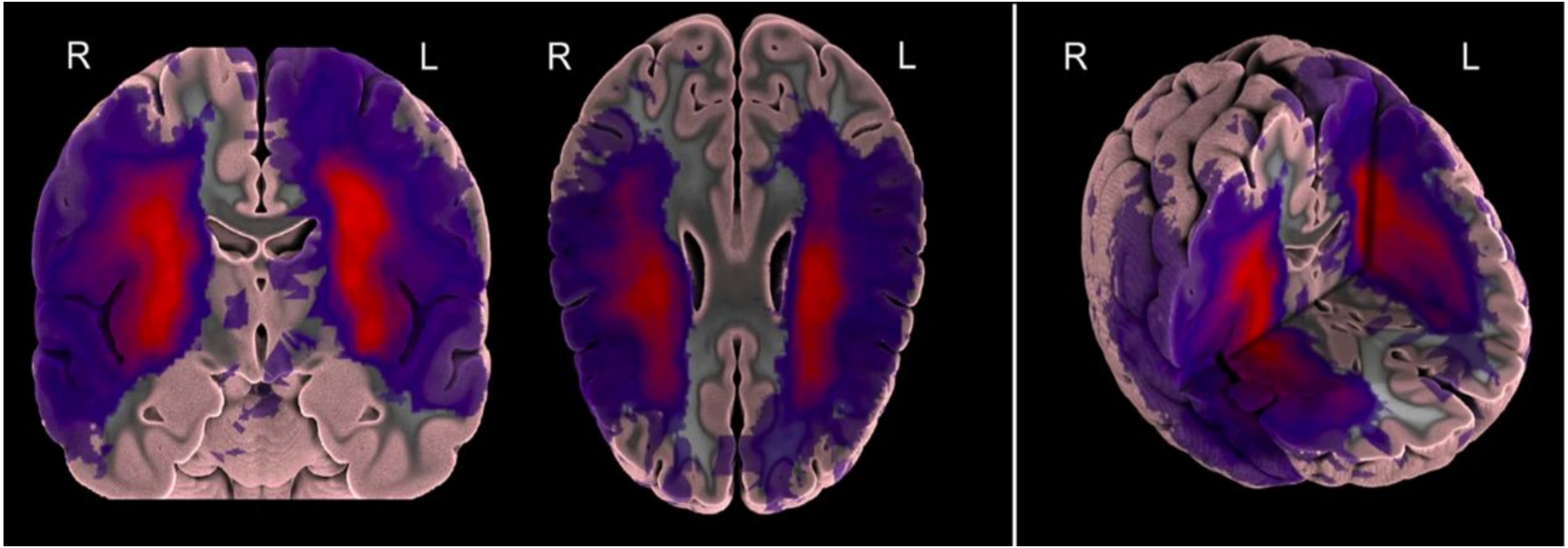
- The research experimental results show that these deep neural networks are absolutely relevant to brain stroke diagnosis.
- Therefore, the accuracy got on proposed method is better than traditional machine learning methods.

DataSet (Hemorrhagic and Ischemic Stroke)

The Anatomical Tracings of Lesions After Stroke (ATLAS) Dataset

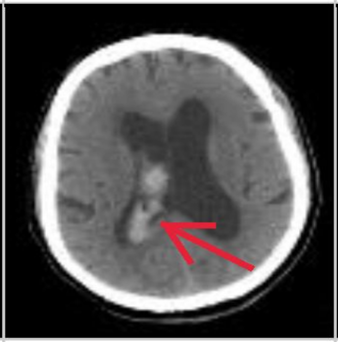
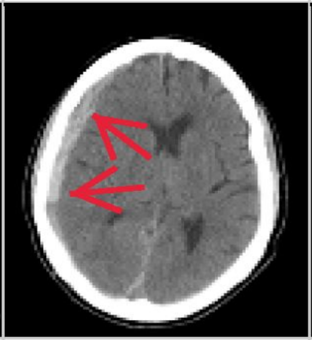
- A large, open source dataset of stroke anatomical brain images and manual lesion segmentations
- is an open-source data collection consisting a total of 304 T1-weighted MRIs (Magnetic Resonance Imaging) with manually segmented diverse lesions and metadata.

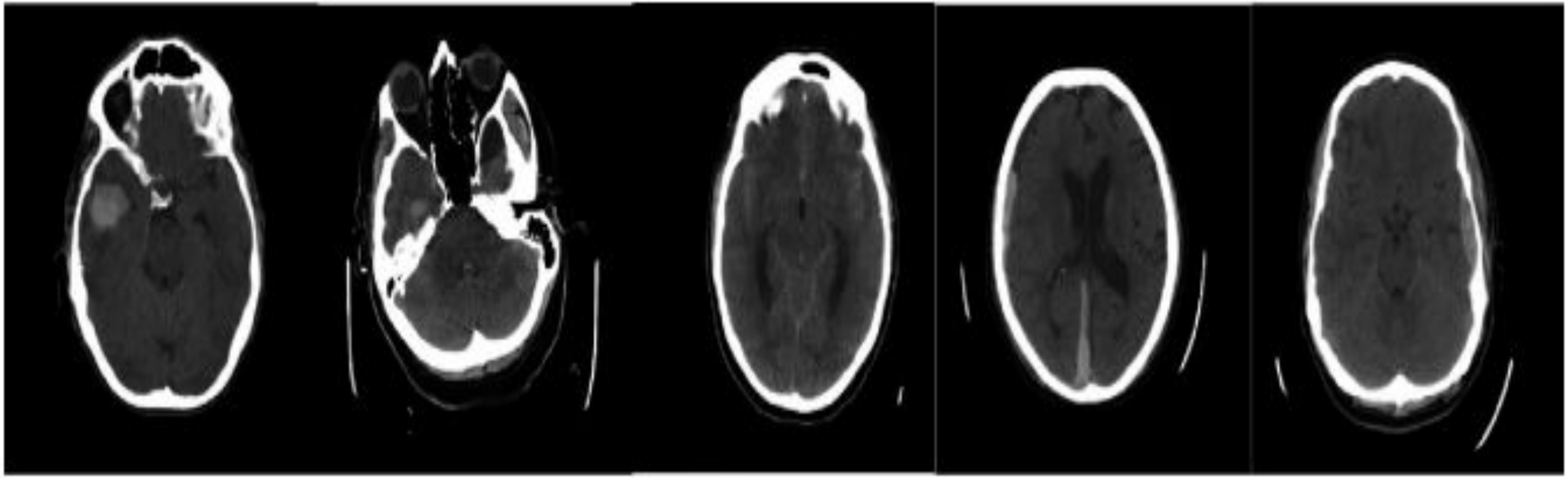
- From 11 cohorts worldwide, 304 MRI images were collected from research groups in the ENIGMA Stroke Recovery Working Group consortium.
- Images consisted of T1-weighted anatomical MRIs of individuals after stroke.



RSNA Intracranial Hemorrhage Detection

- RSNA dataset provided by the Radiological Society of North America (RSNA®) in collaboration with members of the American Society of Neuroradiology and MD.ai.
- All provided images are in DICOM format.
  - DICOM images contain associated metadata.
  - This will include PatientID, StudyInstanceUID, SeriesInstanceUID, and other features.
  - Predict whether a hemorrhage exists in a given image, and what type it is

	Intraparenchymal	Intraventricular	Subarachnoid	Subdural	Epidural
Location	Inside of the brain	Inside of the ventricle	Between the arachnoid and the pia mater	Between the Dura and the arachnoid	Between the dura and the skull
Imaging					
Mechanism	High blood pressure, trauma, arteriovenous malformation, tumor, etc	Can be associated with both intraparenchymal and subarachnoid hemorrhages	Rupture of aneurysms or arteriovenous malformations or trauma	Trauma	Trauma or after surgery
Source	Arterial or venous	Arterial or venous	Predominantly arterial	Venous (bridging veins)	Arterial
Shape	Typically rounded	Conforms to ventricular shape	Tracks along the sulci and fissures	Crescent	Lentiform
Presentation	Acute (sudden onset of headache, nausea, vomiting)	Acute (sudden onset of headache, nausea, vomiting)	Acute (worst headache of life)	May be insidious (worsening headache)	Acute (skull fracture and altered mental status)



intraparenchymal intraventricular subarachnoid subdural epidural

ICH Classes

- ❖ The training data is provided as a set of image Ids and multiple labels:-
  - one for each of five sub types of hemorrhage
  - plus an additional label for any, which should always be true if any of the sub-type labels is true.
- ❖ There is also a target column, Label, indicating the probability of whether that type of hemorrhage exists in the indicated image.
- ❖ There will be 6 rows per image Id. The label indicated by a particular row will look like **[Image Id]_[Sub type Name]**, as follows:

RSNA Intracranial Hemorrhage Detection

ID	Label
1_epidural_hemorrhage	0
1_intraparenchymal_hemorrhage	0
1_intraventricular_hemorrhage	0
1_subarachnoid_hemorrhage	0.6
1_subdural_hemorrhage	0
1_any	0.9

Papers (Hemorrhagic and Ischemic)

Classification of Ischemic Stroke using Machine Learning Algorithms

❖ Date

2016

❖ Used Machine Learning Algorithms

- k- Nearest Neighbor Algorithm
- Decision Trees Algorithm

❖ The Dataset

- The dataset items were collected from several hospitals and medical centers in Sudan.
- The hospitalreport includes the patient number, age, sex, CT, MRI diagnoses, and other variables for all patients hospitalized in the hospitals participated in the study.
- The data used in the dataset include the data of patient of cases from 2013 to 2015.
- The dataset contains 400 patients; their age is mainly between 50 and 88 years. A few cases in the age of 33 years and most of them are male.

Dataset Features

Feature name	The data that the feature contains
A1	The patient number
A2	the age of patient
A3	the sex of patient
A4	if patient have irritability
A5	if patient have convulsions
A6	if patient have left-side weakness
A7	if patient have right-side weakness
A8	patient have mouth deviation
A9	if patient have difficulty in speaking
A10	if patient have unable to walk
A11	if patient have headache
A12	if patient have difficulty in seeing
A13	the result of CT as above
A14	the result of mri as above
A15	the three classes {thrombotic, hemorrhagic embolic}

❖ Accuracy

Decision Trees Algorithm

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	class
0.975	0	1	0.975	0.987	0.995	Thrombotic
1	0	1	1	1	0.992	hemorrhagic
1	0.018	0.958	1	0.979	0.994	embolic
0.987	0.005	0.988	0.987	0.987	0.994	
0.987	0.005	0.988	0.987	0.987	0.994	Weighted Avg

k- Nearest Neighbor Algorithm

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	class
0.975	0.026	0.975	0.975	0.975	0.976	Thrombotic
1	0	1	1	1	0.992	hemorrhagic
0.957	0.018	0.957	1	0.957	0.993	embolic

❖ Conclusion

The results of the experiment revealed that the performance of decision tree classification is better than the performance of KNN algorithm.

Classification of stroke disease using machine learning algorithms

❖ Date

2019

❖ Used Machine Learning Algorithms

- Artificial Neural Network **ANN**
- Decision Trees Algorithm
- Support Vector Machine **SVM**
- Logistic Regression **LR**
- Bagging and boosting

❖ The Dataset

- The data were collected in the form of patient case sheets from Sugam Multi specialty Hospital, India.
- The case sheets contained information from over 507 stroke patients ranging from 35 to 90 years of age.
- A total of 22 unique class labels related to stroke were identified that fell under two major stroke types: ischemic stroke and hemorrhagic stroke.

Dataset Features

Variable name (features)	Extracted feature from the dataset	Variable name (features)	Extracted feature from the dataset
X1	Patient number	X13	Patient with severe headache
X2	Age of the patient	X14	Patient with vomiting
X3	Gender of the patient	X15	Patient with weakness
X4	Patient with numbness	X16	Patient with giddiness
X5	Patient with loss of consciousness	X17	Patient with facial palsy
X6	Patient with diplopia	X18	Patient with nausea
X7	Patient with dysarthria	X19	Patient with aphasia
X8	Patient with difficulty in walking	X20	Patient with altered sensorium
X9	Patient with difficulty in speaking	X21	Patient with hypertension (HT)
X10	Patient with loss of memory	X22	Patient with diabetes mellitus (DM)
X11	Patient with swallowing difficulties	X23	Class of stroke {ischemic (IS), hemorrhage (HE)}
X12	Patient with paralysis		

Sample Dataset

Table 2 Sample dataset

X2	X3	X4	X5	X6	X7	X8	X9	X_{10}	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23
96	M	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	IS
75	M	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	IS
45	F	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	HE

❖ Accuracy

Table 4 Performance measurements for various classification methods

S. no.	Model name	Accuracy	Sensitivity	Specificity	Recall	Precision	Standard deviation
1	Simple tree	90.7	99.1	0	91.4	99.1	38.20
2	Medium tree	89.0	96.5	6.9	91.8	96.5	34.73
3	Complex tree	87.2	94.3	9.3	91.8	94.3	33.14
4	Logistic regression	90.6	99.1	0.0	91.4	99.1	38.19
5	Linear SVM	91.5	1.0	0.0	91.5	1.0	44.50
6	Quadratic SVM	89.5	96.3	11.6	92.1	96.3	32.88
7	Cubic SVM	88.3	95.0	16.2	92.4	95.0	30.68
8	Fine Gaussian SVM	91.1	99.1	4.6	91.8	99.1	36.43
9	Medium Gaussian SVM	91.5	1.0	0.0	91.5	1.0	44.50
10	Coarse Gaussian SVM	91.5	1.0	0.0	91.5	1.0	44.50
11	Ensemble boosted tree	90.9	98.2	11.6	92.3	98.2	33.45
12	Ensemble bagged tree	91.5	99.5	4.6	91.8	99.5	36.55
13	Ensemble RUS boosted tree	63.1	62.9	65.1	95.1	62.9	12.66
14	Artificial neural network	95.3	95.9	60	99.2	95.9	14.69

❖ Conclusion

- The categories of SVM and ensemble (bagged) provided 91% accuracy with 0.0000 negative predictive value.
- ANN trained with the stochastic gradient descent algorithm outperformed other algorithms, with a higher classification accuracy 95% with a lower standard deviation of 14.69.
- This study indicates that stroke is more prevalent in men than in women and in the age group from 40 to 60 years old.

Thank You ツ