


Evolution of Artificial Intelligence in Bone Fracture Detection

Deepti Mishra, G. L. Bajaj Institute of Technology and Management, India*

 <https://orcid.org/0000-0001-5796-6728>

Amit Kumar Mishra, Sharda University, India

ABSTRACT

The objective of the paper is to present the techniques of artificial intelligence based on deep learning that can be applied to detect fractures in bones on x-rays. The paper comprises of discussions of various entities. Initially, there is a discussion on data formulation and processing. Following which, distinguished image processing techniques are presented for fracture detection. Later, there is an analysis of conventional and current neural network methodologies for fracture detection techniques. Furthermore, there is a comparative analysis for the same. Finally, in the end, a discussion is presented in the paper regarding problems and challenges confronted by researchers for fracture detection. The study shows deep learning techniques provide more accuracy in the diagnosis than the conventional methods in fracture detection on x-rays. The paper leads to a path for the researchers to deal with difficulties and issues encountered with the fracture detection on x-rays while using deep learning techniques.

KEYWORDS

Artificial Intelligence, Artificial Neural Network, Bone Fracture, Deep Learning

INTRODUCTION

Artificial Intelligence (AI) plays a crucial role in the domain of medical science such as diagnosis of symptoms and processing of radiological images. AI applies the process of learning that is machine learning, iteratively for improving in response of training procedure, additionally refining outcomes for decision making and predictions. Learning process includes various types (Han, Kamber, & Pei, 2011) (Farooqui & Mehra, 2019). Supervised learning is one of the types where predefined data is trained for applying further predictions on new data (Velchamy, Subramanian, & Vasudevan, 2012). Afterwards, classification technique is executed on afresh entered data for isolating them into groups with the help of trained model and is used to generate decisions (Veni & Rani, 2015).

On the other hand, the techniques of unsupervised learning can be applied to generate decisions without having any trained models (Mishra & Soni, 2014). The process of learning includes various techniques which are used for analysing, computation and implementation of data (Koteeswaran, Visu, & Janet, 2012) (Li, Roy, Khan, Wang, & Bai, 2012). Usually, learning is considered as the

DOI: 10.4018/IJRQEH.299958

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

procedure for training the data set. Like learning, Deep learning can be considered as a combination of techniques based on unsupervised, supervised learning and additionally having networks capability such as artificial neural network. It is an automated computer science technology applied to identify patterns and trends in dissimilar groups, for example, identifying fractures in images of medical field. Deep learning tools plays a crucial role in the domain of image processing, computer visualization and medical domain (Dimililer, 2017).

Concept of artificial intelligence i.e., machine intelligence is reflected in deep learning which is comprised of many layers of networks to extract features to improve the level of perception and precision (CHUNG, et al., Automated detection and classification of the proximal humerus fracture by using deep learning algorithm, 2018). Artificial intelligence requires deep learning for computations and processing. Deep learning is composed of neural networks, which is further comprised of input layers and hidden layers required for processing and calculations (Johari & Singh, 2018). Therefore, deep neural network has now become an empowered technology for medical image diagnosis for enhanced performance and accuracy in diagnosis.

A fracture is a broken bone, which may be a thin crack or major bone battered badly, and still, it further may be a closed or open type (Al-Ayyoub, Hmeidi, & Rababah, 2013). The severity of fracture may vary which depends on external force (Maheshwari & Mhaskar, 2011). Severity also generally depends on the age of person and on the part of Bone. A fracture is a broken bone which is dealt in Orthopaedics branch of medical science. Fracture can be vertical, horizontal, oblique or may be in smaller pieces in bone after contact with some external force or stimulus. Clinical diagnosis of fracture in lack of specialized expertise may steer to consequences, i.e., Failure to diagnose fracture may otherwise lead to grave consequences and poor outcome.

Usually, all type of fractures can be visible and diagnosed in X-Ray. Later, CT and MRI can be done for confirmation of fracture. X-rays is an imaging technique based on electromagnetic radiation to create picture of internal body parts and bones underneath skin. For acquiring the image of body part in X-ray, the patient is positioned accordingly. Computer Tomography is generally known as CT scan, utilizes computer and X-rays machines for creating cross-sectional images and three-dimensional reconstruction images for more accuracy and understanding. MRI also called as magnetic resonance imaging utilizes the paired and unpaired magnetic fields of proton nucleus of hydrogen atoms of water (which is present everywhere in the body) to create images. MRI is especially useful in detecting fractures which have not yet become precipitated or become imminent e.g., Stress Fractures, or where X-rays and CT scan are contraindicated e.g., Pregnancy.

Now a days, machine learning has surfaced as a scientific boom which provides computer-based techniques for medical diagnosis (Farooqui & Mehra, 2019). It incorporates recent advance techniques with resources to provide swiftness in the process of diagnosis with additional accuracy. A missing fracture in X-rays can lead to sever aftereffects for patients consequentially poor recovery.

This paper presents the techniques of Artificial Intelligence that are applied to detect fractures in bones. The paper presents a study to show accuracy in diagnosis of fractures using Artificial Intelligence techniques such as deep learning and artificial neural network. Distinct research works are presented in the paper conforming the guidance of artificial intelligence in medical science to detect fractures by applying its techniques. The paper includes a comparison and discussion among various techniques of artificial intelligence.

Further the next sections of the paper are planned as, section 2 incorporates around the study of bone fractures in human being. Section 3 presents data pre-processing. Further the section 4 in the paper, presents different techniques of artificial intelligence in bone fracture detection. Further the next section 5 presents the comparative study of various techniques followed by discussion. Final section presents comparison and discussions on algorithm following the conclusion and future work.

LITERATURE SURVEY

Bone Fracture

Fractures can be classified in multiple ways. For example, Traumatic Fractures – are those which are caused by trauma viz. Road traffic accidents, assaults, rail traffic accidents, fall etc. etc. Pathological fractures --occur in bones which have been weakened by some underlying disease like bone weakened by cancer spread (bony metastasis) or bone weakened by increased porosity – osteoporosis. Periprosthetic fractures – occur at points of mechanical weakness due continued stress at different parts of implant placed in a bone.

In children and adolescents, the fracture line may be incomplete due to the plastic, less brittle nature of their bones. These incomplete fractures are called greenstick fractures, where one tension cortex fails. If the compression cortex buckles, they are called torus or buckle fractures. Paediatric bone may also simply undergo plastic deformation without a visible fracture line.

The classification methodology can be based on many other points like fractures pattern, displacement of fragments, associated joint dislocation and or subluxation, fragment numbers etc. etc. Below mentioned table beautifully delineates the usual classifications and concomitant comparison of the fractures in between them in different aspects.

Types of fracture may be considered as and are mentioned in figure 1:

1. **Stable and closed fracture:** It is when the broken ends of the bone are ruled up in a line and are hardly out of place.
2. **Open or Compound Fracture:** It is when the bone sticks out of the skin. The bone may or may not be visible in the wound.
3. **Transverse Fracture:** If fracture line is horizontal, then it can be considered as transverse fracture.
4. **Oblique Fracture:** The fracture is visible in different angular arrangement.
5. **Comminuted Fracture:** When the bone breaks into more than two pieces.
6. **Greenstick Fracture:** It can be stated as when bone is not completely broken into pieces. It may happen in the case of children.
7. **Spiral Fracture:** The fracture escalates and spiral down the bone.

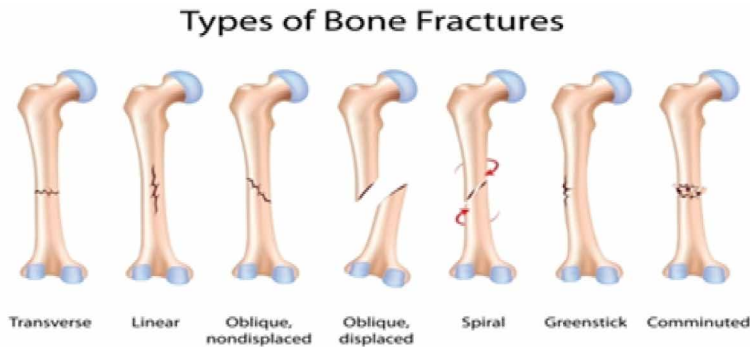
A visible deformity may be recognized by anyone, especially if it is pronounced one, even by naked eyes but the fracture can properly be delineated only on X-rays (Edward, Hepzibah, Wei, & Li-ming, 2017). Fracture description requires name of bone, region of bone, pattern of fracture line, presence of compression, presence of displacement of fracture fragments, type and degree of displacement, any pre-existing pathology and any associated joint pathology like dislocation or subluxation presented in table 1.

Role of AI in Bone Fracture Detection

Artificial Intelligence applying in medical science with respect to X-rays images started as early as in 1960 (Lodwick, Keats, & Dorst, 1963). Initially for computation and analysis, the data was transformed into numerical form. Later AI came into existence with computer aided detection. With time there has been enhancement in computer processing because of development of ability of learning from experience in systems termed as machine learning. The machine learning empowers more enhanced AI systems in various domains. AI models offers more accurate and efficient outcomes on datasets by applying training and learning (Kalmat, et al., 2020) (Rainey, McConnell, Hughes, Bond, & Fadden, 2021). AI has been quite efficient in identifying critical and non-critical cases.

Artificial Intelligence systems trained with machine learning techniques can be applied to detect bone fractures in X-rays. Models enabled with AI can detect bone fractures automatically. Neural

Figure 1. Open fracture illustration (Kumara & Gopura, 2019) and types of bone fracture



network applied on trained data sets provides accurate results on X-ray images. Moreover, AI, specifically deep learning including neural network, empowers medical science domain by providing better visualization on imaging data. Convolutional neural network plays the key role in designing architecture for detection of fractures in bones. It offers both precision and speed while generating outcomes.

DATA PREPROCESSING

Non-medical images can be used for pre-training of the dataset, which further can be used for automated fracture detection on plain radiographs. Dataset of fractured images are preprocessed to improve data features and to conceal undesirable features (Khatik, 2017). It requires good amount of time for pre-processing the data for further learning model. The objective is to prepare data for analysis and computations. Overfitting can cause the problem while retrieving and selecting the data (Hodge & Austin, 2004). Overfitting is the process in which while learning, the data captivates noise also, thus impacting the outcomes negatively. Initially captured dataset is framed for learning process, which itself requires ample amount of time. Thereafter, it needs a labeling process which needs to be overseen properly. The data is further classified as fracture or no fracture. The identification of the fracture is confirmed by expertise surgeon having experience of more than three years. A dataset of radiographs is obtained from a specialized hospital and analyzed under the guidance of expert orthopedic surgeon.

Now a days, there are various visualization tools available for implementation of neural network. R studio provide a good background for implementation of deep neural network. Packages like nnet, neuralnet, deepnet and rnet are available for deep learning in R studio.

DIFFERENT TECHNIQUES OF AI IN BONE FRACTURE DETECTION

Primary Machine Learning Based Algorithm

Functioning of initial systems was based on consideration of sole feature for detection of fracture which was Neck-Shaft Angle (NSA). The value of NSA predicts the fracture or no fracture, such as if value is less than 116° , then fracture is recognized in the image. The single feature is not capable enough to identify fractures accurately generating error rate of approximate 8%. Thus, a new technique is proposed for detecting slight disorders based on classification, by dragging traits in femur X-rays, executing texture analysis of bony trabecular pattern.

Table 1. Comparison of fractures with diagnosis

Type of Fractures	Traits			
	Symptoms	Diagnosis	Consequences if missed	Possibility of correction
Stable and closed	Pain, swelling, slight or no deformity, loss of transmitted movements	X-rays	Angulation, pain,	yes
Open and compound	Pain, open wound, bleeding, deformity, loss of transmitted movements crepitus	X-rays	Malunion, nonunion, infection, osteomyelitis, septic arthritis, Compartment syndrome, Volkmann's ischemia	Less
Transverse	Pain, swelling, deformity, loss of transmitted movements, crepitus	X-rays	Angulation, persistent deformity, stiffness, malunion, nonunion, shortening	less
Oblique	Pain, swelling, deformity, loss of transmitted movements crepitus	X-rays	Angulation, persistent deformity, stiffness, malunion, nonunion, shortening	less
Comminuted	Gross swelling, severe pain, total loss of transmitted movements, gross deformity, crepitus	X-rays	Compartment syndrome, Volkmann's ischemia, neuro vascular injury, mal union, nonunion, Subdeck's osteodystrophy	Less likely
Greenstick	Pain, swelling, deformity, angulation, restricted mobility	X-rays	Angulation, persistent deformity, stiffness, malunion	yes
Spiral	Pain, swelling, deformity, loss of transmitted movements	X-rays	Deformity, angulation	yes

Feature extraction is an integral component, while applying primary machine learning techniques, for fracture detection. Gabor filters, Intensity gradient direction and Markov random field techniques are utilized to extract features in femur X-rays (Lim, et al., 2004) (Lum, Leow, Chen, Howe, & Png, 2005). Some researchers solely used Gabor filter for feature extraction in fracture detection using classifiers (Yap, Chen, Leow, Howe, & Png, 2004). Transformation based on Gradient with homotopic functions are used by some researchers. (Wei, Na, Huisheng, & Fan, 2009). The regression trees and SVM can be applied in various domains such as medical diagnosis and, also in weather prediction. (Ordoñez, Matías, Juez, & García, 2009). Present your perspective on the issues, controversies, problems, etc., as they relate to theme and arguments supporting your position. Compare and contrast with what has been or is currently being done as it relates to the article's specific topic and the main theme of the journal.

Ensemble Based Classification System

Ensemble methods are the combinations of algorithms in a sole predictive paradigm to improve predictions. It benefits in refining the outcomes by associating various models of learning. Hence, ensemble methods are given priority over other techniques. Ensemble based classification system is the collection of various classifiers, to generate a conjoint outcome by combining their individual results (Naz, Zafar, & Khan, 2019). Basically, the ensemble technique has two approaches, first based on different learning methods, unlike the other ones having same learning approach on different training sets. The classifier system has few base classifiers whose outcome is combined for finding a precise one (Jurek, Bi, Wu, & Nugent, 2013).

Ensemble classifiers utilizes selection techniques for choosing optimal subset of classifiers. The approaches for selection are static and dynamic. In static approach, best functioning classifier is chosen for a validation set. Dynamic approach can be performed in two methods- one is dynamic classifier selection and other is dynamic ensemble selection. Dynamic classifier selection in which base classifiers are selected to construct an ensemble. In dynamic ensemble selection, the best combination of ensembles is chosen having high prediction accuracy (AdnanO.M.Abuassba, DezhengZhang, XiongLuo, AhmadShaheryar, & HazratAli, 2017). These ensemble techniques such as bagging, and boosting can be applied in medical science for fracture detection (Vishnu, Prakash, Rengasamy, & Sharmila, 2015).

Researchers have utilized ensemble-based learning approaches for detecting osteoporotic fractures by applying boosting, bagging and random subspaces (Kilic & Hosgormez, 2016). Distinct models are prepared having different set of features which further functions as classifiers. One of the ensemble techniques, 10-fold cross validation is applied to detect fractures in bone (Iliou, Christos-Nikolaos, Anagnostopoulos, & Anastassopoulos, 2014). Initially, researcher uses feature selection before applying machine learning technique. It has the advantage of avoiding needless further testing with bone.

Here, some techniques with their functions on ensemble classifier are discussed.

Bagging

The widespread technique is bagging which partition the dataset in distinct training sets from a sole data set. Initial functioning of different dataset is selected from single dataset k times randomly causing some repeated and common instances.

It offers an advantage that different partitioned data sets which belong to sole one are trained concurrently saving the time, and additionally improving the performance, for example Decision trees. It improves accuracy by conjoining the outcomes of more than one classifier having majority values. It is based on simple average of outcomes of base classifiers. Equation 1 presents the formula for calculation in bagging technique:

$$outcome(O) = 1 / n \sum_{i=1}^n O_i \quad (1)$$

On the other hand, the disadvantage is that the training data sets are not independent.

Algorithm

```

Input:
D: Dataset
C: Number of classifiers
S: Classification Scheme
Output:
Predicted output by Ensemble - A classifier model
Method:
1. For i to C do // create C models
2.   Construct dataset Di from D
3.   Apply Di with Ci and generate model
4. End;
5. Conjoint the outcomes in Ensemble S
6. Generate predicted value

```

Previous Authors have proposed a technique “Wagging (weights aggregation)” {bagging variant} which repeatedly unsettle the training data set instead of applying sampling from it (Bauer & Kohavi, 1999). For weighing process attributes such as Gaussian noise with a given Standard Deviations and zero mean value were added to each weight for computations. Initially, as testing, in all instances, noise was added to the weights of the instances in each iteration process and then one classifier is induced.

The classification technique bagging is applied in detecting fractures (Al-Ayyoub, Hmeidi, & Rababah, 2013). Authors take a set of X-rays after labelling it and applied filters for reducing noise. Further, bagging technique is applied as classification to prepare base classifier.

Boosting

Boosting, is another approach which operates by applying same learning method on different data sets. It is considered as iterative approach. It is based on dynamic division of training sets in addition with classifier accuracy by assigning weights to each instance. To begin with, the base classifiers are prepared and added to ensemble, all the instances are reweighted. The combination of predicted values of multiple weights of base classifiers are considered as concluding prediction. Final prediction is the conjoint of all the predicted weighted values calculated by each base classifier.

Boosting outcomes are based on the calculation of weighted average of results of different base classifiers presented in equation 2:

$$Outcome(O) = \sum_{i=1}^n w^* O_i \quad (2)$$

AdaBoost (Adaptive Boosting) algorithm is one of the kinds of boosting ensemble method.

A technique named MultiBoosting is presented which is expansion of AdaBoost algorithm for decision making (Webb, 2000).

Gradient Boosting machines (GBM) provides more focused predictions which may in turn be very informative and can be applied for medical diagnosis (Atkinson, et al., 2012). Researchers have found that GBM generates robust fracture predictions. Regression techniques are applied before supplying the data to GBM model. Separate model is constructed using R package for predicting fractures.

Stacking

Stacking is the way to combine the outcome of different learning classifiers into a one classifier resulting in less bias or variance error in classifier's prediction. It is a kind of ensemble learning techniques.

During stacking, a problem domain is partitioned for the scope of different learning schemes. Distinct learning models are generated for different partitioned problems. Finally, a predicted outcome is resulted and combined.

Random Forest

It is based on the concept of partitioning the attribute space. It uses the same training data set by applying different feature subspaces. Random forest functions on the large number of different trees to generate ensemble classifiers. It is the method used in classification by building hundreds of decision trees.

Researchers apply Random Forest technique is applied for analysing X-ray images by utilizing feature extraction process (Cao, Wang, Moradi, Prasanna, & Syeda-Mahmood, 2015). The findings are generated after executing experiments on 145 X-rays images. The functioning of proposed methodology utilizes feature extraction process, random forest technique and support vector machine.

Deep Learning-Based Techniques for Fracture Detection

In deep neural network they offer accurate outcomes by combining the predictions of multiple training models (Thian, et al., 2019). They provide better scaling and flexibility in data and offer different sets of weight in each iteration in deep learning thereby generating distinct predictions.

A technique named ADPO (Adaptive Differential Parameter Optimization) is also discussed by the authors which apply artificial neural network in bone fracture detection by classifying and differentiating fractured lines with non-fractured lines (Yang, Cheng, Shimaponda-Nawa, & Zhu, 2019). Algorithm applies probabilistic Hough transform for data pre-processing and artificial neural network for classification for identification of X-rays as fracture and no fractures. Accuracy accomplished was 74.4% additionally with AUC 0.8149.

A deep CNN model is implemented for identifying wrist bone fracture (Lindsey, et al., 2018). The initial dataset consists of 1,35,845 X-rays including images of all body parts, out of which 34,990 X-rays were selected as they were wrist bone X-rays. Two test datasets were created, first test data set having 3500 X-rays which were chosen randomly from 34,990 wrist bone X-rays. The second dataset comprised of 1400 X-rays. The training of the model is done in two steps. Initially it is pretrained and then polished and refined for identifying wrist fractures. The performance evaluation is accomplished by generating AUC curve of approximate 0.967 and 0.975 on different datasets additionally sensitivity 93.9% and specificity 94.5%.

DenseNet which is one of the architectures of neural network, is applied by the researchers for analysing pelvic X-rays functioning on approximate 53,000 images (Gale, Carneiro, Oakden-Rayner, Palmer, & Bradley, 2017). Initially, the labelling process is applied, thereafter a deep neural network of 172 layers having 12 features/units per layer is employed. ROC curve is generated for the model offering accuracy of 95% to over 99%.

The approach of transfer learning, which is a kind of deep convolutional neural network, is utilized for identification of fractures in X-rays images (Kim & MacKinnon, 2018). Initially, the tool trained 11,112 images identifying 695 fracture and 694 no fracture images. The performance evaluation process includes by generating AUC curve approximate 0.954 by calculating sensitivity 0.9 and specificity 0.88.

Artificial neural network approach is also capable to detect proximal humerus fractures in anteroposterior X-rays (CHUNG, et al., Automated detection and classification of the proximal humerus fracture by using deep learning algorithm, 2018). The dataset of 1891 images are trained

using convolutional neural network. Data pre-processing is done on images by cropping them manually in squares and resizing them in 256 X 256 pixels. Thereafter, data is distributed in 10 subsets such as subset is used as data set and rest for training the model and a deep CNN model is applied. The model shows 95% accuracy with AUC 0.996 additionally sensitivity 0.99 and specificity 0.97.

Deep learning is applied in efficient manner for bone fracture detection (Dimililer, 2017). The objective of researchers is to categorize X-rays images as fractured bone or unfractured bone applying neural network including 1024 neurons on X-rays of various body parts. The functioning of methodology was based on 30 trained samples and finally executed on 70 sets. Two basic pre-processing techniques are applied named Haar wavelet transform and SIFT. SIFT is scale-invariant feature transform which is comprised of some steps such as feature extraction and orientation. The methodology applying deep learning in form of neural network has accuracy approximately 94%.

Researchers apply deep neural network to identify hip fractures in X-rays (Cheng, et al., 2019). Deep convolutional neural network is trained by various datasets, initial is pretrained dataset of 25,505 limb X-rays including ankles, elbows, feet, and wrists. Another dataset includes frontal pelvic dataset of 3,605 X-rays images. Furthermore, a dataset of 100 pelvic X-rays obtained for a second test. Pre-processing is applied to dataset as all images are converted to 512X512 pixels. Densenet is used to model the system for fracture identification (Huang, Liu, Maaten, & Weinberger, 2018). There are various ranges of accuracy and AUC depending on the dataset. Initially, in pretraining accuracy is 99.5%. But after applying directly to hip fracture images accuracy was 90% and AUC 0.98. Additionally, mean accuracy generated after a questionnaire with 4 orthopaedic surgeons and two radiologists, the sensitivity 98% and specificity 87.7%.

COMPARISON AND ANALYSIS

In this section we have tried for a comprehensive comparative study of few methodologies of computer science for bone fracture detection. A comparison shows that usual pre-existing or primary methods for bone fracture usually depends on sole feature detection (Cao, Wang, Moradi, Prasanna, & Syeda-Mahmood, 2015). Primary techniques applied for bone fracture detection does not provide a precise outcome relative to detection related to deep learning methods. The motivation of study in table 2, indicates the empowerment of deep learning implies in artificial neural network offers a worthy outcome in bone fracture detection in comparison with primary learning techniques.

Ensemble techniques such as bagging and boosting are also applied in bone fracture detection whereas both having a similarity that is, the process of generating training data sets for the classifiers.

The table 3 represents the study done by various authors in deep learning, which is part of artificial intelligence, shows, it can be applied for identifying fractures in bones. While observing comparative findings, it is clearly visible that deep learning provides precise results in detection of fracture in bones. The study is done on various research works applied in the field of bone fracture detection employing artificial intelligence. The comparative table shows that artificial neural network offers a robust environment in orthopaedics. The data presented in table having worthy AUC value with sound specificity and sensitivity percentage.

The data presents the summary of study done in domain of artificial intelligence while showing that how its wings can be spread in every subject.

PERFORMANCE ASSESMENT

Till now in the current paper, data pre-processing, different learning techniques have been discussed so far. In this section, a glimpse of performance evaluation is demonstrated in corresponding with deep learning in bone fracture detection. Usually, these mentioned techniques apply the concept of probability and statistics.

Table 2. Comparison of automated techniques for fracture detection

Approach		Properties				
		Methods	Outcome	No. of feature at a time	Computation	Advantage / Disadvantage
Primary Techniques		Based on consideration of single feature	% Error is high	one	Not complex	Outcome is not precise
Ensemble techniques	Bagging	Based on training the base learners on sets of databases.	More precise than Primary	multiple	Complex	Better accurate results.
	Boosting	Apply weight to data sets for learning. It applies weighted	Better than Bagging	multiple	Complex	More clear and accurate results
Deep Learning		Learned from large sets of data and then applied for predictions	Precise	multiple	Expensive computation	Most sensitive, specific, and precise outcomes.

Table 3. A comparative study of deep learning by various authors for fracture detection in bone discussed in section

Article	Type of fracture/ X-rays	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
A.Y Yang (Yang, Cheng, Shimaponda-Nawa, & Zhu, 2019)	Long Bone fractures	0.849	74.4	-	-
Lindsey (Lindsey, et al., 2018)	Wrist Bone fractures	0.967	-	93.9	94.5
Gale (Gale, Carneiro, Oakden-Rayner, Palmer, & Bradley, 2017)	Pelvic fractures	0.994	95-99	95	
D.Kim (Kim & MacKinnon, 2018)	Wrist X-rays	0.954	-	90	88
S.W. Chung (CHUNG, et al., Automated detection and classification of the proximal humerus fracture by using deep learning algorithm, 2018)	Proximal Humerus Fractures	0.996	95	99	97
K. Dimililer (Dimililer, 2017)	Various body Parts X-rays	-	94	-	-
C.T. Cheng (Cheng, et al., 2019)	Hip Fractures	0.98	99.5 (90 when applied to hip fractures)	98	84

Confusion Matrix

It is kind of performance calculation for learning problems in which result may be differentiated in two or more classes. Rows represent the predicted values while column represents the actual values in the matrix as shown below in table 4.

Now, from the table it can be concluded for fracture identification as TP for True positive fracture, TN for true negative fracture, FP for false positive and FN for false negative. FN can also be stated as observation for fracture is positive but predicted as negative. Like as FN, FP can be stated as fracture is not detected but it is predicted as positive. FP and FN can be identified as Type 1 error and Type 2 error.

Values obtained from confusion matrix are classification accuracy, recall and precision mentioned in equation 3, 4 and 5:

$$\text{Accuracy } A = \frac{TN + TP}{TN + TP + FN + TN} \quad (3)$$

$$\text{Recall } R = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Precision } P = \frac{TP}{TP + FP} \quad (5)$$

Chi-Square Test

It is statistical method which can be applied in learning to verify the correlation among two variables. This can be applied to ascertain that selected features are appropriate for predicted outcome. Initially chi square scores are calculated for selected features and thereafter a rank is generated for them. And then top ranked features are chosen for training the model. Equation 6 presents chi-square formula:

$$X^2 = \sum_{i=0}^n \frac{(O_i - E_i)^2}{E_i} \quad (6)$$

where O_i is observed value and E_i is expected value.

Feature selection plays a crucial role in bone fracture detection (Bandyopadhyay, Biswas, & Bhattacharya, 2014). The decision models and classifiers improve and can provide accuracy in the outcomes if their calculations are applied on best chosen features.

Table 4. For confusion matrix

		Actual Values	
Predicted Values		Positive	Negative
	Positive	TP (True Fracture)	FP (False predicted fracture)
	Negative	FN (Fracture Not predicted)	TN (Fracture not detected)

It is non-parametric test to test whether data belongs to fracture or no-fracture group and fits according to expectation. Chi-square test is applied to compare between fracture and no-fracture group.

AUC and ROC Curve

Receiver operating characteristic curve (ROC) and Area under ROC curve (AUC) both are used for performance evaluation for classification problems of classifiers based on threshold settings (Tanzi, Vezzetti, Moreno, & Moos, 2020). ROC is probability curve and AUC are based on degree and measurement of separability. It usually represents the capability of model so that how much it is robust to distinguish between classes.

Both can be used to distinguish between two classes that is fracture and no fracture in the bone. The values on X axis and Y axis are retrieved from the values of confusion matrix. X axis measures FPR (False positive rate) and Y axis concludes TPR (True Positive Rate) further which are calculated from confusion matrix mentioned above. Both formulae presented in equation 7 and 8:

$$FPR = \frac{FP}{TN + FP} \quad (7)$$

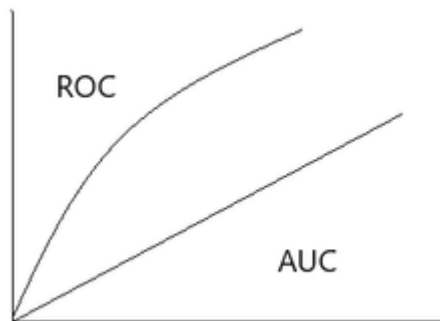
$$TPR = \frac{TP}{TP + FN} \quad (8)$$

ROC curve provides a comparison and diagnostic potential for fracture detection for the given data samples. Figure 2 represents the visualization of ROC and AUC curve.

DISCUSSION

From our previous knowledge we all know that the adult human body comprises of 206 bones of various shapes, size, and structures. Bone fracture is a common problem in humans. With increasing urbanization, road traffic accidents top the lists of fractures causation, followed by fall and many more. Fractures can occur in almost all and every bone of our body. From skull to toe every bone can be fractured, although the incidence and causation mechanism varies. Fractures can be classified in multiple ways.

Figure 2. ROC curve



Limb deformity secondary to gross bony fracture can be appreciated by the naked eyes and even a layman can assume or presume about the presence of a fracture. However, the fracture or its type can't be seen by the naked eyes and therefore either of an imaging modality is required to detect the fracture, namely X-ray or CT scan (Computed Tomography), Magnetic Resonance Imaging (MRI), and sometimes Ultrasound even. Of all these X-rays being the most used because of their widespread availability followed by the CT scan. Clinicians utilize X-rays films to see the presence and absence of fractures, location, and their types as well injuries to the joints in vicinity. But sometimes these images lack enough details needed to diagnose the fracture. Many times, the fracture may be a hairline one, difficult to detect or interpret. Sometimes fracture is not clear in x-rays. Least, it may be even missed in a casual, cursory view if it is a hidden one.

Pseudocode for bone fracture detection applying artificial intelligence:

1. Patient registered
2. Entry in the dataset
3. Analysis and computation on the entered dataset
4. Examination of X-ray
5. Matching the Outcomes
6. Analysis of results

This paper presents a study of artificial intelligence techniques for detection of bone fractures. Index paper provides study of various methods for detection of bony fractures using artificial intelligence and aids in designing new techniques of fracture detection with improved accuracies. This paper presents a detailed discussion of many methodologies utilizing neural network-based systems for fracture detection with their merits and demerits. The study in the paper shows the multiple ways to represent accuracy of deep learning methodology to detect bone fractures. This methodology also has ability to decrease error while diagnosing fractures in X-ray. The study reflects that deep learning techniques offers more robust system contrasted to clinical or to those X-rays seems to be difficult to physicians. The study reveals about the systems approaching deep learning for X-rays which need be followed through different phases. Phases can be stated as pre-processing, analysis, and decision generation. The phases are also designated through pseudocode. During pre-processing stage X-rays are automatically furnished for visualization. After pre-processing, the X-rays are passed through neural network system to generate outcomes. In the last decision making is done while applying post processing and analyses on the outcome.

Applying Artificial Intelligence model is also a challenging task while detecting fractures in bones. Few significant challenges are 1. Detecting various types of fractures having different shapes and sizes 2. Finding best neural network architecture for fracture detection. These challenges can be resolved in various steps- 1. First detecting presence of fracture whether it is present or not. 2. In next step to trace the fracture at exact location. 3. Apply AI as a science to design architecture to provide accurate outcome. Convolutional neural network offers precise results.

Deep neural network can detect various kinds of bone fractures occurring in different body parts. Deep learning functions efficiently on fracture datasets and offers accurate results.

CONCLUSION AND FUTURE WORK

Artificial neural network now a days is an innovative thought in medical science. It can be considered as boom in medical science. The current paper presents a deep discourse of artificial neural network for bone fracture detection. A survey is discussed in the paper regarding application of neural network in detection of fracture in bones. As per the comparison, it has been found that Deep learning approaches

performed better when compared with other techniques that are applied for fracture detection. The paper leads to a path for the researchers to deal with difficulties and issues encountered with the fracture detection on X-rays while using deep learning techniques.

The modern tools in vogue nowadays can precisely unearth information about human body in a feasible and cost-effective way. The new emerging discoveries provided via both hardware and software demands new techniques and refinements of pre-existing techniques to be invented. It can be easily assumed that there is no universal recipe or cookbook method that can be made to investigate or manage every human body part and different techniques are meant for each body part individually.

FUNDING AGENCY

Publisher has waived the Open Access publishing fee.

REFERENCES

- Abuassba, Zhang, Luo, Shaheryar, & Ali. (2017). Improving Classification Performance through an Advanced Ensemble Based Heterogeneous Extreme Learning Machines. *Computational Intelligence and Neuroscience*, 1-11. 10.1155/2017/3405463
- Al-Ayyoub, M., Hmeidi, I., & Rababah, H. (2013). Detecting Hand Bone Fractures in X-Ray Images. *Journal of Multimedia Processing and Technologies*, 4, 155–168.
- Atkinson, E. J., Therneau, T. M., Melton, L. J. III, Camp, J. J., Achenbach, S. J., Amin, S., & Khosla, S. (2012). Assessing Fracture Risk using Gradient Boosting Machine (GBM) Models. *Journal of Bone and Mineral Research*, 27(6), 1397–1404. doi:10.1002/jbmr.1577 PMID:22367889
- Bandyopadhyay, O., Biswas, A., & Bhattacharya, B. B. (2014). Long-Bone Fracture Detection in Digital X-ray Images Based on Concavity Index. In *International Workshop on Combinatorial Image Analysis IWCIA 2014* (pp. 212-223). Springer. doi:10.1007/978-3-319-07148-0_19
- Bauer, E., & Kohavi, R. (1999). An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants. *Machine Learning*, 36(1/2), 105–139. doi:10.1023/A:1007515423169
- Cao, Y., Wang, H., Moradi, M., Prasanna, P., & Syeda-Mahmood, T. F. (2015). Fracture detection in x-ray images through stacked random forests feature fusion. In *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE. doi:10.1109/ISBI.2015.7163993
- Cheng, C.-T., Ho, T.-Y., Lee, T.-Y., Chang, C.-C., Chou, C.-C., Chen, C.-C., Chung, I.-F., & Liao, C.-H. (2019). Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. *European Radiology*, 29(10), 5469–5477. doi:10.1007/s00330-019-06167-y PMID:30937588
- Chung, S. W., Han, S. S., Lee, J. W., Oh, K.-S., Kim, N. R., Yoon, J. P., Kim, J. Y., Moon, S. H., Kwon, J., Lee, H.-J., Noh, Y.-M., & Kim, Y. (2018). Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthopaedica*, 89(4), 468–473. doi:10.1080/17453674.2018.1453714 PMID:29577791
- Dimililer, K. (2017). IBFDS: Intelligent bone fracture detection system. *9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017*, 120, 260-267. doi:10.1016/j.procs.2017.11.237
- Edward, C. P., Hepzibah, H., Wei, Z., & Li-ming, Z. (2017). *A Survey for Fracture Detection in Hand Bone Xray Images*. Academic Press.
- Farooqui, N. A., & Mehra, R. (2019). Design of A Data Warehouse for Medical Information System Using Data Mining Techniques. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*. IEEE. doi:10.1109/PDGC.2018.8745864
- Gale, W., Carneiro, G., Oakden-Rayner, L., Palmer, L. J., & Bradley, A. P. (2017). *Detecting hip fractures with radiologist-level performance using deep neural networks*. Computer Vision and Pattern Recognition.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining Concepts and Techniques*. Elsevier.
- Hodge, V. J., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22(2), 85–126. doi:10.1023/B:AIRE.0000045502.10941.a9
- Huang, G., Liu, Z., Maaten, L. v., & Weinberger, K. Q. (2018). Densely Connected Convolutional Networks. *Computer Vision and Pattern Recognition*. arXiv:1608.06993.
- Iliou, T., Anagnostopoulos, C.-N., & Anastassopoulos, G. (2014). Osteoporosis Detection Using Machine Learning Techniques and Feature Selection. *International Journal of Artificial Intelligence Tools*, 23(5), 1450014. Advance online publication. doi:10.1142/S0218213014500146
- Johari, N., & Singh, N. (2018). Bone Fracture Detection Using Edge Detection Technique. In *Soft Computing: Theories and Applications* (pp. 11-19). Springer. doi:10.1007/978-981-10-5699-4_2
- Jurek, A., Bi, Y., Wu, S., & Nugent, C. (2013). A survey of commonly used ensemble-based classification techniques. *The Knowledge Engineering Review*, 1–31. doi:10.1017/S0269888913000155

Kalmet, P. H., Sanduleanu, S., Primakov, S., Wu, G., Jochems, A., Refae, T., Ibrahim, A., Hulst, L., Lambin, P., & Poeze, M. (2020). Deep learning in fracture detection: A narrative review. *Acta Orthopaedica*, 91(2), 215–220. doi:10.1080/17453674.2019.1711323 PMID:31928116

Khatik, I. (2017). A Study of Various Bone Fracture Detection Techniques. *International Journal Of Engineering And Computer Science*, 6(5), 21418–21423. doi:10.18535/ijecs/v6i5.38

Kilic, N., & Hosgormez, E. (2016). Automatic Estimation of Osteoporotic Fracture Cases by Using Ensemble Learning Approaches. *Journal of Medical Systems*, 40(3), 61. Advance online publication. doi:10.1007/s10916-015-0413-1 PMID:26660692

Kim, D., & MacKinnon, T. (2018). Artiðcial intelligence in fracture detection: Transfer learning from deep convolutional neural networks. *Clinical Radiology*, 73(5), 439–445. doi:10.1016/j.crad.2017.11.015 PMID:29269036

Koteeswaran, S., Visu, P., & Janet, J. (2012). A review on clustering and outlier analysis techniques in data mining. *American Journal of Applied Sciences*, 254–258.

Kumara, E., & Gopura, R. C. (2019). *Development of an External Bone Fracture Fixator with adjustable and self-locking pins*. <https://www.researchgate.net/publication/335380795>

Li, J., Roy, P., Khan, S. U., Wang, L., & Bai, Y. (2012). Data Mining using Clouds: An Experimental Implementation of Apriori over Mapreduce. *12th international conference on scalable computing and communications*.

Lim, S. E., Xing, Y., Chen, Y., Leow, W. K., Howe, T. S., & Png, M. A. (2004). Detection of femur and radius fractures in x-ray images. *Proc. 2nd Int. Conf. on Advances in Medical Signal and Information Processing*.

Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M., Sicular, S., . . . Potter, H. (2018). Deep neural network improves fracture detection by clinicians. *Proc. Natl. Acad. Sci.*, 11591–11596. doi:10.1073/pnas.1806905115

Lodwick, G. S., Keats, T. E., & Dorst, J. P. (1963). The Coding of Roentgen Images for Computer Analysis as Applied to Lung Cancer. *Radiology*, 81(2), 185–200. doi:10.1148/81.2.185 PMID:14053755

Lum, V. L., Leow, W. K., Chen, Y., Howe, T. S., & Png, M. A. (2005). Combining classifiers for bone fracture detection in X-ray images. In *IEEE International Conference on Image Processing*. IEEE. doi:10.1109/ICIP.2005.1529959

Maheshwari, J., & Mhaskar, V. A. (2011). *Essential Orthopaedics*. Jaypee. doi:10.5005/jp/books/12250

Mishra, D., & Soni, D. (2014). A comprehensive Overview on Data Mining Approaches and Applications. *International Journal of Computer Science and Information Technologies*, 7814-7816.

Naz, M., Zafar, K., & Khan, A. (2019). Ensemble Based Classification of Sentiments Using Forest Optimization Algorithm. *DATA, MDPI*, 76(4), 76. Advance online publication. doi:10.3390/data4020076

Ordoñez, C., Matías, J. M., Juez, F. J., & García, P. J. (2009). Machine learning techniques applied to the determination of osteoporosis incidence in post-menopausal women. *Mathematical and Computer Modelling*, 50(5-6), 673–679. doi:10.1016/j.mcm.2008.12.024

Rainey, C., McConnell, J., Hughes, C., Bond, R., & Fadden, S. (2021). *Artificial intelligence for diagnosis of fractures on plain radiographs: A scoping review of current literature*. Intelligence Based Machine. doi:10.1016/j.ibmed.2021.100033

Tanzi, L., Vezzetti, E., Moreno, R., & Moos, S. (2020). X-Ray Bone Fracture Classification Using Deep Learning: A Baseline for Designing a Reliable Approach. *Applied Sciences (Basel, Switzerland)*, 10(1507), 1–17. doi:10.3390/app10041507

Thian, Y. L., Li, Y., Jagmohan, P., Sia, D., Chan, V. E., & Tan, R. T. (2019). Convolutional Neural Networks for Automated Fracture Detection and Localization on Wrist Radiographs. *Radiology. Artificial Intelligence*, 1(1), e180001. Advance online publication. doi:10.1148/ryai.2019180001 PMID:33937780

Velchamy, I., Subramanian, R., & Vasudevan, V. (2012). A five step procedure for outlier Analysis in Data Mining. *European Journal of Scientific Research*.

- Veni, C. K., & Rani, T. S. (2015). Ensemble based classification using small training sets: A novel approach. In *IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*. Orlando, FL: IEEE. doi:10.1109/CIEL.2014.7015738
- Vishnu, V. A., Prakash, J. D., Rengasamy, S., & Sharmila, T. S. (2015). Detection and classification of long bone fractures. *2nd International Conference on Innovations in Information, Embedded and Communication systems (ICIIECS)*, 10, 18315-18320.
- Webb, G. I. (2000). MultiBoosting: A Technique for Combining Boosting and Wagging. *Machine Learning*, 40(2), 159–196. doi:10.1023/A:1007659514849
- Wei, Z., Na, M., Huisheng, S., & Fan, H. (2009). Feature Extraction of X-ray Fracture Image and Fracture Classification. *International Conference on Artificial Intelligence and Computational Intelligence*. doi:10.1109/AICI.2009.40
- Yang, A. Y., Cheng, L., Shimaponda-Nawa, M., & Zhu, H.-Y. (2019). Long-Bone Fracture Detection using Artificial Neural Networks based on Line Features of X-ray Images. In *IEEE Symposium Series on Computational Intelligence (SSCI)*. Xiamen, China, China: IEEE. doi:10.1109/SSCI44817.2019.9002664
- Yap, D. W.-H., Chen, Y., Leow, W. K., Howe, T. S., & Png, M. A. (2004). Detecting femur fractures by texture analysis of trabeculae. In *Proceedings of the 17th International Conference on Pattern Recognition*. Cambridge, UK: IEEE. doi:10.1109/ICPR.2004.1334632

Deepti Mishra, PhD, Associate professor, G L Bajaj Institute of Technology & Management, Greater Noida, India. She completed her PhD from Noida International University, Greater Noida, India. During her PhD she worked on Outlier detection in data sets applying clustering techniques. Her research area includes data mining, machine learning, artificial intelligence. She has participated at many professional conferences and workshops. She has published research articles and book chapters in national and international publications.

Amit Kumar Mishra who has did his MBBS from S N Medical College, Agra, one of the oldest medical schools in Indian subcontinent in 2000 with distinctions in 5 subjects. He completed his post graduate diploma in orthopaedic surgery from J N Medical College, Aligarh Muslim University, Aligarh, again a legendary medical school of unparalleled reputation, in 2004, securing 1st position. He has vast experience of more than 16 years in orthopaedics. He has been an avid teacher, and has been a jubilant person participating and encouraging researches. He has been affiliated with Sharda hospital, school of medical sciences and research, Greater Noida for more than 16 years.