Comparative study of Image Classification Algorithms

Utkarsh Pandey[1] | Himanshu Aneja[2]| Deepanshu Jindal[3] | Ajay Tiwari[4]

Department of IT & Maharaja Agrasen Institute Of Technology, Delhi, India

To Cite this Article
Utkarsh Pandey, Himanshu Aneja, Deepanshu Jindal and Ajay Tiwari, "Comparative study of Image Classification Algorithms", *International Journal for Modern Trends in Science and Technology*, Vol. 07, Issue 01, January 2021, pp.-88-92.

ABSTRACT

This investigation analyzed five common machine learning techniques for performing image classification included Support Vector Machines (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), Binary Decision Tree (BDT) and Discriminant Analysis (DA). AlexNet deep learning model was utilized to fabricate these machine learning classifiers. The structure classifiers were executed and assessed by standard execution models of Accuracy (ACC), Precision (P), Sensitivity (S), Specificity (Spe) and Area Under the ROC Curve (AUC). The five strategies were assessed utilizing 2608 histopathological pictures for head and neck cancer. The examination was directed utilizing multiple times 10-overlay cross validation. For every strategy, the pre-trained AlexNet network was utilized to separate highlights from the activation layer. The outcomes outlined that, there was no contrast between the consequences of SVM and KNN. Both have the equivalent and the higher accuracy than others were 99.98 %, though 99.81%, 97.32% and 93.68% for DA, BDT and NB, separately. The current examination shows that the SVM, KNN and DA are the best techniques for classifying our dataset images.

KEYWORDS*: Machine learning methods, AlexNet, ConvNet*

I. INTRODUCTION

Lately, a few exploration focus on the utilization of deep learning strategy for medical image Computer-aided diagnosis (CAD) techniques that utilization deep neural organizations to learn examples of the picture dependent on a huge preparing informational index. Deep learning has been accounted for to essentially outflank old style AI strategies for object grouping and has been progressively utilized for clinical picture investigation. Up until this point, the utilization of deep learning for clinical pictures include the detection and segmentation of lesions from histopathology images ([5]).

The deep learning technique was the convolutional neural organization (CNN), which is a deep neural network committed for picture classification, it is likewise named the ConvNets in certain literary works ([5]). CNN has more advantages than traditional machine learning strategies; the advantages are:

- CNN doesn't accept hand-made highlights as info; it wiped out the requirements for tumor division and highlight determination, making

the entire cycle considerably more helpful and less inclined to client inclination [5].

- CNN takes a picture fix of n × n pixels as info and afterward does characterization as indicated by the example of the picture fix; it learns the examples of fix appearance from a lot of preparing patches. The yields of CNN are the scores for various classes, and the class with the most noteworthy score is considered as the characterization result.
- CNN likewise tries not to utilize the discussed surface highlights which are influenced by tumor size.
- CNN uses the picture appearance design around the tumor. The appearance design incorporates data of neighborhood contrast, close by tissues, limit sharpness, and so forth Such data is not the same as yet as ground-breaking as the symptomatic highlights.
- CNN are not influenced by the size of the tumor, since they are figured from the whole picture fix which incorporates both the tumor and its encompassing tissues.

Wang and partners (2017) showed that SVM has a place with the piece based classifier family, which verifiably maps the information highlights into a higher dimensional element space utilizing a bit work that quantifies the distance between include focuses in the planned space, so SVM can accomplish much preferable characterization execution over customary straight grouping techniques [5].

This examination expected to analyze the presentation of numerous AI strategies for characterizing head and neck cancer from histopathological images.

II. MATERIAL

A. Data Sets

In this examination, informational indexes were gathered from two sources. The HNSCC and Normal histology pictures were gotten from The Ethical Tissue division at the University of Bradford, while Salivary Glands pictures from WebPathology web-page. An all out 2608 histopathological pictures, 1184, 1184 and 240 for HNSCC, salivary organs and ordinary histology, individually. These pictures were put away in jpg design.

B. Data Augmentation

Dosovitskiy et al. (2013) explored the part of information enlargement in profound figuring out how to get enough various examples which expected to prepare a CNN from the pictures [3]. The dataset pictures were pivoted (90, 180 and 270), flipped left to right on a level plane and afterward vertically to make a bigger example size and to cause the way to deal with perceive tumor cells in various directions (figure1).
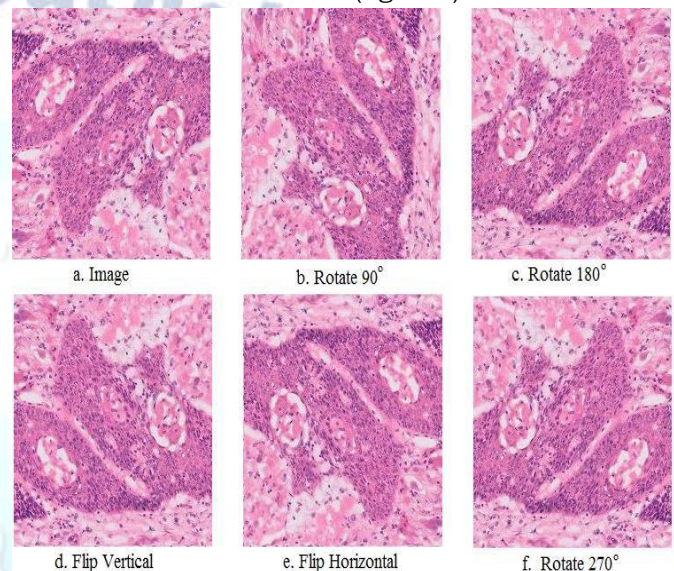

a. Image b. Rotate 90° c. Rotate 180°

d. Flip Vertical e. Flip Horizontal f. Rotate 270°

Figure 1 shows Data Augmentation of the original image

III. METHODOLOGY

A. Feature Extraction

Feature extraction is a way to use pre-trained networks without consuming time and effort into training. Learned features were extracted from a pre-trained network, and then used to train a classifier.

B. AlexNet Model

AlexNet model (Krizhevsky et. al.,2012) is trained on 1.2 million images of the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) 2010 dataset [1]. The ILSVRC 2010 training set contains 1000 distinct classifications, speaking to everything articles, for example, blossoms, vehicles, creatures, etc [1]. As portrayed in figure 2, the profound AlexNet design is made of a pile of eight layers [1]: the initial five layers are convolutional and the leftover three are completely associated. The info layer is designed to take a fixed estimated 227 x 227 RGB picture as an information; they standardize all the preparation pictures to get similar scope of qualities for every one of the information highlights. The first convolutional

layer channels the info picture with 96 pieces (size: 11x11x3) with a step of 4 pixels. The second convolutional layer channels the yield (standardized and pooled) of the first convolutional layer with 256 portions (size: 5x5x48). The third, fourth, and fifth convolutional layers are associated with no mediating pooling or standardization layers. The third convolutional layer has 384 bits (size: 3x3x256), the fourth convolutional layer has 384 portions (size: 3x3x192), and the fifth convolutional layer has 256 pieces (size: 3x3x192). The completely associated layers have 4096 neurons each. The yield of the last completely associated layer is taken care of to a 1000 classes. A softmax layer creates the likelihood dispersion for the yields of the last completely associated layer by changes them over to genuine qualities somewhere in the range of zero and one with entirety one.
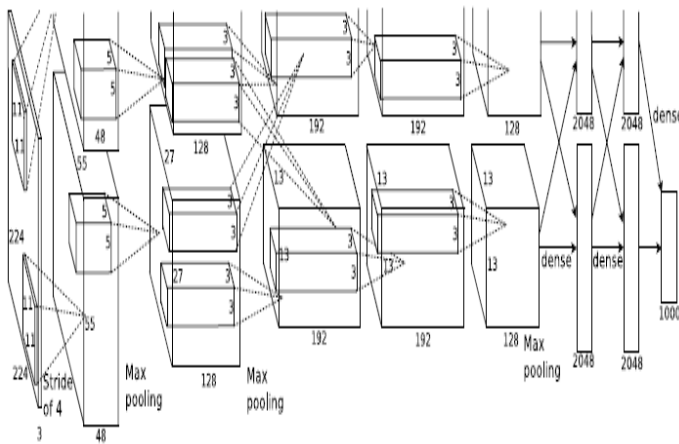

Figure 2 shows AlexNet Architecture

C. *Machine Learning Algorithms*

- SVM uses an *error-correcting output codes (ECOC)* algorithm to classify multiclass models. ECOC decreases the issue of grouping with at least three classes to a bunch of parallel classifiers. ECOC models can improve grouping exactness, even contrasted with other multiclass models. ECOC grouping allots a one-versus-one coding plan, which decides the classes that the double students train on, and an interpreting plan, which decides how the outcomes (expectations) of the parallel classifiers are totaled . Leave M alone the coding plan lattice with components m_{kl}, and s_l be the anticipated characterization score for the positive class of student l. A novel perception is appointed to the class ($\bar{k}$) that limits the collection of the misfortunes for the L parallel students. That is,

$$\hat{k} = \arg\min_k \frac{\sum_{l=1}^{L} |m_{kl}| g(m_{kl}, s_l)}{\sum_{l=1}^{L} |m_{kl}|} \tag{1}$$

- The k-nearest neighbor algorithm estimates the predictors inside each class by looks for in the perception to discover the closest focuses to predictor points and response values to those nearest points, and then it classifies an observation by estimating the posterior probability for each class and expected classification cost; Let y is the predicted classification, K is the number of classes, P(k|x) is the posterior probability of class k for observation x and C(y|k) is the cost of classifying an observation as y when its true class is k .

$$\hat{y} = \arg\min_{y=1,\ldots,K} \frac{\sum_{k=1}^{K} P\langle k | x \rangle C\langle y | k \rangle}{\sum_{k=1}^{K} |m_{kl}|} \tag{2}$$

- The discriminant analysis algorithm estimates the predictors within each class by using a multivariate normal distribution. It accepts that indicator has a Gaussian blend conveyance. For straight discriminant examination, the model has a similar covariance framework for each class, just the methods change. It utilize expected order cost (condition 2) for expectation .

- Naive Bayes algorithm estimates the densities of the predictors within each class, models posterior probabilities according to Bayes rule, and then it classifies an observation by estimating the posterior probability for each class, and then assigns the observation to the class yielding the maximum posterior probability. Let Y is the random variable corresponding to the class index of an observation, $X_1,\ldots,X_P$ are the random predictors of an observation and π(Y=k) is the prior probability that a class index is k

$$\hat{P}(Y = k|X_1, \ldots, X_p) = \frac{\pi(Y = k)\prod_{j=1}^{P} P(X_j|Y = k)}{\sum_{k=1}^{K} \pi(Y = k)\prod_{j=1}^{P} P(X_j|Y = k)} \quad for\ all\ k = 1, \ldots, K$$

(3)

☐

● The Binary Decision Tree algorithm (Coppersmith et. Al., 1999) represents the observation in binary tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. It gauges the indicators inside each class by following the parts of perception until it arrives at a leaf hub If arrives at a leaf hub, it restores the characterization of that hub. Anticipated Class Label was ascertain be limit the normal grouping cost (Equation 1). For trees, the score of a characterization of a leaf hub is the back likelihood of the grouping at that hub. The back likelihood of the grouping at a hub is the quantity of preparing successions that lead to that hub with the arrangement, isolated by the quantity of preparing successions that lead to that hub [4].

D. Pre-processing Step

The study was performed in a MATLAB environment. For
classification, it should be set down that 80% of data were
randomly chosen for training. The remaining 20% were used
for testing. Bukar & Ugail (2017) demonstrated that training set of data should be considerably larger in order to give more accurate results [2]. AlexNet model requires input pictures of size 227-by-227-by-3, yet the dataset pictures have various sizes. So the training and test images were resized to stature 227 and width 227 before they are input to the pre-training network.

E. Standard Performance

The machine learning classifiers were implemented and evaluated according to standard performance such as Accuracy (ACC), Precision (P) and Sensitivity (S). Sokolova and Lapalme (2009) extracted these terms from the confusion matrix [6]. The standard execution estimations were figured as portrayed in the accompanying conditions [6]; where exactness assesses the general viability of a classifier; Precision assesses the class arrangement of information names with the positive names given by the classifier; Sensitivity

assesses the adequacy of a classifier to distinguish positive names; Specificity assess the adequacy of a classifier recognizes negative marks; Area Under Curve assesses Classifier's capacity to dodge false classification.

$$Accuracy = \frac{(tp + tn)}{(tp + fn + fp + tn)} \qquad (4)$$

$$Precision = \frac{tp}{tp + fp} \qquad (5)$$

$$Sensitivity = \frac{tp}{tp + fn} \qquad (6)$$

$$Specificity = \frac{tn}{tn + fp} \qquad (7)$$

$$AUC = \frac{1}{2}\left\{\frac{tp}{tp + fp} + \frac{tn}{tn + fp}\right\} \qquad (8)$$

IV. RESULTS

The performance results of using AlexNet model to build machine learning classifier were reported in the table 1. The five methods were implemented using the functions of MATLAB R2018a. These results used 2 times 10-fold cross-validation to evaluate the classifiers. For each of cross-validation, the performance values were calculated for each feature set based on the nine folds of training samples, via grid search in the parameter space. Therefore, each cross-validation might have slightly different values, and the average optimal value was reported. The Classifier used a linear function as the kernel function, Layer 'fc7' to extract features from dataset images, and the sequential minimal optimization method to find the separating hyper plane, its average kernel size was 2.0. A 3 X 3 confusion matrix was shown in figure 3.

Classifier	ACC(%)	P	S	Spe	AUC
SVM	99.86	.9986	.9986	.9986	.9986
KNN	99.86	.9986	.9986	.9986	.9986
NB	93.68	.8982	.8839	.8839	.8911
BDT	97.32	.9692	.9502	.9502	.9597
DA	99.81	.9861	.9972	.9972	.9917

Table 1 illustrates the standard performance for Classifiers

(a) Support Vector Machine Classifier
(b) Binary Desision Tree Classifier
(c) K-Nearest Neighbor Classifier
(d) Naïve Bayes Classifier
(e) Discriminant Analysis Classifier
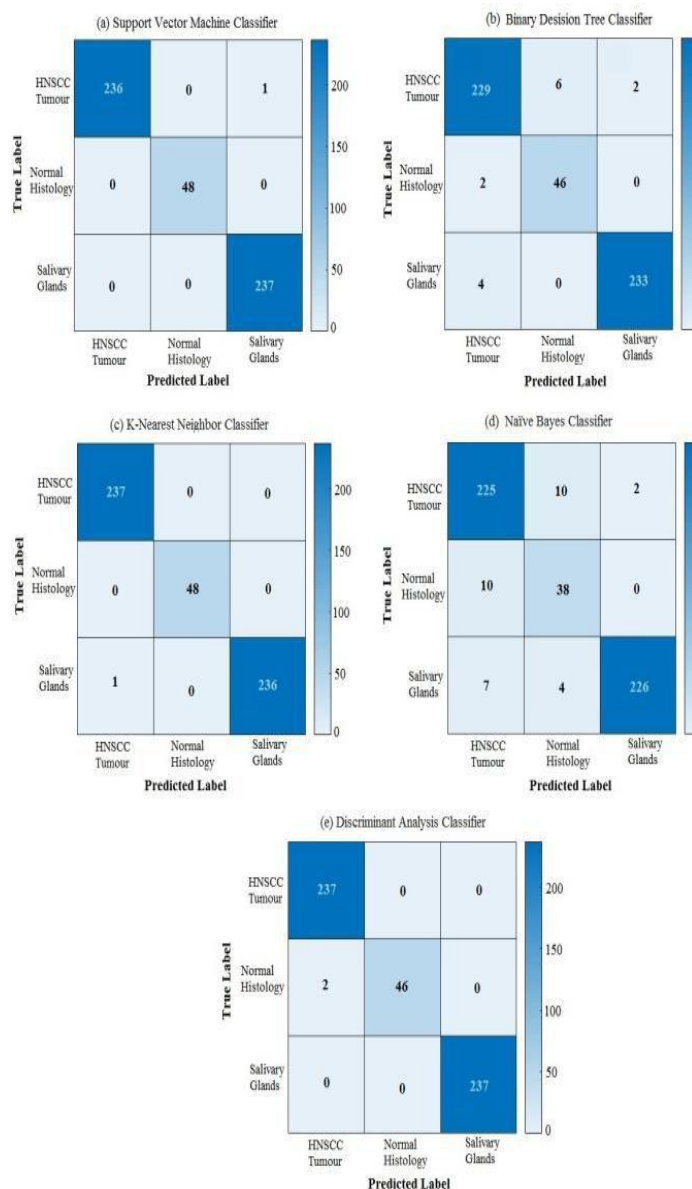
Figure 3 shows the confusion matrix for Classifiers

V. ACKNOWLEDGMENT

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." Advances in neural information processing systems. 2012.

[2] A.M. Bukar, and H. Ugail. "Convnet Features for Age Estimation", Bradford, UK, ISBN: 9978-989-8533-66-1. 2017.

[3] A.Dosovitskiy, J. T. Springenberg and T. Brox."Unsupervised feature learning by augmenting single images". CoRR, abs/1312.5242. 2013pdf2text library https://github.com/syllabs/pdf2text

[4] D.Coppersmith, S. J. Hong, and J. R. M. Hosking. "Partitioning Nominal Attributes in Decision Trees." Data Mining and Knowledge Discovery, Vol. 3, 1999, pp. 197–217.

[5] H. Wang, Z. Zhou, Y. Li, Z.Chen, P. Lu, W. Wang, W.Liuand L.Yu. "Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18F-FDG PET/CT images",EJNMMI Research 7:11, DOI 10.1186/s13550-017-0260-9. 2017.

[6] M. Sokolova and G. Lapalme. "A systematic analysis of performance measures for classification tasks". Inf. Process. Manage., 45(4):427–437, July 2009.