



# MULTI LABEL CLASSIFICATION FOR AN IMAGE USING CONVOLUTIONAL NEURAL NETWORKS

<sup>1</sup>N. Lakshmi Prasanna; <sup>2</sup>R. Vaishnavi; <sup>3</sup>V. Prasanna Lakshmi; <sup>4</sup>V. Dakshayani; <sup>5</sup>T. Keerthana

<sup>1</sup>Associate Professor, Dept of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur (Dist.), Andhra Pradesh, INDIA

<sup>2,3,4,5</sup>B-Tech, Dept of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur (Dist.), Andhra Pradesh, INDIA

DOI: 10.47760/ijcsmc.2021.v10i07.001

**Abstract**— The machine learning has many capabilities one of them is classification. Classification employed in many contexts like telling hotel reviews good / bad, or inferring the image consists of dog, cat etc. As the data increases there is a need to organize it, for that purpose classification can be helpful. Modern classification problems involve the prediction of multiple labels simultaneously associated with a single instance known as "Multi Label Classification". In multi-label classification, each of the input data samples belongs to one or more than one classes or labels. The traditional binary and multi-class classification problems are the subset of the multi-label classification problem. In this paper we are implementing the multi label classification using CNN framework with keras libraries. Classification can be applied to different domain such as text, audio etc. In this paper we are applying classification for an image dataset.

**Keywords**— classification, machine learning, multi label classification, CNN (convolutional neural networks), keras.

## 1. INTRODUCTION

**Classification** is a datamining function that assigns items in a collection to a target labels or classes. It predicts the target class for each case in data. This task begins with dataset in which the class assignment are known. The historical data for a classification project is typically divided into 2 data sets: one is for building model and second one is for testing model. **Multi label classification** is the process in which each sample is mapped to a set of target labels among the set of training labels.



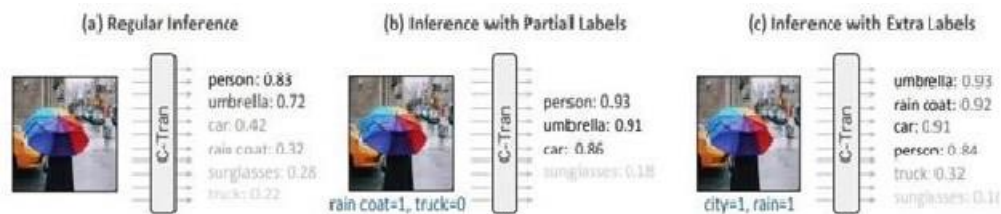
1.1: Image 1



1.2: Image 2

To be more clear on multi label classification definition, consider the above 2 images. Fig:1.1 consists of only one object/feature (i.e; car), so to label the image 1 only one label “car”. “Binary Classification” is the process which involves only 2 outcomes (training set: {car} output set: {yes/no}). To classify image1 binary classification is sufficient. Otherwise we can go with “multi class classification”, which involves many training labels but at the time of label assignment only 1 label should be assigned (training set: {car, trees, house} output set: {car}).To classify Fig:1.2 these two classifications will not sufficient as it consisting of many objects, here comes the need of “multi label classification”.

Images in real-world applications generally consists of many objects/features and can portray complex situations. Multi-label image classification is a visual recognition task that goals to predict a set of labels corresponding to objects, attributes, or actions given in an input image. This objective goes beyond the more well studied single-label multi-class classification problem where the goal is to extract and associate image features with a single concept per image. In the multi-label setting, the output set of labels has some structure that reflects the structure of the world. For example, dolphin is unlikely to co-occur with grass, while knife is more likely to appear next to a fork. To this end, we present the Classification Transformer (C-Tran), a multi-label classification framework that leverages a Transformer encoder. Transformers have demonstrated a remarkable capability of being able to exploit complex and rich dependencies among sets of inputs using multiple layers of multi headed self-attention operations. In our approach, a Transformer encoder is trained to rebuild a set of target labels/classes given an input set of masked label embeddings and a set of features obtained from a convolutional neural network (CNN). Unlike the Transformer encoders used for language modelling, C-Tran uses a label mask training objective that allows us to represent the state of the labels as positive, negative, or unknown. At the time of testing, C-Tran is able to predict a set of target labels using only input visual features by masking all the input labels as unknown. Figure 1 gives an overview of this strategy. We demonstrate that this approach leads to superior results on a number of benchmarks compared to other recent approaches that exploit label relations using graph convolutional networks and other recently proposed strategies. Beyond obtaining state-of-the-art results on the introduced regular multi-label classification task, we also claim that C-Tran is a more general model for reasoning under prior label observations. Because our approach models the uncertainty of the labels during training, it can also be used at test time with partial or extra label annotations by setting the state of some of the labels as either positive or negative instead of masking them out as unknown.



**1.3: Sample Image with feature assignment**

Consider the example image shown in Figure 1.3(a) where a model is able to predict person and umbrella with relatively high accuracies, but is not confident for categories such as rain coat, or car that are clearly present in the picture. Suppose we know some labels and set them to their true positive (for rain coat) or true negative (for truck) values. Provided with this new information, the model is able to predict car with a high confidence as it moves mass probability from truck to car, and predicts other objects such as umbrella with even higher confidence than in the original predictions (Figure 1.3(b)). In general, we consider this setting as realistic since many images also have metadata in the form of extra labels such as location or weather information (Figure 1.3(c)). This type of conditional inference is a much less studied problem. Our general approach to multi label image classification with Transformers is able to naturally handle all these scenarios under a unified framework. We compare our results with a competing method demonstrating superior results under variable amounts of partial or extra labels.

## 2. RELATED WORK

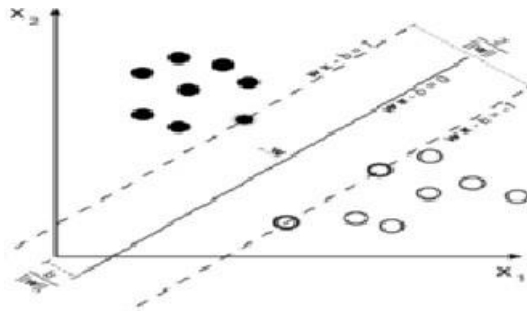
### 2.1 MACHINE LEARNING

Machine learning (ML) is the study of computer algorithms which can improve automatically through experience by executing different tasks. Machine Learning is a sub field/set of artificial intelligence (AI). The algorithms of machine learning build a mathematical model based on data provided, known as "training data", in order to make predictions or decisions without being explicitly programmed, to do so ML algorithms are used in a various number of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop conventional algorithms to perform the needed real world tasks.

The mathematical optimization study gives methods, theory and application domains to the field of machine learning. Data mining is a field of study related to ml, focusing on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning can be called as predictive analytics.

Early days classifications for machine learning approaches divided them into 3 categories, depending on the nature of the signal or feedback available to the learning system. These were:

**Supervised learning:** The computer is provided with the sample data called "training data" , by accessing and learning from it the computer can generate the output



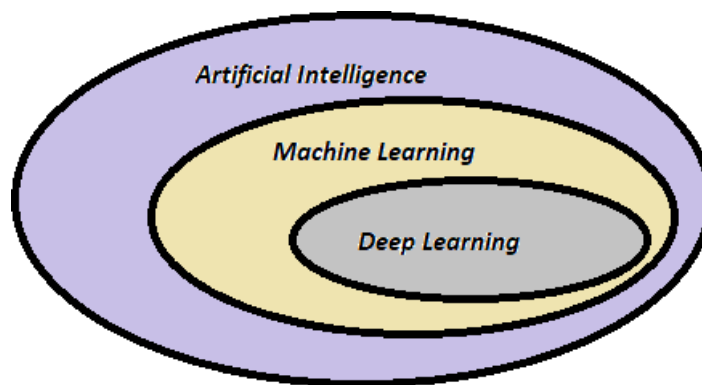
**2.1: graph for supervised learning**

**Unsupervised learning:** No labels are given to the learning algorithm (no training data), leaving it on its own to find structure in its input. Unlike supervised learning Unsupervised learning can be a goal in itself or a means towards an end (feature learning).

**Reinforcement learning:** A machine program (lines of code) interacts with a dynamic environment and there it must perform a certain task/goal as it navigates its problem space, the program is provided feedback that's analogous to rewards/outputs, which it tries to maximize.

Other approaches or processes developed but don't fit neatly into this three-fold categorization, and sometimes more than one approach is used by the same machine learning system. As of 2020, deep learning had become the best approach for much ongoing work in the field of machine learning.

**Deep Learning** is a subset of machine learning concerned with algorithms evolved by the structure and function of the brain called artificial neural networks (ANN). In addition to scalability, another usage of deep learning models is the ability to perform automatic feature extraction from raw data, also called feature learning.



**2.2: Basic Structure of ML**

## 2.2 RELATION TO DATA MINING

Machine learning and data mining utilize the same methods, but while machine learning focuses on prediction, based on known properties which are learned from the training data given, data mining mainly focuses on the discovery of (previously) unknown properties in the data (this is the analysis step of knowledge discovery in databases). Data mining utilizes many number of machine learning methods or properties, but with different goals; on the other hand, machine learning also uses data mining methods or features as "unsupervised learning" or as a pre-processing step to improve learner accuracy.

### 2.3 MULTI LABEL CLASSIFICATION

Classification is one of the data mining feature or function that assigns the items in a collection to a target categories or classes. It predict the target class for each case in data. This task begins with datasets in which the class assignments are known. Classification models are tested by comparing the predicted values to target values in a set of test data. The data for classification is divided into 2 data sets, one for building a model second one is for testing a model. A classifier is the algorithm that maps the input data to specific category. These classifiers can be of discriminative or generative. A feature is an individual measurable property of process. Multi label classification is the process in which each sample is mapped to a set of target labels.

i.e, training set {car, house, tree}, a particular image can be labelled with 1 or 2 or all 3 from the training data set.

### 2.4 DATASET

There are various datasets that you can leverage for applying convolutional neural networks like images, text, videos etc. We collected 306 images of 256x256 manually consisting the bit depth (number of bits used to represent each pixel in an image) of 24 (8bits\*3channels) and the features desert, mountain, sea, sunset, trees. After collecting the Images we prepared an excel sheet with 5 features and filenames of images and filled the values of 0&1 for those features manually (or these images can be represented using the csv file consisting of the labels for images with comma separated e.g: 4.jpeg desert, mountain). Created a csv file with those features to the 167 images, which are going to be used as training data set and 39 images as validation data set for our model. Remaining 100 images are for testing. These images are given to code as an input in the .csv file format. For that we manually created a .csv file and entered the comma separated values into it for each and every image as shown in fig:2.3. On passing this .csv file to the code it will read file as an array of arrays and process it.

	A	B	C
1	Filenames	labels	
2	1.jpg	desert	
3	2.jpg	trees	
4	3.jpg	desert	
5	4.jpg	desert,mountains	
6	5.jpg	desert	
7	6.jpg	desert	
8	7.jpg	desert,mountains	
9	8.jpg	desert	
10	9.jpg	desert	
11	10.jpg	desert,mountains	

2.3: Input data file

### 3. REVIEW OF LITERATURE

Pooja.V.Magdum, Mahadev.S.Patil [1] published a paper which tells about the comparison of neural networks, Deep learning is sub field/set of machine learning, which aims to learn hierarchy of features from input data. Deep learning technologies now a days are becoming major approaches for natural signal and information processing like image classification and speech recognition. Deep learning is a technology inspired by working/functioning of human brain. Convolutional neural network (CNN) becomes very popular for image classification in deep learning. In this paper, it is discussed about the deep learning's convolutional structures based on Keras. Four different structures of CNN are compared on CPU system with four different combinations of classifiers and activated functions. Yun Chen et al [2] explored, proposed and improved different neural network architectures to take contextual data. Different neural networks are applied to text present in Chinese public text dataset and got good results. Initially, the architecture called inception is designed to extract features of n-gram and then a simple modification is done on top of RNN models to extract more information.

Xinsheng Wang et al [3], proposed an improved CNN via hierarchical dirichlet process (HDP) model to manage the multi-label classification problem in NLP. At first an HDP model is applied to remove the words which are semantically less important. Finally, CNN is trained based on word vectors. Experimental results have

shown that this method is better than traditional multi-label classification methods and Text CNN in terms of performance. Rajni Jindal *et al* [4], proposed a new method for automated categorization of text documents. The proposed method is constructed on lexical and semantics concepts. It is a new idea in the area of categorization of research articles. It can be applied to both single-label and multi-label text documents. The proposed method is experimented and tested with a dataset of IEEE research articles. It has given a good performance with an accuracy of 75%. In 2019, MD. Aslam Parwez *et al* [5], presented various neural network models for multi-label classification of micro blogging data. These models are based on convolutional neural network (CNN) architectures, which make use of pre-trained word embedding's from generic and domain-specific textual data sources. To predict class labels, the word embedding's are handled separately and in different combinations through different channels of CNN. A comparative analysis of the proposed CNN models with state-of-art machine learning models and one of the existing CNN architectures is done. The proposed models are assessed over a real Twitter dataset, and the results have shown their efficiency to classify micro blogging texts with improved accuracy.

Che-Ping Tsai *et al* [6], proposed a creative structure to train a multi-label classifier. A Generative Adversarial Network (GAN) is used to imitate the label dissemination for multi label classification. This structure is built upon a conditional GAN (cGAN). The classifier acts as a conditional generator, which has some instance as input and it outputs a set of labels like a multi label classifier. This framework also has a discriminator which is trained to model label dependency, that is it takes an object and a set of labels as input, and outputs a score. The discriminator learns to discriminate the real and generated label sets. The generator then learns to fool the discriminator by generating label sets with what seems to be the correct dependencies, given an input instance by the discriminator. The classifier and discriminator are learned iteratively as in a typical GAN. Wang, jiang & yang [7] proposed a unified framework for multi label image classification. CNN is a type of feedforward artificial neural network (ANN) with variations of multilayer perceptions designed to use minimal amounts of preprocessing. RNN unlike feed forward neural networks, can use their internal memory to process arbitrary sequences of input data. NUS-WIDE, MS\_COCO, PASCAL VOC 2007 are the datasets used. Algorithms used are rmsprop is an optimization algorithm, K nearest neighbor search for classification of labels. CNNRNN models can be achieved by using the cross entropy loss on the softmax normalization of score and employing back proposition through time algorithm. Experimental results on these datasets demonstrate that the proposed approach achieves superior performance to the state-of-the-art method. It is an interesting direction to not only predict the labels, but also predict the segmentation of the objects by constructing an explicit attention models.

Ying hong [8] published a paper on multi label image classification on clothing images with convolutional neural networks. Online shopping is taking a crucial part in today's lifestyle so they worked on the clothing image sets to classify them according to the user interest. In this paper, he crawled the clothing images from the internet. The dataset contains six categories and it was divided into training set and validation set. In our work, he built a convolutional neural network model and trained a multi-label classifier to predict both clothing type and color. In 2019, Yan Luo, Ming Jiang, Qi Zhao [9] proposed a paper on visual attention in multi label classification i.e. learning of representative features that capture the rich semantic information in a cluttered scene is the significant task in image classification. Specifically, they propose a dual-stream neural network that consists of two subnetworks: one is a conventional classification model and the other is a saliency prediction model trained with human fixations. Features that compute with the two sub-networks are trained separately and then fine-tuned together using a multiple cross entropy loss. Experimental results show that the additional saliency sub-network improves multi-label image classification performance on the MS COCO dataset as per our model. This improvement is consistent across various levels of scene clutteriness. The applications of multi label classification are present in the paper published by Ayesha Mariyam, SK Althaf Hussain Basha, S Vishwanadha Raju in 2020 [10]. In this paper they explored and discussed various recent advancements in the field of multi-label text classification which presents future scope as well and the evaluation metrics. The applications and research areas of text classification are also mentioned.

#### 4. ALGORITHM AND FLOW MODEL

##### 4.1 MODEL

Convolutional Neural Network (ConvNet/CNN) is one of the Deep Learning algorithms which can take an input image, assign importance (learnable weights and biases) to various features in the image and be able to differentiate one from the other. The pre processing required in a CNN is much lower as compared to other classification algorithms for an image. While in primitive methods, the filters are hand engineered, with enough training data, ConvNets have the ability to learn these filters/characteristics effectively. A ConvNet/CNN is successfully take the Spatial and Temporal dependencies of an image through the application of relevant filters.

The architecture of a ConvNet/CNN is corresponding to that of the connectivity pattern of Neurons in the human brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted part of the visual field known as the Receptive Field. A collection of such fields combine to cover the entire visual area.

## 4.2 PROCESS

As a scientific endeavor, machine learning grew out of the quest for artificial intelligence. In the early days of AI as an academic discipline, some researchers were working on a concept of machines learn from data. They attempted to approach the problem with various symbolic methods, as well as what were then termed "neural networks".

We describe the label set-based solutions under a general framework of meta-labels and provide some theoretical justification with convolution neural networks.

Machine learning and data mining often use the same methods and overlap significantly, but while machine learning focuses on prediction based on known properties learned from the training data, data mining focuses on the discovery of (previously) unknown properties in the data.

This project was going to be implemented by using "keras" in "Anaconda tool".

There are 3 scenarios

1. Segmentation model: only for small size files and the size should be fixed for all the training as well as test images ex.256x256.
2. Feature identification: In this model the feature maps or activation maps are generated using horizontal, vertical sizes of the objects in image.
3. ROI (Region Of Interest) calculation: This model is used to classify labels for images but no same size criteria i.e, the size of images is not considered.

In this project the domain specific dataset is taken and stored in csv files so that it is easy to process. We can add extra features to the file accordingly as we need.

By using 3 scenarios which were described above are used to find the probability values for a specific testing image set with the help of training image data set. If that probability value is  $\geq 0.5$ , then we will consider that feature for that testing image.

This was the criteria considered for this multi label classification.

## 4.3 STEPS FOR MULTI LABEL CLASSIFICATION:

Convolutional neural network is a class of deep neural networks that is mostly used for computer vision or analyzing visual imagery tasks. Convolutional layer: computers read images as pixels and it is expressed as a matrix (n $\times$ n $\times$ 3) (height by width by depth). Generally, images make use of three channels (RGB), so that is why we have a depth of three.

Steps used to build your Multi-Label Image Classification Model

**1. The first step is to get our data in a structured format** . This can be applied to both binary as well as multi-class image classification.

You should have a folder containing all the images on which you want to train your model. Now, for training this model, we require the true labels of images. So, you should also have a .csv file which contains the names of all the training images and their all true labels.

Once the data is ready in .csv file format, we can divide the further steps as follows.

- Python has pandas to manipulate the data sets which can be imported easily and used.

### 2. Load and pre-process the data

Load all the images from dataset taken and then pre-process them as per our project's requirement. To check how our classification model will perform on unseen data (test data), we create a validation set of images. We train our model on the training set of images and validate it using the validation set of images (standard machine learning practice)

- With the function `read_csv()` in pandas the csv file can be loaded into the code as an array.

### 3. Define the model's architecture

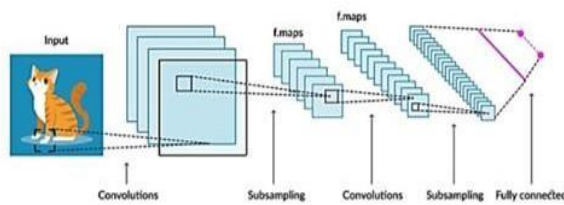
The next step is to define the architecture of the model. This includes deciding the number of hidden layers, number of neurons in each layer, activation function, and so on useful for dividing image into layers.

• By using keras, the model can be created and the necessary features can be added with `add()` function or can be deleted with `pop()` function.

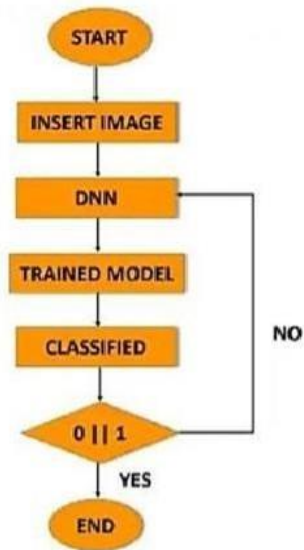
#### 4. Train the model

After creating and defining the model we have to train our model on the training set! We pass the training images and their corresponding labels/classes to train the model previously created. We also pass the validation images here which can validate how well the model will perform on unseen data.

- We are using the rmsprop algorithm to find the root mean square values while compiling the model. We pass the validation images here which validate how well the model will perform on unseen data. use the keras image data generators to augment images.



4.1: convolutional neural networks model



4.2: flow chart for image classification

#### 5. Make predictions

we use this trained model to get predictions on testing images.

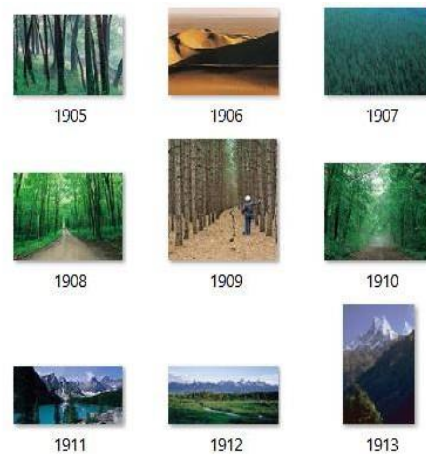
- Predict the generator with a condition of probability > 0.5. That is if the predicted value is at least 50% closer to the original values the particular image. The resultant data stored in a file with further statements.

#### 5. RESULTS

Input provided to our code is the image data set. The fig:5.1 shows some of the images used for training and testing from our dataset gathered. These images may not consist all the 5 features that we have considered and may not consists of even one features our set.



5.1(a): training data set images



5.1(b): testing data set images

On running the code as it's first step it will read the image datasets by giving the below lines to terminal. The dataset that considered for this project gives 0.46 accuracy. Our model is executed with data set consisting of 2000 images and it is working fine with the accuracy of 0.55. One more thing that we observed with this model is if the testing image has very near features/labels of training features then automatically the accuracy is increasing. The testing data set that we considered have different images consisting of different objects (other than training data set labels eg. Sand, sky etc) in each so our code accuracy is 0.46.

```

Anaconda Prompt (anaconda3)
(base) C:\Users\Venkatesh>cd Desktop
(base) C:\Users\Venkatesh\Desktop>cd miml_dataset
(base) C:\Users\Venkatesh\Desktop\miml_dataset>activate tensorflow
(tensorflow) C:\Users\Venkatesh\Desktop\miml_dataset>python sample.py
Found 167 validated image filenames belonging to 5 classes.
Found 39 validated image filenames belonging to 5 classes.
Found 100 validated image filenames.
1/5 [====>.....] - ETA: 1:00 - loss: 0.72
2/5 [=====>.....] - ETA: 4s - loss: 1.1146
3/5 [=====>.....] - ETA: 2s - loss: 1.1721
4/5 [=====>.....] - ETA: 1s - loss: 1.1558
5/5 [=====] - ETA: 0s - loss: 1.1176
5/5 [=====] - 22s 2s/step - loss: 1.0922 - accuracy: 0.4544 - val_loss: 0.5850 - val_accuracy: 0.4000
100/100 [=====] - 3s 28ms/step
    
```

**5.2: output of the code**

To store result create an empty csv file with name used in code (results.csv), after running the code the file is filled with 2 columns(filename, labels) and 100 rows (data) as shown below. It consists of more than one label assigned to each class (some of them have 2, 3) which is said to be multi label classification.

	A	B	C
9	1908.jpg	mountains,trees	
10	1909.jpg	mountains,trees	
11	1910.jpg	mountains,trees	
12	1911.jpg	mountains,trees	
13	1912.jpg	mountains,trees	
14	1913.jpg	mountains,trees	
15	1914.jpg	mountains,trees	
16	1915.jpg	mountains,trees	
17	1916.jpg	mountains,trees	
18	1917.jpg	mountains,trees	
19	1918.jpg	mountains,trees	
20	1919.jpg	mountains,trees	
21	1920.jpg	mountains,trees	

5.3: result data file



## Conclusion and Future Work:

The multi-label classification task is the most critical problem. We explored and discussed various recent advancements in the field of multi-label image classification which presents future scope as well and the evaluation metrics. A sparse set of existing algorithms has been explained. We have focused on the lazy learner based algorithms which use CNN as the learning algorithm. In conclusion multi-label classification techniques have been reviewed. As a future work, will carry out an extensive analysis of our proposed algorithm on various dimensions of datasets. In future, we can improve the performance of our proposed algorithm.

# REFERENCES

- [1]. Pooja.V.Magdum, Mahadev.S.Patil, "Comparison of Different Convolutional Neural Network Structures Based on Keras", *vol5 issue1 Journal of Electronics and communication systems*, 2020.
- [2]. YunChen, BoXiao, ZhiqingLin, ChengDai, ZuochaoLi and LipingYan, "Multi-label Text Classification with Deep Neural Networks," *Proceedings of IC-NIDC, 2018*, pp.409-413.
- [3]. Xinsheng Wang, Lijun Sun, Zhihua Wei, "An Improved Convolutional Neural Network Algorithm for Multi-label Classification", *ICALIP, IEEE, 2018*, pp.113-117.
- [4]. Rajni Jindal, Shweta, "A Novel Method for Efficient Multi-Label TextCategorization of research articles", *IEEE International Conference on Computing, Power and Communication Technologies, 2018*, pp.333-336.
- [5]. MD. Aslam Parwez, Muhammad Abulaish, Jaheeruddin, "Multi-Label Classification of Micro blogging Texts Using Convolution Neural Network," *IEEE ACCESS, 2019*, pp.68678-68691.
- [6]. Che-PingTsai,Hung-YiLee, "Adversarial Learning of Label Dependency: A Novel Framework for Multi-Class Classification", *IEEE, 2019*, pp.3847-3851.
- [7]. wang, jiang & yang, "CNN- RNN : A unified framework for multi label image classification" *IEEE 2018*.
- [8]. Ying Hong, Research on multi label clothing image classification based on convolution neural network in 2020.
- [9]. Yan Luo, Ming Jiang, Qi Zhao: visual attention in multi label image classification in 2019 *IEEE papers*.
- [10].Ayesha Mariyam, SK Althaf Hussain Basha, S Vishwanadha Raju, "Applications of Multi-Label Classification" 2020.
- [11].He, S., Xu, C., Guo, T., Xu, C., & Tao, D. (2018). "Reinforced Multi-Label Image Classification by Exploring Curriculum", *Vol. 32 No. 1 (2018): Thirty-Second AAAI Conference on Artificial Intelligence*.
- [12].Mingkui Tan, Qinfeng Shi, Anton van den Hengel, Chunhua Shen, Junbin Gao, Fuyuan Hu, Zhen Zhang; "learning graph structure for multi label classification via clique generation" *2015*, pp. 4100-4109.
- [13].J. Mao, W. Xu, Y. Yang, J. Wang, and A. Yuille. Deep captioning with multimodal recurrent neural networks (m-rnn). In *ICLR, 2015*.
- [14].Wu, F., Wang, Z., Zhang, Z., et al.: Weakly semi-supervised deep learning for multi-label image annotation. *IEEE Trans. Big Data* **1**(3), 109–122 (2017).
- [15].Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1097–1105. *Curran Associates Inc. (2012)*.
- [16].Y. Wei, W. Xia, J. Huang, B. Ni, J. Dong, Y. Zhao, and S. Yan. Cnn: Single-label to multi-label. *arXiv preprint arXiv:1406.5726, 2014*.
- [17].X. Li, F. Zhao, and Y. Guo. Multi-label image classification with a probabilistic label enhancement model. *UAI, 2014*.
- [18].Ajay Shrestha and Ausif Mohmood (April, 2019), "Review of Deep learning algorithms and architectures", *IEEE, Volume 7*, pp. 53040-53065, DOI: 10.1109/ACCESS. 2019.2912200
- [19].Connor shorten and Taghi M. Khoshgoftaar (2019), "A survey on image data augmentation for deep learning", *Journal of Big Data, Volume 6, Article 60*, DOI: 10.1186/s40537-019-0197-0.
- [20].P. L. F. de Carvalho and A. Freitas, "A Tutorial on Multi-label Classification Techniques", in *Foundations of Computational Intelligence, Volume5, Springer Berlin Heidelberg, 2009*, pp.177-195.