2023 4th International Conference for Emerging Technology (INCET)
Belgaum, India. May 26-28, 2023

Enhanced Zero shot Learning using Deep Neural Network ResNet50

Ansari Shaista Khanam
EXTC Department
Rajiv Gandhi Institute of Technology
Mumbai, India
shaista.khan@vcet.edu.in

Dr.Poonam . N.Sonar
EXTC Department
Rajiv Gandhi Institute of Technology
Mumbai, India
poonam.sonar@mctrgit.ac.in

2023 4th International Conference for Emerging Technology (INCET) | 979-8-3503-3575-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/INCET57972.2023.10170025

***Abstract*— Computer vision has advanced with lot of development in visual recognition systems, which poses restrictions to expand for huge numbers of image classes. This restriction is due to huge image classes with unlabeled images which cannot be used in training the machine learning algorithm. As Traditional machine learning method of classification are based on the classification of categories which are available at the time of training. Technique of Zero-shot learning (ZSL) recognizes categories of test sets which is not appearing while training the model. The enhanced ZSL technique proposed is based on deep visual semantic embedding method. In this method Visual and semantic features are used for the classification of unknown categories. The extraction of visual features is accomplished with convolutional neural network (CNN) and ResNet 50. FastText is used to convert labels of classes into word embedding vectors. Visual and word embedding features are mapped. The model is predicting the Top 5 labels for an unknown image category (zero-shot class). The experiments are performed on standard datasets SUN and AWA2. The proposed technique of enhanced ZSL with ResNet 50 gives better accuracy and reduced model loss then CNN model.**

Keywords*— *word embedding, CNN, ResNet50 Zero shot learning, visual semantic embedding, word embedding, CNN, ResNet50

I. INTRODUCTION

Conventional Classification approaches have accomplished extensive achievement in many areas [1]. Though, there are some constraints for these methods under this learning model. Sufficient labeled training examples are required for each class. The trained model can only categorize the examples of categories covered by the training data and cannot recognize unseen types. Nevertheless, it is not always possible to have labeled categories in practical scenarios. The lack of labeled data is due to large data sets which are expensive to annotate and due to rare data classes.

Zero-shot learning (ZSL) overcomes this problem by identifying unseen categories of the test set absent in the training set. ZSL is based on learning some transitional features from training data, and learned features are used to identify unknown categories during testing. To learn intermediate features, each class uses semantic features in a set of attributes, word embedding of class labels or sentences using natural language processing. Attributes are characteristics of categories that describe the features of the classes. Labels or text are used to describe the class semantically. This semantic information concatenated with Visual elements recognize unseen classes (zero-shot classes).

This paper proposed a method to identify unknown categories by extracting visual features with deep neural networks and semantic features (word embedding) using word2vec of training data. The model gives the top 5 predictions of the most suitable labels for unknown categories.

II. LITERATURE REVIEW

The main objective of ZSL is to discriminate unknown categories with the association between known and unknown categories. ZSL employs certain kinds of semantic information in the form of attributes and word vectors. Using semantic output code, ZSL was first proposed in [1] to predict the classes excluded in the training set. ZSL conquers the problem of classification of unlabeled data. In 2009 Lampert [2] et al. proposed attribute-based classification for unseen classes. Attributes are manually made features for groups of classes. Attributes like feather type, body structure, animal habitat etc., are used as auxiliary information for unseen categories that are not available in training. Attribute-based zero-shot image classification was introduced in [3]. This method uses direct attribute prediction (DAP) and indirect attribute prediction (IAP), which are probabilistic classifiers. Attribute label embedding [ALE] suggested in [4] works better than DAP. This mechanism introduces a compatibility function between the image and the label based on which model is evaluated. This method can work with any side information encoded in a vector. Frome [5] has developed deep visual semantic embedding (DeViSE), which extracts visual features with a convolutional neural network and semantic features using a skip-gram language model. The trained model is checked for its prediction using the Hinge loss function. Cross-model Transfer (CMT) [6] is a ZSL model based on a mixture of seen and unseen classes. This method is not based on attributes; it maps Visual features with semantic word vectors of comparable class. Zero-Shot Learning by the Convex Combination of Semantic Embeddings (ConSE) method [7] maps the image to its label embedding using the convex method. The model can be used without additional training.

Zero-shot classification by understanding the compatibility of Input embedding and output embedding is given in [8]. Input Embeddings used Fisher vectors (FV) and Deep CNN Features (CNN). The author discovers combinations of five types of output embeddings supervised

Jain College of Engineering
Belagavi, India

IEEE
BANGALORE SECTION

4th International Conference of Emerging Technology (INCET 2023)

26th – 28th May 2023

Certificate

This is to certify that Dr./Prof./Mr./Ms. ***Ansari Shaista Khanam*** *has presented paper entitled* ***Enhanced Zero shot learning using deep neural network ResNet50*** *in 4th International Conference of Emerging Technology (INCET 2023) during 26th to 28th May 2023.*

VIDYAVARDHINI'S COLLEGE OF ENGINEERING & TECHNOLOGY
EXTC Engg.

HEAD
Dept. of Electronics and Telecommunication Engg.
Vidyavardhini's College of Engineering & Technology
Vasai Road 401202

Dr. Krupa Rasane
Convener - INCET 2023

Dr. J. Shivakumar
General Chair - INCET 2023