SHRI G. S. INSTITUTE OF TECHNOLOGY AND SCIENCE, INDORE

Fix Match : Semi Supervised Learning

Author:
HARSHIT TRIPATHI
JEETENDRA SHARMA
NAMAN MALPANI

Supervisor:
DR. NEERAJ K. RATHORE
MR. UPENDRA SINGH

A PROJECT REPORT (Phase-I)

submitted to ***Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal***

In partial fulfillment of the requirements

for the degree of ***Bachelor of Engineering***

in the

INFORMATION TECHNOLOGY

May 1, 2021

Declaration

We, HARSHIT TRIPATHI, JEETENDRA SHARMA, NAMAN MALPANI declare that this project report titled, "Fix Match : Semi Supervised Learning" and the work presented in it are our own. We confirm that:

- This work was done wholly while in candidature for a bachelor degree at this University.
- Where any part of this project has previously not been submitted for a degree or any other qualification at this University or any other institution.
- Where we have consulted the published work of others, this is always clearly attributed.
- Where we have quoted from the work of others, the source is always given. With the exception of such quotations, this project is entirely our own work.
- We have acknowledged all main sources of help.

Signed: H Tripathi JSharma Naman Malpani

Date: **31/03/2021**

Acknowledgements

We would like to express our special thanks and gratitude to our Head of Director Dr. Sunita Verma, who gave us the golden opportunity to submit this report on **Fix-Match : Semi Supervised Learning** which also helped us in doing valuable research and draw insights about the existing technology.

Also,we would like to thank our mentor Dr. Neeraj K. Rathore , who guided us during the completion of the entire project work.We are really thankful to them. We would also like to thank our parents and friends who helped us a lot in finalizing this report within the limited time frame....

Contents

List of Figures

List of Tables

List of Abbreviations

SSL	**S**emi **S**upervised **L**earning
CIFAR	**C**anadian **I**nstitute **F**or **A**dvanced **R**esearch
VGG	**V**isual **G**eometry **G**roup
ResNet	**Res**idual **N**etworks

List of Symbols

τ	confidence threshold
μ	unlabeled data ratio
λ_u	unlabeled loss weight

Chapter 1

Introduction

Semi-supervised learning (SSL) [6] provides an effective mens of leveraing unlabeled data to improve a model's performance. Here in the paper, we demonstrate the power of a simple combination of two common SSL [6] methods: consistency regularization [2] and pseudo-labelling. We have use the concept of FixMatch algorthim [8] and try to study the accuracy, loss of the following convolution model such as VGG, Googlenet, etc. To get the better accuracy with minimal labelled samples and utilizes the resources of unlabelled samples which are being provided to us.

1.1 Objective

Deep neural networks have become the most successful way in the field of computer vision application. Deep neural networks often overrule the the tradition methods of supervised learning such as decision tree, apriori, etc, which requires the labeled data set to generalize them based on the distinct features of inter-class and common features of intra-class. The performance of the neural network generally depends on the vast variety of data set which is being used but for accumulating such a vast variety of data with the accurate labelled data often requires a significant man-power, which will increase the cost of accumulating the data and used for training the model such as in medical application of detection of tumor.

Therefore the Semi Supervised Learning act as a bridge for producing the artificial pseudo label of unlabelled images based on the observation of labelled data set and accumulating the knowledge of the unlabelled data set to enlarge the cluster of la bled data based on the similarity on in the image shape, size, patter, etc features which will act as input for the the training model to further generalizing the above features of images and to categorizes them based on that. For above to achieve we will use the two most popular techniques to pseudo-labelling which uses the model's class to predict the label to train against which is beyond the threshold. Similarly, consistency regularization obtains an artificial label using the model's predicted distribution after randomly modifying the input or model function, and also stops the model from over fitting the leaning curve to concerted on generally patter and not on precise patterns

1.2 Scope

SSL mitigates the requirement for labeled data by providing a means of leveraging unlabeled data. Since unlabeled data can often be obtained with minimal human labor, any performance boost conferred by SSL often comes with low cost. This has led to a plethora of SSL methods that are designed for deep networks.

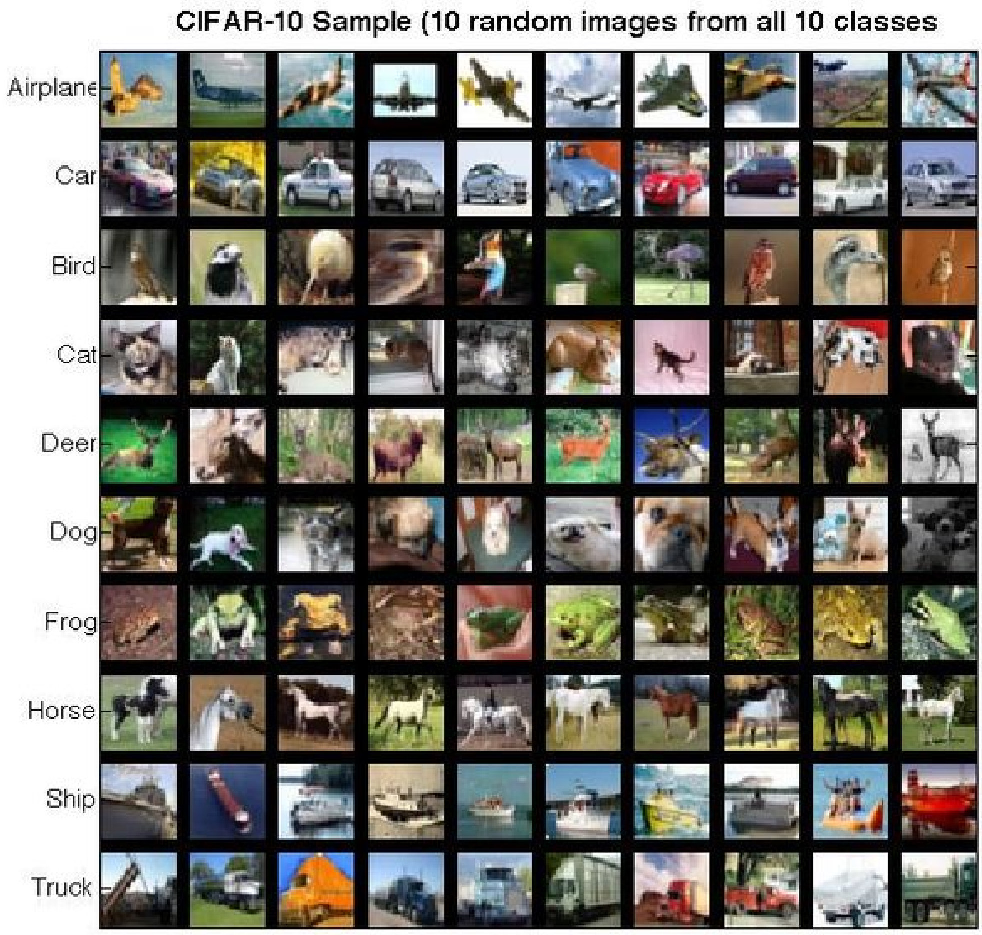

FIGURE 1.1: Images in CIFAR-10 [3] dataset.

FixMatch [8] works on the principle and can be used in computer vision tasks where the size of labelled data is insufficient to train to high accuracy and the cost of labelling the data is also very high.

1.3 Problems in existing system

The original paper **FixMatch: Simplifying Semi-Supervised Learning** [8] with Consistency and Confidence uses ResNet-28-2 with 1.5M parameters for CIFAR-10 [3] only. In our approach we tried using multiple models such as VGG-16,Google net, etc and performed comparative study on them to find which model works best on the dataset and why.

1.4 CIFAR-10 dataset

CIFAR-10 data set is a assemblage of images that are there in our day-to-day life such as car,ships,etc. This data set is commonly used to train the convolution model which is being proposed as set as a bench mark for the comparative study of the various model on the same data set. The CIFAR-10 data set contains approximate an 60,000 color images with the pixel of each image in width and height is 32*32 and distributed with the shades of RGB color palette from values ranging from 0-255. Here the data set contains the 6,000 images per each class.

We will use to evaluate the standard dataset of CIFAR 10 [3] with the different convolution models as well as different confidence threshold factors whose details are being provided in the chapters to follow.

Chapter 2

Literature Survey

Here we have came across the different semi-supervised approach which consist of both machine learning aspects of supervised as well as unsupervised learning method to use the unlabelled samples of the images to have a information gather out from their prediction with the closed to the labelled class. It is very useful for having the better knowledge from the data which we have for reducing the cost of data collection (generally in tumor detection in medical science we need to consult to 2-3 doctors for the accurate labeled to the image which is much more costly).

2.1 Semi-Supervised Learning

Semi-supervised learning is approach to machine learning that combines a small fraction of labeled data with a large fraction of unlabeled data during training. Semi-supervised learning is a bridge between unsupervised learning (without labeled training data) and supervised learning (with only labeled training data).

Unlabeled data, when used with the small data set of labeled data, can produce improvement in learning accuracy with lower cost.

2.2 Different Semi-Supervised Approaches

Over the period of time, different tactics have been used to tackle the problem of labelling the unlabeled to gain more data for training the model and improving its accuracy.

2.2.1 Self Training

The basic ideology behind using unlabeled data in classification is making machine self-learning to find the similarity between labelled data with unlabelled data to enlarge the data set for supervised learning with a small labeled data set , which is also known as self-training [6], self-labeling, or decision-directed learning.

This wrapper-algorithm that used the iteration of supervised learning method. It starts by training on the labeled data only. At after each epoch a part of the unlabeled point is labeled points is labeled according the current decision function, then the supervised method is retrained using its own prediction as pseudo label (additional label).

2.2.2 Co-training

Co-training [5] is a semi-supervised learning technique that requires two the different views of the data in which a first part contains the labelled data set and other

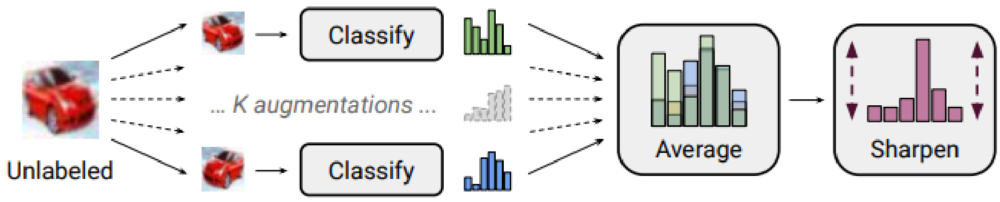

FIGURE 2.1: MixMatch Algorithm [1]

contains the unlabelled data set and both the distinct features set provides the complementary or related information about the instances or the label of the class which are present the the first part of the data. Co-training first uses the labelled part of data to separate classifier for each label and then using the general principle and with predicting the class of unlabelled data set which is most similar to the labelled data set will be included in each iteration to form the cluster of images of the same class which could be further train to evaluate the other unlabelled images.

2.2.3 Graph-based SSL

Graph-based SSL [9] methods can be classified in the three separate steps: graph creation, graph weighting and inference.

Here the graph construction phase comprises of the following two steps of the graph based SSL that are the graph creation, graph weighting where the graph creation deals with the connecting the nodes of graph with each other and the graph weighting deals with assigning the weights to the edge between the nodes based on their similarity or relation between the two nodes.

After the graph is constructed we will have an adjacency matrix of n*n where n is the number of nodes and and each cell of matrix is assigned a weight based on the the weight assigned to that two nodes set.

When the graph is constructed it is used to predict the unlabelled data points based on the similarity of the data with the weights assigned to each cell. The general objective functions for transductive graph-based methods contains one component which is responsible for penalizing predicted labels or matching the diversity between the images and other components for matching the similarities or the correlation between the labels prediction for connected data points.

2.3 Recent Development in SSL

In recent years, the amount of unlabelled data has increased ten-folds so has increased the research for the utilisation of this data. Many new SSL algorithms have shown comparable result to that of Supervised learning.

2.3.1 MixMatch

MixMatch [1] The main idea behind the use of MixMatch is based on the fact that if even slightly modify the image of the labelled samples than the prediction of the same would also resembles the same labelled class because the intra-class similarity will exceed the inter-class distances. Here it use both unlabelled as well as labelled

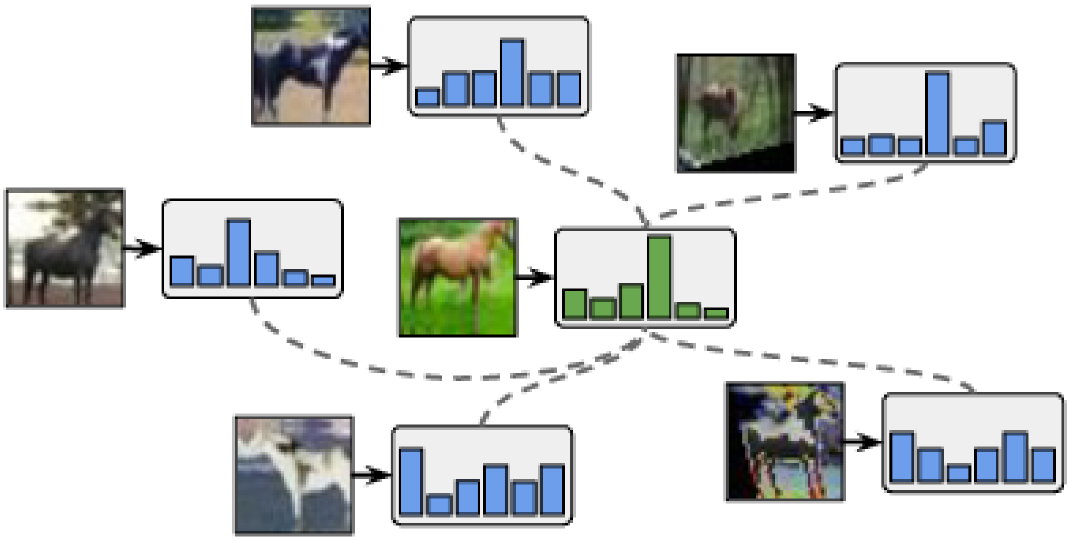

FIGURE 2.2: ReMixMatch Algorithm [2]

samples in training, at each iteration it will assign the pseudo label to the unlabelled class and if it exceeds over the beyond thresholds means the intra-class similarity is maximized and falls in the same category of the respective class. Hence therefore the total loss term for unlabeled data that reduces the entropy measure of the disorder or uncertainty while maintaining the consistency to overcome the over-fitting and compatible with traditional regularization techniques.

2.3.2 ReMixMatch

ReMixMatch [2] it is an enhanced improvised version of "MixMatch" algorithms here we consider the two techniques as distribution alignments which ensures the the distribution of the prediction unlabeled data need to be close to the ground truth of labels. Aug-mention anchoring is used to provide the multiple augmentation of the image vector based on transformation of image matrix, projecting in different plane with an angle, modifying the saturation level of image, many other ways of the input to the prescribed model and encourages that the output of each need to close to the predicted label for the input image or the small manipulation or weak augment the image vector, because both augmentation of the same image always falls in same category which is being used to provide the general cluster for the image based on same principle irrespective of the image orientation being considered.

2.3.3 FixMatch

Fix Match [8], an algorithm that is significant simplification of existing Semi-Supervised Learning methods. Fix Match approaches the task as follows, while training the model at each iteration it will generates the pseudo-labels using the model's prediction for the weakly augmented unlabelled image samples. For a given images, if the pseudo-label is similar to the threshold i.e. high-confidence prediction. The model is then trained to predict the pseudo-label when the strong-augmentation such as translation, light, saturation operation perform on image vectors of the same input

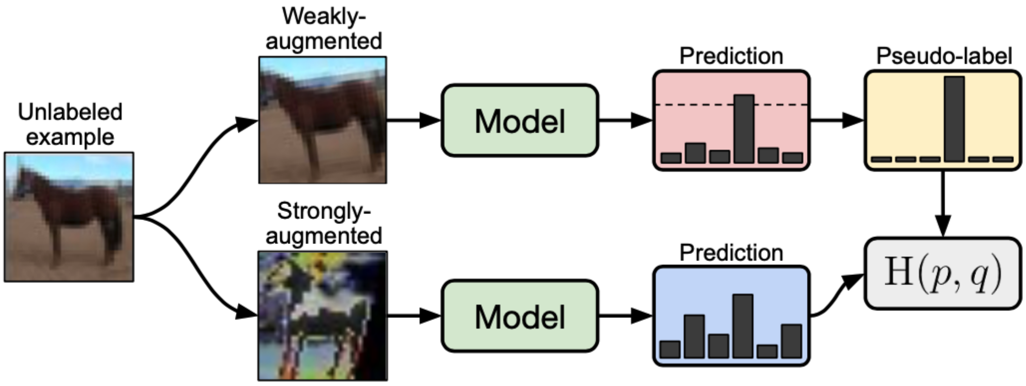

FIGURE 2.3: FixMatch Algorithm [8]

image vector. Despite the simple idea, Fix Match [8] tend to perform across a variety of semi-supervised learning benchmarks.

2.4 Our Proposed Model

In our proposed model, we used FixMatch [8] algorithm on different convolution neural networks architectures and used different augmentation policies to perform a comparative study, how and why different architecture behaved as such.

Here by studying about the following approaches of semi-supervised learning of Mix-Match [1], ReMixMatch [2], Fix Match we have choose from the above to use FixMatch [8] algorithm which has a better accuracy with the CIFAR-10 [3] trained over ResNet [11] so we would like to explore other convolution models with the same prescribed algorithm.

Chapter 3

System Requirement Analysis

We have study the following connvolution model such as AlexNet [12], VGG-16[7], ResNet[11], GoogleNet [10] which have a remarkable accuracy over the CIFAR-10 [3] dataset and we will try to use them for FixMatch [8] algorithm.

3.1 Information Gathering

In the recent years, enormous amounts of information has become available most notably unstructured and semi-structured data available from the internet. In order for this information to be of greater use, more structure needs to be discovered in it.

We began journey by finding the ways to utilise unlabelled data for supervised learning and came to know about Semi-Supervised Learning(SSL). We went through the Self-training, Co-training, MixMatch , ReMixMatch , FixMatch (which has been discussed in section 2.2). After that we studied the following convolutional neural network architecture from their respective research paper.

3.1.1 AlexNet

AlexNet [12] architecture consists of 5 convolutional layers, 3 max-pooling layers, 2 normalization layers, 2 fully connected layers, and 1 softmax layer. Each convolutional layer consists of convolutional filters and a nonlinear activation function ReLU. The pooling layers are used to perform max pooling.

3.1.2 VGG-16

VGG [7] convolution model follows the principle of Alex Net with the variation of decreasing the window sizes and strides after the starting convolution layer, and by increasing the depth of the architecture by reducing the max pooling between the convolution layer and increase the kernel level size drastically from the Alex-Net. In overall it contains 16 layers which contain the trainable parameters and other layers like Max pool layer which will reduce the image vector size when transition from one convolution layer to lower one with increased in kernel size and this layer do not contain any trainable parameter their utility is to reduce the image pixel or image vector

3.1.3 ResNet

Resent [11] has its intuition behind the human cerebral cortex which is the outer layer of neural tissue and to define the convolution layer how the information passes from neuron to neuron in humans. Residual neural network achieve this by using

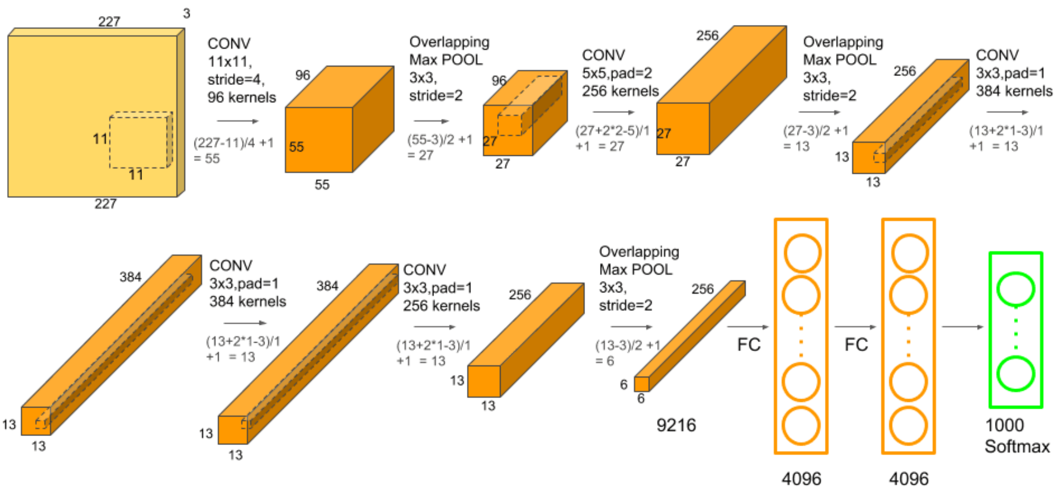

FIGURE 3.1: Different layers in AlexNet [12]

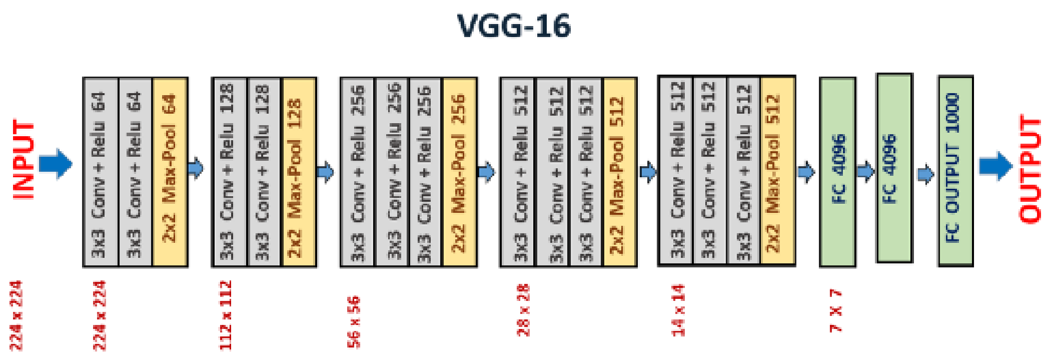

FIGURE 3.2: VGG-16 architecture [7]

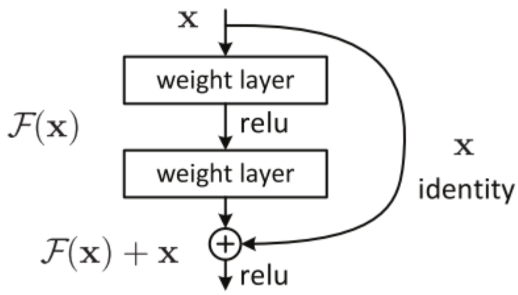

FIGURE 3.3: General Resnet architecture [11]

the skip connections, shortcut jumps between the layers so not only depend on the just previous layers but the two to three back layers and combing their information which will act as the input for the next layer it contains the Rectilinear unit function and batch normalization between layer. In addition with tat it contains an weight matrix which is used to learning the skip weights, they are also knows as Highways Nets. When same model is modified with the parallel skips they are referred to as Dense Nets.

3.2 System Feasibility

3.2.1 Economical

The system being developed is extremely economical as compared to the supervised learning which in some cases requires domain experts to label the data set. It is one of the key feature of this system.

3.2.2 Technical

The project can be developed with the help of machine learning algorithms and libraries on Python. A good laptop is needed to demonstrate and understand the use of this technology.

3.2.3 System Feasibility

The system proposed is very practical and achievable as well as solves a major problem in deep learning field.

3.2.4 Hardware

Currently Laptop is only required hardware, which is necessary for implementing algorithm. We are also using Google cloud computing platform for training the model.

3.2.5 Software implementation language/technology

The project is developed in Python technology using OpenCV, Tensorflow, Pil, Keras, Pandas, Numpy, scipy packages.

3.3 Information Flow Representation

3.3.1 Usecase Diagram

FIGURE 3.4: Usecase digram of the system

3.3.2 Use Cases

The different usecases of our system are as follow:

- **Choose model**: The system trains a set models having different architecture and allows the user to choose any one of them.
- **Data augmentation**: The system uses different augmentation techniques for the data extension and pertubation.
- **Label dataset**: The main aim of the system is to label the unlabelled dataset which may be used in supervised learning algorithms.
- **Predict label**: The system can also be used for predicting machine learning model.

The system design of our system is very simple as it is just a application of unsupervised learning algorithm on different model architectures and there are no flow of data between different classes and any database involved.

Chapter 4

System Design

In our project, we are performing comparative study SSL algorithm due to which the system design is very simple. It is consist of two parts:

- Different deep learning models.
- FixMatch algorithm.
  FixMatch is a combination of two approaches to SSL: Consistency regularization and pseudo-labeling. Its main novelty comes from the combination of these two ingredients as well as the use of a separate weak and strong augmentation when performing consistency regularization.

4.1 Architectural Design

Architectural Design of the project follows a simple linear path with a conditional branch for the case of labelled and unlabelled data. Each case having its own categorical cross-entropy loss.

4.1.1 Description of Architectural Design

1. First we have considered the image dataset of CIFAR-10 [3] with 10 different classes and with the input image of (32*32*3).
2. With the image data generator we will manipulate the image and enlarge the dataset with the slight transformation with the image and its vector to different operation which are same as mentioned in Random-Augmentation.
3. Now we will train the model with the different Convolution Models such as GoogLeNet , ResNet, VGG, etc.
4. Now at each epoch of training the dataset we will do the unlabelled dataset with the weak augmentation and strong augmentation of image and by combining the loss of the both and will minimize.

4.1.2 Internal Data Structures

A Tensor is a generalization of vectors and matrices to potentially higher dimensions. Internally, TensorFlow represents tensors as n-dimensional arrays of base datatypes. Each element in the Tensor has the same data type, and the data type is always known. The shape (that is, the number of dimensions it has and the size of each dimension) might be only partially known.

In our system , the images are converted in 3 dimensional array using numpy which are further converted in tensors and are passed through the model to predict the label.

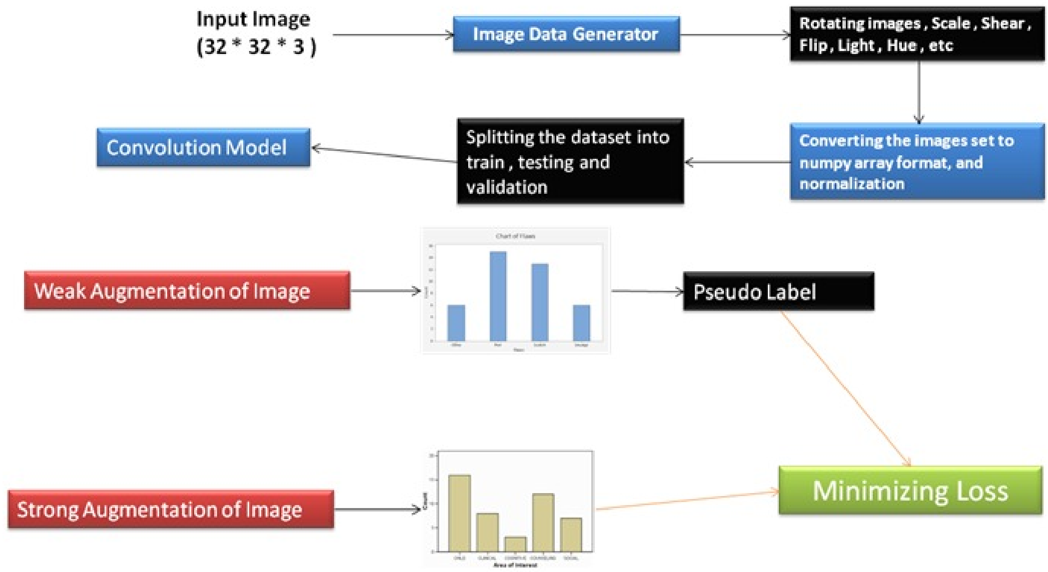

FIGURE 4.1: Architecture Design

4.1.3 Algorithm design

We present the complete algorithm for FixMatch:

Algorithm 1: FixMatch algorithm [8]

Input: Labeled batch
$\chi = \{(x_b, p_b) : b \in (1, ..., B)\}$, *unlabeled batch* $U = \{u_b : b \in (1 ..., \mu B)\}$, *confidence threshold* τ, *unlabeled data ratio* μ, *unlabeled lossweight* λ_u

$l_s = 1/B \sum_{b=1}^{B} H(p_b, (x_b))\{Cross - entropy\ loss\ for\ labeled\ data\}$

for b = 1 **to** μB **do:**

$q_b = p_m(y|\alpha(u_b); \theta)$

{ Compute prediction after applying weak data augmentation of u_b}

end for

$l_u = 1/_u B \sum_{b=1}^{uB} 1\{max(q_b) > \top\}\ H(argmax(q_b), pm(y|A(u_b))$

{Cross-entropy with psudo label and confidence for unlabeled data }

return $l_s +_u l_u$

The algorithm takes labelled and unlabelled images from the CIFAR-10 [3] dataset. For labelled images, it perform supervised learning on them with categorical-cross-entropy as loss function. For unlabelled images, the images are augmented with weak and strong augmentation functions. Prediction is performed on weakly augmented images , if model is confident above the confidence threshold then those prediction is treated as pseudo-labels and the categorical-cross-entropy loss is calculated on same strongly augmented images with pseudo-labels as correct labels. Finally the total loss is the sum of unsupervised and supervised loss.

4.2 Data Design

The CIFAR-10 [3] dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

4.2.1 Data collection and preprocessing

From the CIFAR-10 [3] dataset , we will be taking 250 images of each class as labelled images from the training set and other images from the training set as unlabelled images. Further these 250 will be augmented to drive more data for training. After that all the images will be converted to numpy array and normalised. Finally they will be fed to the algorithm where they will be weakly augmented as well as will be strongly augmented as per FixMatch[8] algorithm.

4.2.2 Data objects and resultant data structures

After the algorithm has successfully run, we get two data objects or outputs:

1. Labels of the unlabelled dataset.
2. Trained models which can be used to for prediction.

The model obtained after FixMatch[8] algorithm training has better accuracy as compared for supervised learning with same amount of data.

Apart of this, from our study we will get which model performed better on Cifar-10 [3] dataset and why. we will also try to find the impact of different augmentation of images on the prediction.

Chapter 5

Results

5.1 Accuracy on various Convolution Model

Here we have performed the FixMatch[8] algorithm on the various supervised standard model as well as on their modified architectures. We have altered the value of confidence threshold to understand its effect and we also changed the ratio of unlabelled dataset to labelled dataset to analyse its effect.

After the training of models following outcomes were achieved and are shown in below-given format with varying threshold and convolution model.

Model	Epoch	Threshold	Optimizer	U/L Ratio	Training Loss	Test Loss	Accuracy
VGG 16	250	0.95	Adam	0.1	2.1123908	1.921999	0.7402
VGG 16	250	0.95	Adam	0.2	2.141766	1.953861	0.739
VGG 16	250	0.95	Adam	0.3	2.145502	2.293342	0.7376
VGG 16	250	0.95	Adam	0.5	2.1516054	1.656397	0.7261
VGG 16	250	0.9	Adam	0.1	2.1712048	1.628632	0.7435
VGG 16	250	0.9	Adam	0.2	2.1887937	2.430666	0.7419
VGG 16	250	0.9	Adam	0.3	2.113949	1.954267	0.7345
VGG 16	200	0.9	Adam	0.5	2.1779883	2.012911	0.7317
VGG 16	250	0.8	Adam	0.1	2.225982	1.920046	0.7391
VGG 16	250	0.8	Adam	0.2	2.198515	1.86853	0.7309
VGG 16	250	0.8	Adam	0.3	2.1702876	1.968906	0.729
VGG 16	250	0.9	Adam	0.5	2.1887937	2.430666	0.7027

TABLE 5.1: Result of VGG-16 Models

Above shows the result of training VGG-16[7] model at the end 250 epochs.The model was trained for 3 different values of confidence threshold (0.95,0.9,0.8) as well as 4 different values of unlabelled and labelled dataset ratio.

Model	Epoch	Threshold	Optimizer	U/L Ratio	Training Loss	Test Loss	Accuracy
Resnet - 50	150	0.95	Adam	0.1	2.1312473	2.2012012	0.593
Resnet - 50	150	0.95	Adam	0.2	2.3550835	1.9903095	0.5807
Resnet - 50	150	0.95	Adam	0.3	2.28239	2.1253781	0.5904
Resnet - 50	150	0.95	Adam	0.4	2.2836332	1.1810992	0.5852
Resnet - 50	150	0.95	Adam	0.5	2.4420755	1.5309395	0.5821
Resnet-18	150	0.95	Adam	0.1	2.307897	2.0707574	0.668
Resnet-18	150	0.95	Adam	0.2	2.2099125	2.1152635	0.6241
Resnet-18	150	0.95	Adam	0.3	2.2539325	2.1534734	0.6314
Resnet-18	150	0.95	Adam	0.4	2.2322989	2.594476	0.6256
Resnet-18	150	0.95	Adam	0.5	2.2083755	2.5991275	0.6183

TABLE 5.2: Result of ResNet Models

Above shows the result of training Resnet[11] models(50 layers, 18 layers) at the end 150 epochs. The models was trained for 5 different values of unlabelled dataset and labelled dataset ratio.

Model	Epoch	Threshold	Optimizer	U/L Ratio	Training Loss	Test Loss	Accuracy
VGG Mod.	150	0.95	Adam	0.1	2.4417548	0.5268971	0.8236
VGG Mod.	150	0.95	Adam	0.2	2.3434238	0.56225884	0.8128
VGG Mod.	150	0.95	Adam	0.3	2.1886437	0.55986637	0.8321
VGG Mod.	150	0.95	Adam	0.4	2.2964838	0.6137788	0.8237
VGG Mod.	150	0.95	Adam	0.5	2.2138987	0.7720934	0.8131
VGG Mod.	150	0.95	Adam	0.6	2.259274	0.8634428	0.8005
VGG Mod.	150	0.95	Adam	0.7	2.2356193	0.96561986	0.7817
VGG Mod.	150	0.95	Adam	0.8	2.2380457	1.080498	0.7529
VGG Mod.	150	0.95	Adam	0.9	2.3302784	1.301985	0.6914
VGG Mod.	150	0.95	Adam	0.95	2.2321618	2.0022128	0.6203
VGG Mod.	150	0.9	Adam	0.1	2.439519	0.6776542	0.8012
VGG Mod.	150	0.9	Adam	0.2	2.452824	0.55453026	0.8162
VGG Mod.	150	0.9	Adam	0.3	2.3387198	0.5732669	0.8316
VGG Mod.	150	0.9	Adam	0.4	2.191008	0.6723652	0.8257
VGG Mod.	150	0.9	Adam	0.5	2.2786562	0.64873946	0.8289
VGG Mod.	150	0.9	Adam	0.6	2.1820295	0.81414205	0.8079
VGG Mod.	150	0.9	Adam	0.7	2.2066555	0.9345428	0.7914
VGG Mod.	150	0.9	Adam	0.8	2.2611156	1.1566179	0.748
VGG Mod.	150	0.9	Adam	0.9	2.1567755	1.605863	0.6973

TABLE 5.3: Result of Modified VGG-16 Models

Above shows the result of training Modified VGG-16[7] model at the end 150 epochs. The model was trained for 1 different values of confidence threshold (0.95,0.9) as well as 9 different values of unlabelled and labelled dataset ratio.

5.2 Result Analysis

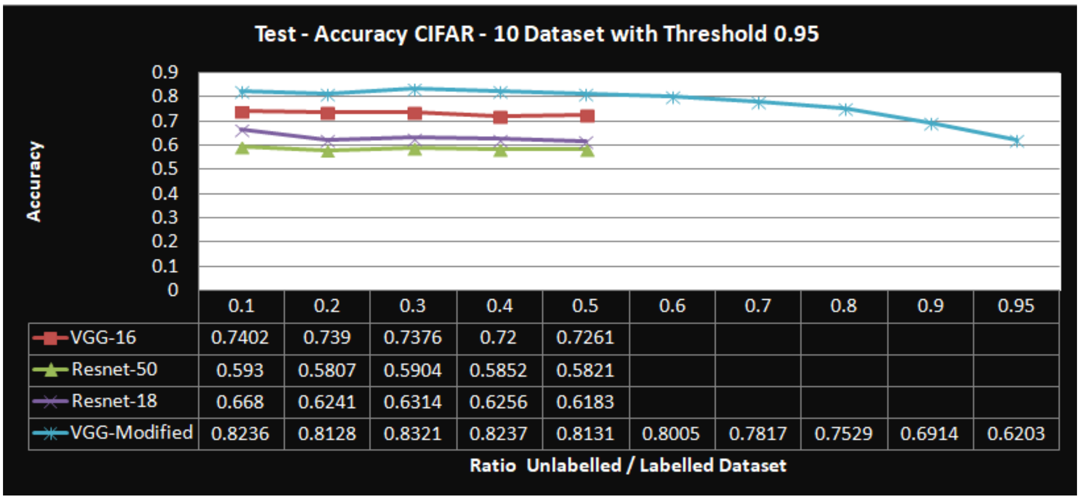

FIGURE 5.1: Unlabellled/Labelled Ratio

As we increase the threshold value from range 0.1 to 0.9 we have observed that there is a decrement in accuracy of the model but the deviation is quite minimal till the Unlabelled/Labelled Ratio of dataset is 0.8 which is nearly by 2% , afterwards the decrement is considerably more. As the shallower network of VGG - 16[7] and its modified version with the dropout layer is outperformed from the deeper Convolution Model such as Resnet - 50[11] and Resnet - 18[11] which is due the over-fitting in the graph for deeper model as compared to VGG-16[7] which can be concluded from the above graph.

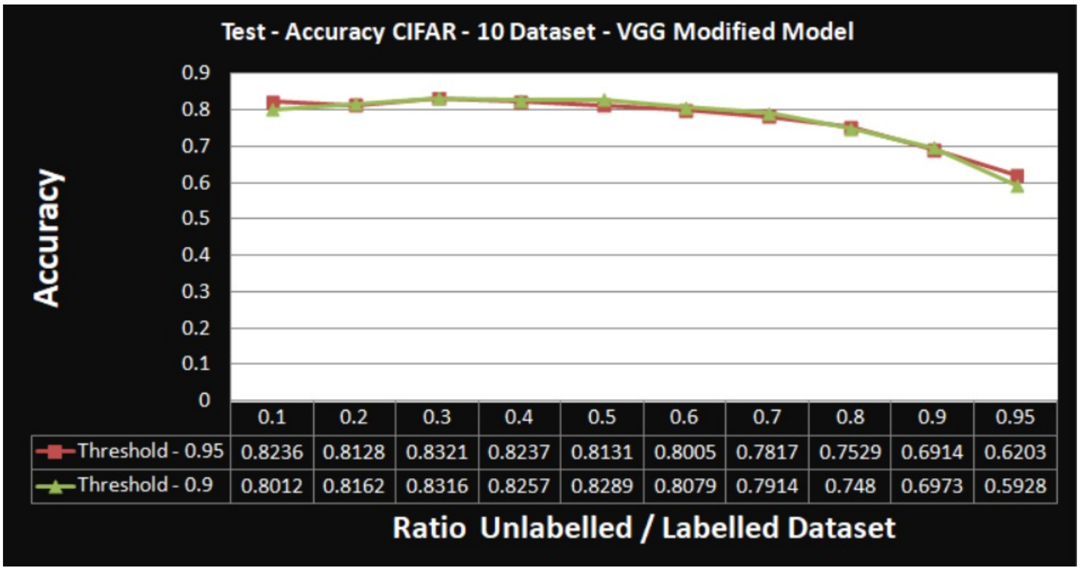

FIGURE 5.2: VGG Modified with varying Thresold

Since from the above model VGG-modified (with addition dropout layer with the original flow of tradition VGG - 16 and some layer addition to it) we have achieve the remarkable accuracy now since we like to depict the accuracy by varying the threshold from 0.95 to 0.9 we have observed that with threshold 0.9 it has a better accuracy, by consider all the other parameter such that optimizer, epochs to be same for both the results.

After thorough analysis of all the result of the several models,we realised that our implementation showed similar trend as shown in the original paper of FixMatch[8].

Chapter 6

Conclusion and Future Scope

The main aim of our project was to implement the research paper FixMatch[8] and observe the effect of different model architecture on FixMatch[8] algorithm implemented model. We also varied the confidence threshold and unlabelled/labelled dataset ratio and analysed its effect on FixMatch[8] algorithm.

6.1 Conclusion

Here we have come to the outcome that as we go to deeper model for training we have observed that there is over fitting in the graph and thus we have tried to go for the shallower convolution model which have a great outcome. Here we have observed that when we are having just the 10% of the training dataset as compared for the supervised model we are able to achieve greate accuracy by 69%. Here we have also able to achieve the accuracy with very less deviation from the accuracy based on the supervised learning model which is 30% of the total training dataset and a decrement in just 2% of the accuracy, which is a great achievement for making the use of unlabeled dataset to extract information from them.

6.2 Future Scope

Since it has a great achievment in field of machine learning for using the unlabelled images all around the internet to use them for training the machine using fix-match alogrithm since this will reduce the cost for making the large dataset for the training purpose for supervised learning and here the semi-supervised learning[4] has an edge for get the remarkable accuracy with least deviation and with large number of unlabel data with some label data.

Our implementation of FixMatch[8] algorithm has shown same trends in case confidence threshold and unlabelled ,labelled data ratio. In our implementation of algorithm shallow models performed better than deeper models.

References

[1] David Berthelot et al. "Mixmatch: A holistic approach to semi-supervised learning". In: *Advances in Neural Information Processing Systems*. 2019, pp. 5049–5059.

[2] David Berthelot et al. "Remixmatch: Semi-supervised learning with distribution matching and augmentation anchoring". In: *International Conference on Learning Representations*. 2019.

[3] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. "CIFAR-10 (Canadian Institute for Advanced Research)". In: (). URL: `http://www.cs.toronto.edu/~kriz/cifar.html`.

[4] Hugo Larochelle et al. "Exploring strategies for training deep neural networks." In: *Journal of machine learning research* 10.1 (2009).

[5] Siyuan Qiao et al. "Deep co-training for semi-supervised image recognition". In: *Proceedings of the european conference on computer vision (eccv)*. 2018, pp. 135–152.

[6] Y Reddy, Viswanath Pulabaigari, and Eswara B. "Semi-supervised learning: a brief review". In: *International Journal of Engineering Technology* 7 (Feb. 2018), p. 81. DOI: `10.14419/ijet.v7i1.8.9977`.

[7] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv:1409.1556* (2014).

[8] Kihyuk Sohn et al. "Fixmatch: Simplifying semi-supervised learning with consistency and confidence". In: *arXiv preprint arXiv:2001.07685* (2020).

[9] Amarnag Subramanya and Partha Pratim Talukdar. "Graph-based semi-supervised learning". In: *Synthesis Lectures on Artificial Intelligence and Machine Learning* 8.4 (2014), pp. 1–125.

[10] Christian Szegedy et al. "Going deeper with convolutions". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9.

[11] Sasha Targ, Diogo Almeida, and Kevin Lyman. "Resnet in resnet: Generalizing residual architectures". In: *arXiv preprint arXiv:1603.08029* (2016).

[12] Zheng-Wu Yuan and Jun Zhang. "Feature extraction and image retrieval based on AlexNet". In: *Eighth International Conference on Digital Image Processing (ICDIP 2016)*. Vol. 10033. International Society for Optics and Photonics. 2016, 100330E.