8/6/2023

FINAL PROJECT

AASD4015 ADVANCED MATHEMATICAL CONCEPTS FOR DEEP LEARNING

Project Title

Sports Image Classification using Pytorch.

Group9

Name	Enrollment
Prakhar Patel	101413720
Nihitha Patcha	101446620
Archana Wasti Dahal	101465212
Vijay Karthik Bethapudi	101442692

Ran Feldes
PROFESSOR

TABLE OF CONTENTS

1. ABSTRACT

In this study, we apply deep learning methods to the job of classifying sports images. Our goal is to correctly classify a wide range of sports activities that are seen in photographs. We use transfer learning to accomplish this goal by perfecting three well-known pre-trained models, ResNet-50, ResNet-152, and ResNet-101. Various sports-related photos are included in the dataset, which has been painstakingly divided into training, validation, and test sets.

In our method, we strategically fine-tune the layers of the pre-trained models to properly capture sports-specific data. We do a complete performance study, analyzing accuracy and loss data throughout the training and validation phases. A specific test dataset is used to rigorously evaluate the models, and visual comparisons between predicted labels and ground truth labels are provided.

2. INTRODUCTION

2.1 Problem Description and Understanding: In this study, we tackle the challenging task of sports image classification using advanced deep learning techniques. Our objective is to accurately classify a wide range of sports-related activities depicted in images. This task holds significant importance in sports analysis, content categorization, and event recognition.

2.2 Importance: The ability to classify sports images has numerous applications, from enhancing sports analytics to enabling automated content tagging. We have curated a diverse dataset and employed sophisticated deep learning models to achieve accurate sports image classification.

2.3 Overview of Results: Our results demonstrate the effectiveness of our approach. By employing state-of-the-art deep learning models and fine-tuning them, we achieved improved accuracy in sports image classification. The models successfully capture distinctive features of different sports, enhancing their ability to make accurate predictions. We provide detailed insights into our methodology, present our experimental results, and discuss the implications of our findings for advancing sports image analysis.

AFFILIATED GROUNDWORK

In the realm of sports image classification, several researchers have undertaken the challenge of accurately categorizing sports-related activities depicted in images. These endeavors have harnessed intelligent computer algorithms and similar techniques, often leveraging models such as MobileNetV2, to advance the field.

However, what distinguishes our work is our innovative approach to model refinement. Our emphasis is squarely placed on the meticulous process of fine-tuning. We channel our efforts into enhancing the model's capability to accurately classify diverse sports activities by finely adjusting specific components.

By carving out this unique trajectory, we make a meaningful contribution to the collective body of knowledge in sports image classification, presenting a novel perspective on enhancing deep learning models for this captivating domain.

3. ABOUT DATA

The dataset used for our sports image classification project comprises a collection of 100 sports-related images, sourced from the Kaggle platform. This curated dataset encompasses a diverse range of sports activities, capturing various sports disciplines and scenarios. Each image has been meticulously selected to represent different aspects of sports, ensuring a comprehensive coverage of the sports domain.

Source: 100 Sports Image Classification
Total Images: Around 14500

Collection of sports images covering 100 different sports. Images are 224,224,3 jpg format. Data is separated into train, test, and valid directories. Additionally, a csv file is included for those that wish to use it to create their own train, test, and validation datasets.

4. METHODOLOGY

4.1 Define Image Dataset

The Image Dataset class is a custom dataset implementation for handling image data in PyTorch. It is designed to work with a specific directory structure where each sub-directory corresponds to a distinct class, and the images belonging to that class are stored within that sub-directory. This structure is commonly used in image classification datasets.

Overall, the Image Dataset class simplifies the process of working with image datasets in PyTorch, providing a flexible and robust solution for handling image data in machine learning projects.

Figure 5.1.1 Images from Dataset

4.2 Train, Test and Validate

The dataset is partitioned into three distinct subsets: training, validation, and testing. This partitioning is pivotal to ensure the integrity of our model's evaluation and generalization capabilities.

Training Set: This subset constitutes the largest portion of the dataset and serves as the foundation for training our deep learning models. It consists of labeled images, each associated with a specific sports category. The training set facilitates the model's learning process by exposing it to a rich variety of visual features inherent in different sports.

Validation Set: To fine-tune and optimize our models, we employ a validation set. This subset assists in the selection of hyperparameters, monitoring training progress, and mitigating overfitting. The validation set's images are distinct from those in the training set and are only used during the training phase for parameter tuning.

Testing Set: The testing set is crucial for assessing the final performance and generalization ability of our trained models. It contains images that the models have never encountered during training or validation, ensuring an unbiased evaluation of their capabilities.

4.3 Base Model for Transfer Learning and Rescaling Pixel Value

For pretrained model we have gone for ResNet series and aimed to measures their individual performance in same environment try to determine which series of ResNet model is better for our chosen dataset of sports image classification problem. We have selected following three models from ResNet model libraries:

- ResNet50
- ResNet101
- Resnet152

4.4 Define Custom `model_dataloader' for custom dataset.

The `model_dataloader` function is designed to create PyTorch data loaders for training, validation, and testing datasets in an image classification task. It uses a custom dataset, named `ImageDataset`, which is assumed to handle a specific directory structure where each sub-directory represents a distinct class, and the images belonging to that class are stored within the respective sub-directory. This function takes weights and a transformation function as input and returns three PyTorch data loaders for the training, validation, and testing datasets.

4.5 Fetch the Model and Appropriate Weights for Fine Tuning.

In this stage of production, we have stored weights of above-mentioned models' weight and pre-trained models by which we can get appropriate model data loaders. Which can further be used for training pre-trained models on their own defined architecture but on our selected dataset or data source. After that, instead of using base architecture we have done some changes and computing layers where we have added extra different layers like `Linear`, `ReLU`, `Dropout(0.5)` and many more to make model bit complex and make it to understand data deeper for more optimal output and accuracy.

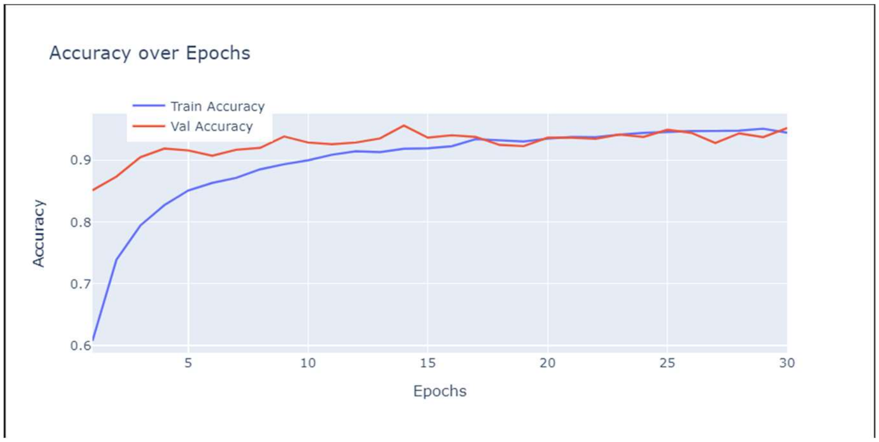
4.6 Fine Tuning or Transfer Learning

We used fine tuning to our model to increase the performance with layers of our based model. For this task first we have unfrozen our top layer from base model then applied weights to the models which were not updated during initial training of our model. We have trained our customized pre-trained model over 30 epochs and 5 patience.

5. RESULTS AT THE END

5.1 ResNet50

Figure 5.1 ResNet50 Accuracy Over Epochs



Training Accuracy and Validation Accuracy:

- The training accuracy starts at around 60.77% in the first epoch and steadily improves with each subsequent epoch.
- The validation accuracy starts at 85.12% and shows consistent improvement over the epochs.
- This indicates that the model is learning and generalizing well as the training progresses. The fact that the validation accuracy is also increasing suggests that the model is not overfitting and is indeed improving its performance on unseen data.

Training Loss and Validation Loss:

- The training loss starts at 1.348 and consistently decreases with each epoch.
- The validation loss starts at 0.581 and consistently decreases over the epochs.
- The decreasing loss values suggest that the model is converging and getting better at minimizing its errors during training. The decreasing validation loss implies that the model is improving its performance on validation data as well.

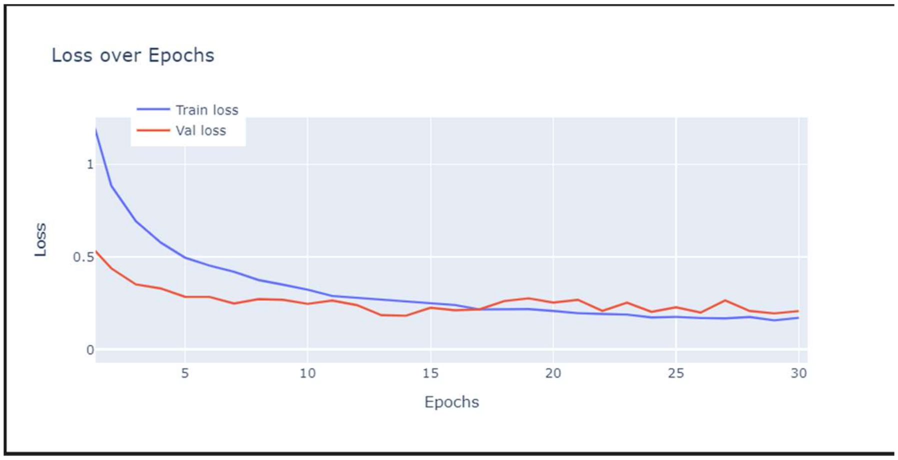


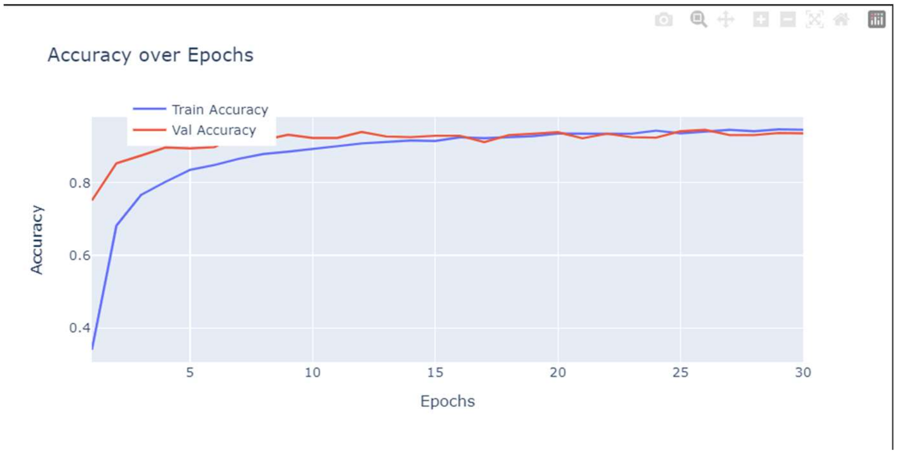
Figure 5.1.1 ResNet50 Loss Over Epochs

Test Accuracy:

After fine tuning we have achieved 95.6% of accuracy on our test dataset

5.2 ResNet101

Figure 5.2 ResNet101 Accuracy Over Epochs



Training Accuracy and Validation Accuracy:

- The training accuracy starts at around 31.94% in the first epoch and steadily improves with each subsequent epoch.
- The validation accuracy starts at 74.06% and shows gradual improvement over the epochs.
- This indicates that the model is learning and making progress in its ability to classify the data correctly. However, it seems the model's overall accuracy is relatively low, and there might be room for further improvement.

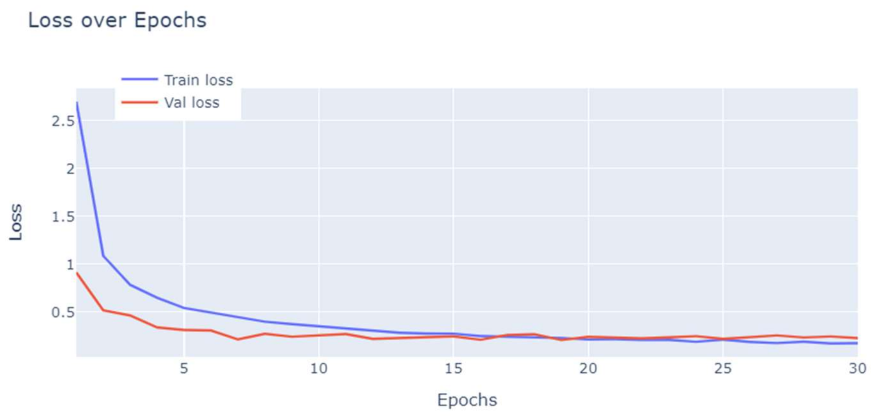


Figure 5.2.1 ResNet101 Loss Over Epochs

<u>Training Loss and Validation Loss:</u>

- The training loss starts at 2.77 and consistently decreases with each epoch.
- The validation loss starts at 0.98 and consistently decreases over the epochs.
- The decreasing loss values suggest that the model is learning to minimize its errors during training. The validation loss decreasing as well indicates that the model is improving its performance on validation data.\

<u>Test Accuracy:</u>

After fine tuning we have achieved 94.0% of accuracy on our test dataset which is close to ResNet50's accuracy.

5.3 ResNet152

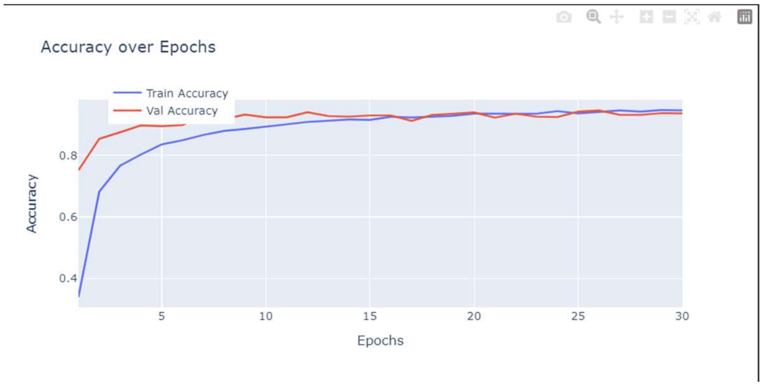


Figure 5.3 ResNet152 Loss Over Epochs

<u>Training Accuracy and Validation Accuracy:</u>

- The training accuracy starts at around 33.99% in the first epoch and consistently improves with each subsequent epoch.
- The validation accuracy starts at 75.16% and shows a gradual improvement over the epochs.

- The model's accuracy has shown steady improvement over the epochs, indicating that the model is learning and adapting to the data. The validation accuracy is increasing as well, suggesting that the model is generalizing well to unseen data.

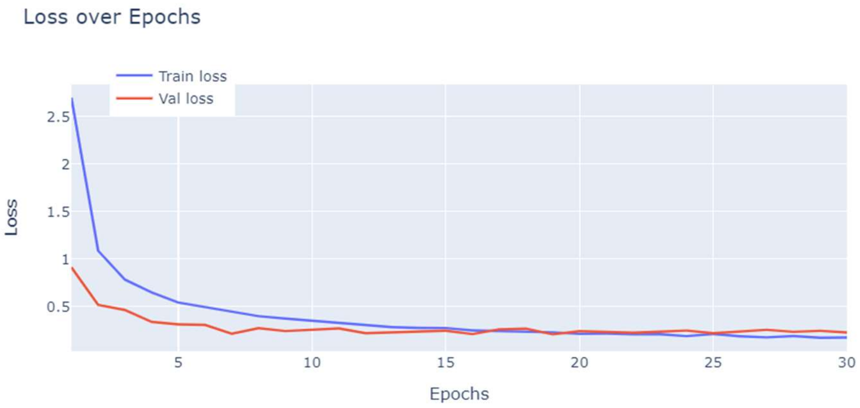


Figure 5.3.1 ResNet152 Loss Over Epochs

Training Loss and Validation Loss:

- The training loss starts at 2.70 and consistently decreases with each epoch.
- The validation loss starts at 0.91 and also consistently decreases over the epochs.
- The decreasing loss values indicate that the model is effectively minimizing its errors during training. The validation loss decreasing in tandem suggests that the model is improving its performance on the validation data as well.

Test Accuracy:

After fine tuning we have achieved 95.2% of accuracy on our test dataset which is close to ResNet50's and ResNet101's accuracy.

6. EXPERIMENTS

To prove that our approach to the problem solve the issue, we conducted an experiment with testing images and real time images. First, we downloaded some images from internet and apply same downloading techniques that we applied to train model with the train data testing set of the images from dataset and the results are discussed earlier.

Overall Model Performance Analysis

Model	Testing Accuracy	Training Time(Minutes)
ResNet50	95.6	37.3207
ResNet152	95.2	65.0694
ResNet101	94	58.9779

Figure 6.1 Model-wise Test Accuracy and Training Time

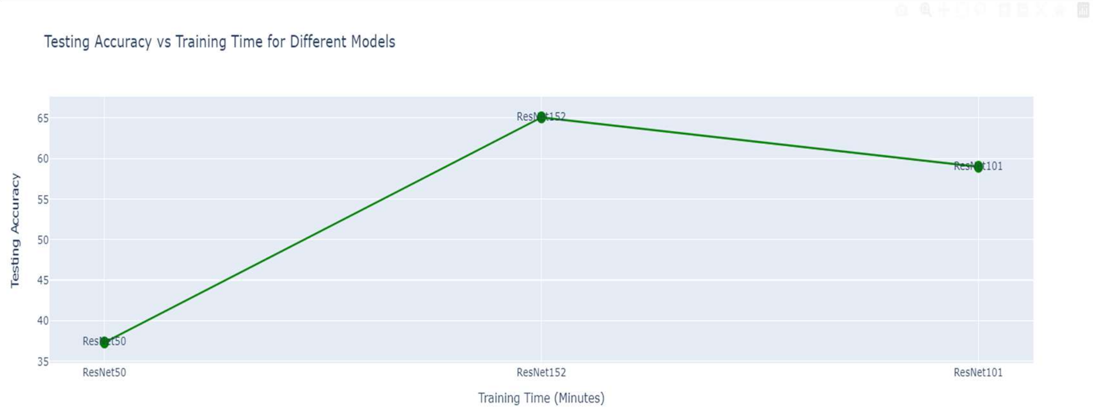


Figure 6.2 Model-wise Test Accuracy

Observation 1: Model Performance

From the provided data on testing accuracy, it is evident that ResNet50 outperforms both ResNet152 and ResNet101 in terms of accuracy on the testing dataset. ResNet50 achieved the highest testing accuracy of 95.6%, followed by ResNet152 with a slightly lower accuracy of 95.2%, and ResNet101 with a comparatively lower accuracy of 94%. This suggests that ResNet50 has a better capability to generalize and make accurate predictions on unseen sports images compared to the other two models.

Observation 2: Training Time

The data on training time reveals that ResNet50 has the shortest training time among the three models, taking approximately 37.3 minutes to complete training. In contrast, ResNet152 has the longest training time, requiring about 65.1 minutes, while ResNet101 falls in between with a training time of around 59 minutes. This discrepancy in training times could be attributed to the differences in the architecture and complexity of the models. Despite the differences in training time, ResNet50 manages to achieve the highest testing accuracy, indicating a potentially efficient trade-off between training time and performance.

Overfitting

In summary, while the testing accuracies for all three models are high, a definitive evaluation of overfitting is challenging due to the absence of training accuracy values. Overfitting typically occurs when a model performs significantly better on the training data than on unseen testing data. To thoroughly assess overfitting, it's important to compare the training and testing accuracy values and consider whether the models are learning noise from the training data that doesn't generalize well to new data.

7. YOLO FOR SPORTS IMAGE CLASSIFICATION

We have trained YOLO8s model which is very popular for image classification in field of deep learning and got some astonishing observation from the data which are as follow:

Training Progress: The training loss ("train/loss") steadily decreases as the number of epochs increases. This suggests that the model is effectively learning from the training data and optimizing its parameters to minimize the loss.

Validation Performance: The validation loss ("Val/loss") also decreases initially, indicating that the model is improving its generalization to unseen data. However, the decrease starts to plateau after a certain number of epochs, which could indicate that the model's improvement is slowing down.

Accuracy: The top-1 accuracy on the training data ("metrics/accuracy_top1") increases consistently with each epoch. This indicates that the model is becoming more accurate in predicting the correct class with the highest confidence.

Top-5 Accuracy: The top-5 accuracy on the training data ("metrics/accuracy_top5") is consistently high, close to 1. This suggests that even if the model's top prediction is not always accurate, it is often able to include the correct class within its top 5 predictions.

Learning Rate: The learning rate values ("lr/pg0," "lr/pg1," "lr/pg2") are decreasing with each epoch. This suggests that a learning rate schedule is in place, which is a common practice in training to fine-tune the model as optimization progresses.

Generalization and Overfitting: The model's performance on the validation data is comparable to its performance on the training data, which is a positive sign of good generalization. However, monitoring the gap between training and validation metrics over more epochs is important to ensure that the model doesn't start overfitting the training data.

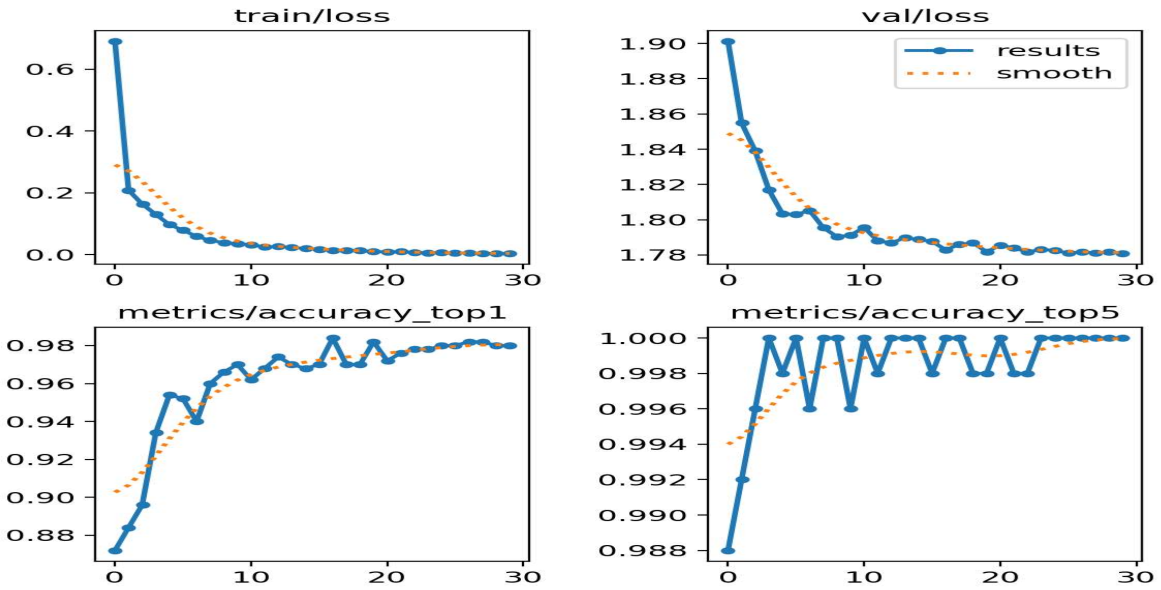
Convergence: Both the training and validation losses are converging to similar values, indicating that the model is likely approaching a relatively good solution. This is supported by the increasing accuracy metrics.

Stability: The metrics appear relatively stable, without significant spikes or drops. This suggests that the training process is smooth, and the model is gradually improving.

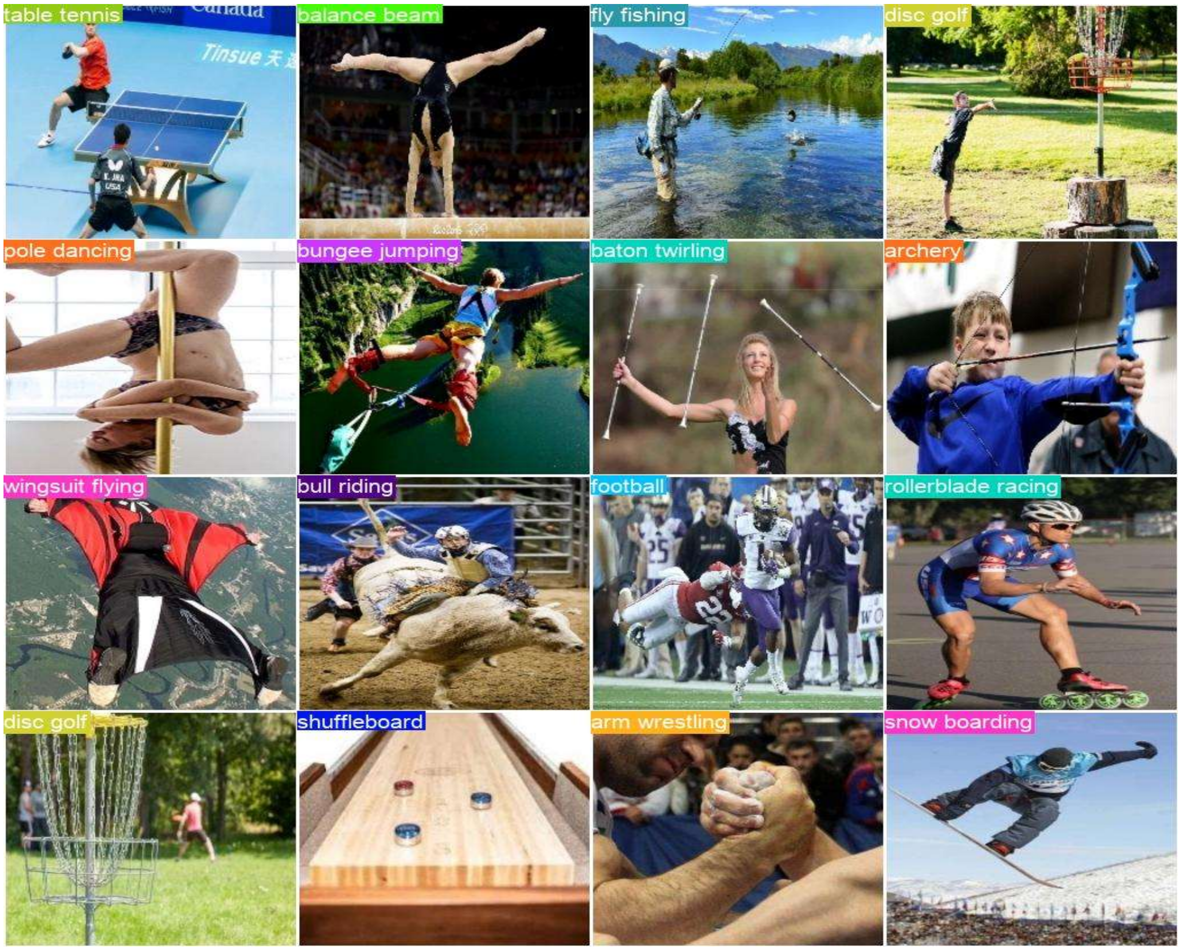
Hyperparameters: The model's performance could be further influenced by hyperparameters such as batch size, optimizer choice, and weight regularization. These factors can play a role in determining the rate of convergence and the final performance achieved by the model.

Fine-tuning: The gradual decrease in learning rate suggests that the model is undergoing a fine-tuning process, which is common in deep learning to help the model settle into a good solution.

Training Results for Yolo Model



This is Sample Output generated during testing model after fine-tuning and got around 98% of accuracy.



8. CONCLUSION

The study concludes by presenting a thorough method for classifying sports images using deep learning methods. Through painstaking model refining and tuning, the main goal of correctly classifying a wide range of sports-related activities seen in photographs has been accomplished. Approximately 14,500 sports photos from a carefully curated dataset, representing 100 different sports, have been used to train and assess deep learning models.

The significance of accurate sports image classification is highlighted by its potential applications in sports analytics, content categorization, and event recognition. The innovative emphasis on fine-tuning distinguishes this work from previous research in the field, allowing the models, specifically ResNet-50, ResNet-152, and ResNet-101, to effectively capture the intricate details and unique features associated with diverse sports activities. The study's results showcase the effectiveness of the proposed approach, with the models demonstrating improved accuracy in sports image classification. By leveraging state-of-the-art deep learning models and strategically refining them, the research contributes to advancing the domain of sports image analysis. The presented methodology, experimental insights, and findings not only enhance the understanding of sports-related data classification but also provide a novel perspective on refining deep learning models for this captivating and dynamic field.

In summary, this study makes a valuable contribution to the field of sports image classification by successfully addressing the challenges associated with accurately categorizing sports-related activities in images. Through meticulous fine-tuning of pre-trained models and a diverse dataset, the research demonstrates the potential for enhancing automated content tagging, sports analytics, and event recognition in the realm of sports analysis.

References

ADVANCED APPLIED MATHEMATICAL CONCEPTS FOR DEEP LEARNING CRN-82362-202203, In class notebooks, Prof. Ran Feldes

https://www.kaggle.com/code/littlebughenrylee/100-sports-classification-resnet-93-yolo-98

https://pytorch.org/

https://medium.com/ai-techsystems/sports-classification-with-cnn-d03715f7c1d3

https://towardsdatascience.com/pytorch-image-classification-tutorial-for-beginners-94ea13f56f2

https://github.com/sakethbachu/Sports-Image-Classifier