ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Leveraging Image Recognition for Sports Analytics

Machine Learning and Content Analytics Assignment

Table of Contents

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Team “Golf” Composition

- Κόκκαλης Ιγνάτιος - p2822212
- Λύσσας Χρήστος - p2822215
- Μπαμπίλης Κωνσταντίνος - p2822217
- Πελεκανάκης Χρήστος - p2822220

Introduction

In the modern era, the fusion of technology and sports has opened avenues for innovation and enhanced user engagement. This project embodies this synergy, exploring the domain of sports image classification through the lens of deep learning and business analytics.

Our project

This initiative focuses on the classification of images from different sports categories, leveraging deep learning architectures for precise categorization. Moving beyond a strictly technical endeavor, the project envisages implications in business analytics, and use cases, fostering personalized user experiences and informed business strategies.

Our Vision/Goals

Our goal is to develop an accurate, efficient, and scalable image classification system that can be seamlessly integrated into various business applications, thereby revolutionizing the sports industry dynamics.

Potential Business Uses

1. Tailored Marketing Campaigns

Marketers can leverage this tool to tailor campaigns more accurately by recognizing the sports depicted in user-generated content or social media platforms. It can facilitate brand alignments with particular sports, thus creating more resonant and targeted marketing strategies.

2. Content Personalization in Media Platforms

Media platforms can employ the tool to personalize content for users based on their preference for certain sports. By analyzing users' uploaded or liked images, platforms can suggest content, advertisements, or communities that align with the identified sports, enhancing user engagement.

3. Educational Tools for Physical Education

Educational institutions can use this tool in physical education settings as an interactive educational tool. It can help students learn about various sports by analyzing images and providing information on the rules, techniques, and history associated with the recognized sports.

4. Data Analytics for Sports Retailers

Sports retailers can implement this tool to analyze consumer preferences and trends by recognizing the sports depicted in customer-shared images. This data can be vital in shaping inventory decisions, designing sports gear, and creating targeted promotions.

5. Event Organizers and Venue Selection

Event organizers can utilize this tool to assess the popularity of different sports in various regions by analyzing images shared on social platforms. This could guide them in selecting appropriate venues and planning events that resonate well with the local audience.

6. Interactive Exhibitions in Museums

Museums and exhibitions can create interactive exhibits where visitors can learn about various sports through image recognition. Visitors can upload or select an image, and the tool will provide detailed information about the sport depicted, making the learning experience more engaging.

7. Fitness and Wellness Apps

Fitness apps can integrate this tool to help users explore different sports. By analyzing images users upload, the app can suggest similar sports to try, potentially fostering a community of users with similar interests and encouraging a healthy lifestyle.

8. Travel and Tourism

Travel and tourism agencies can utilize this tool to offer tailored packages to sports enthusiasts. By recognizing the sports depicted in user's images or preferences, agencies can suggest travel packages that include attending events or experiencing activities related to the identified sports, enhancing the personalization of travel experiences.

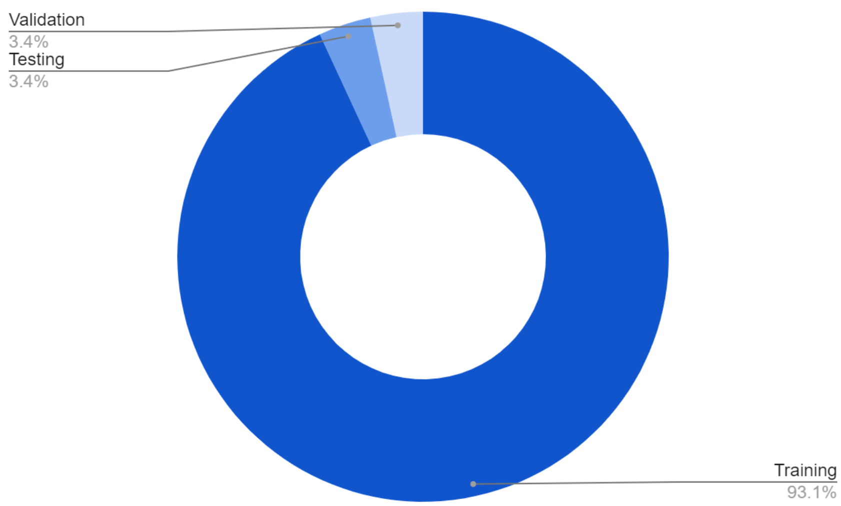
Methodology & Tools

The project utilized the Python programming language and Keras library for crafting the image classification models. The process encompassed model selection and training followed by validation to ensure optimum performance.

Dataset Overview

The dataset used was found on Kaggle under the title "100 Sports Image Classification"

- Total Images: 13,493 (Training), 500 (Testing), 500 (Validation)
- Image Dimensions: 224x224x3
- Image Format: JPG
- Directory Structure: Train, Test, Validate



The dataset encompasses a diverse array of images depicting various sports and activities, spanning from conventional sports such as archery, arm wrestling, bowling, football, water polo, and weightlifting, to more unconventional ones like wingsuit flying and NASCAR racing.

Models and Selection Rationale

In this project, we leveraged three renowned deep-learning models known for their efficacy in image classification tasks: ResNet, InceptionV3, and MobileNetV2. Below, we provide a brief description of each model and the rationale behind their selection:

1. ResNet: A deep convolutional neural network, ResNet stands out for its ability to mitigate the problems of vanishing gradients in deep networks using its residual learning framework. Its architecture aids in the training of much deeper networks without compromising accuracy, making it a suitable choice for our project that demands both depth and precision.

2. InceptionV3: Noted for its complex and efficient architecture, InceptionV3 optimizes both speed and computational cost. Its usage of inception modules facilitates numerous operations being performed concurrently, enhancing the model's capability to focus on various aspects of an image simultaneously, thus contributing to a more nuanced classification.

3. MobileNetV2: Tailored for mobile and embedded vision applications, MobileNetV2 offers an excellent balance between computational efficiency and accuracy. Its inverted residual structure and linear bottlenecks enable it to maintain a high level of accuracy while being computationally less demanding, aligning well with our project's requirements for efficiency and precision.

During the initial phase, we also explored the potential of integrating the EfficientNet model into our project, a scalable and efficient network known for achieving state-of-the-art accuracy. However, due to its intensive computational demands, it proved too taxing for our current computational resources, necessitating its exclusion from our selection to maintain a sustainable and efficient project workflow.

The chosen models, thus, represent a balanced blend of depth, efficiency, and accuracy, catering well to our project's unique demands and computational constraints.

Experiments – Setup, Configuration

ResNet

In our study, we experimented with two variations of models built upon the ResNet architecture to perform image classification tasks. The objective was to evaluate the efficiency and accuracy of different configurations and fine-tuning strategies.

Initial Model

Our initial model utilized the ResNet model and introduced a Flatten layer. The metrics we monitored were accuracy, loss, and the f1 score.

Fine-Tuned Model

For our fine-tuned model, we introduced a Global Average Pooling layer, to try and improve the model's metrics.

InceptionV3

In our second approach, we utilized the ResNet architecture to perform image classification. We experimented with two distinct approaches: initial and fine-tuned models, seeking to optimize accuracy and efficiency in the classification task.

Initial Model

The initial model capitalized on the ResNet pre-trained on the ImageNet dataset. Following the ResNet layers, we implemented a Flatten layer and a series of densely connected layers, culminating in a 100-class softmax output. This setup aimed to provide a strong baseline performance leveraging pre-trained feature representations.

Fine-Tuned Model

In contrast, the fine-tuned model introduced Global Average Pooling (GAP) prior to the densely connected layers, facilitating a focus on the most relevant features for classification. This modification, paired with fine-tuning of the ResNet layers, aimed to enhance accuracy and reduce loss by enabling more nuanced adaptations to our specific dataset.

MobileNetV2

Similarly, to the previous sections, we used two variations of models, but this time built upon the MobileNetV2 architecture.

Initial Model

The initial model integrated the MobileNetV2 as a base model, Following the base model, a Flatten layer was introduced to transition into the fully connected section of the network. During the training phase, an early stopping callback was introduced to monitor the loss and prevent potential overfitting.

Fine-Tuned Model

Subsequently, a fine-tuned version of the model was developed. This version retained the MobileNetV2 base but introduced a Global Average Pooling (GAP) layer to reduce the likelihood of overfitting.

Results & Quantitative Analysis (incl. visualizations)

In the results and quantitative analysis section, we aim to assess the performance of our models using various metrics that provide a comprehensive view of their behavior during the training and validation phases. Let us briefly understand each metric that has been employed to analyze the models' performance:

Loss (Training Loss)

The loss, specifically training loss, is a representation of the error between the predicted outcomes by the model and the actual ground truth during the training phase. A lower loss value indicates that the model is able to predict outcomes more accurately, with a good convergence towards the optimal solution.

Accuracy (Training Accuracy)

Training accuracy is a metric that depicts the proportion of correctly classified instances among the total instances in the training dataset. A higher accuracy percentage represents a model that is adept at identifying the correct categories during training.

Validation Loss (val_loss)

Validation loss, similar to training loss, measures the error between the predictions and the actual labels but in the validation dataset, which is a subset of data not seen by the model during training.

This metric helps in identifying how well the model is generalizing to new, unseen data. A model performing well should exhibit a decrease in validation loss over epochs.

Validation Accuracy (val_accuracy)

Validation accuracy mirrors the concept of training accuracy but is calculated over a separate validation dataset. A high validation accuracy indicates that the model is not only learning the underlying patterns in the training set but is also able to apply this understanding effectively to new, unseen data, showcasing good generalization capabilities.

F1 Score

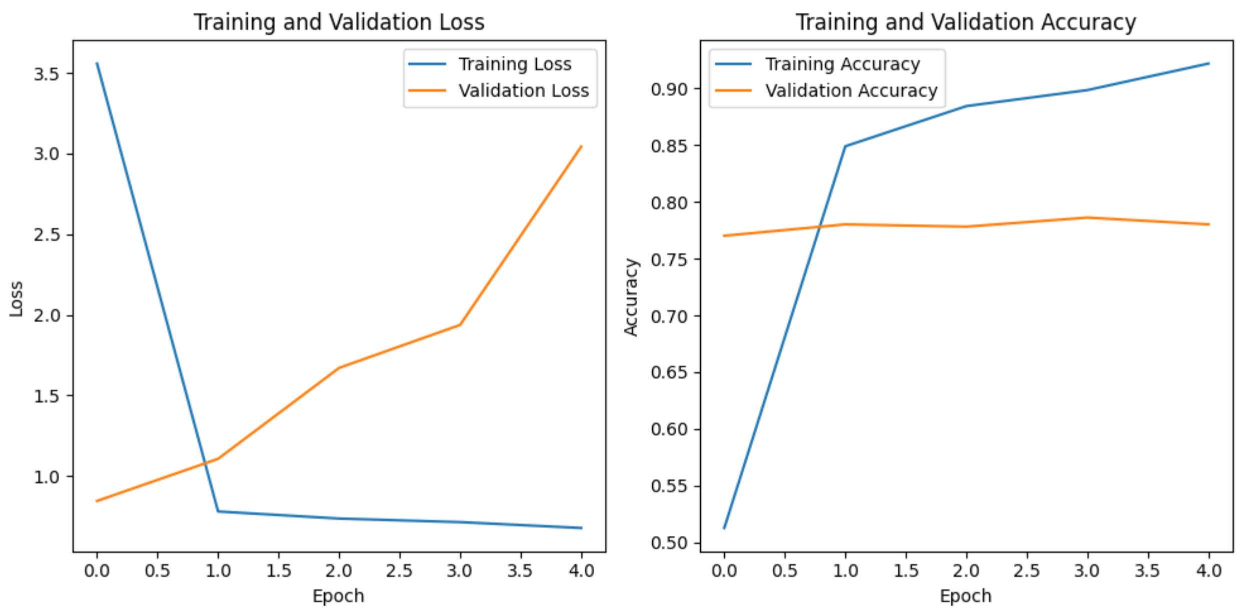
The F1 Score is a harmonic mean of precision and recall, two fundamental metrics in classification tasks. It serves as a balanced measure, especially in datasets where class imbalance might be a concern. An F1 score ranges between 0 and 1, where a higher value indicates better performance with a balanced consideration for both false positives and false negatives.

It is crucial to observe these metrics collectively, as they provide a well-rounded understanding of the model's performance, helping identify areas of strength and potential improvement. In the following section, we will apply these metrics to quantitatively analyze our models' performance, backed by a detailed discussion on the observed trends and their implications.

ResNet

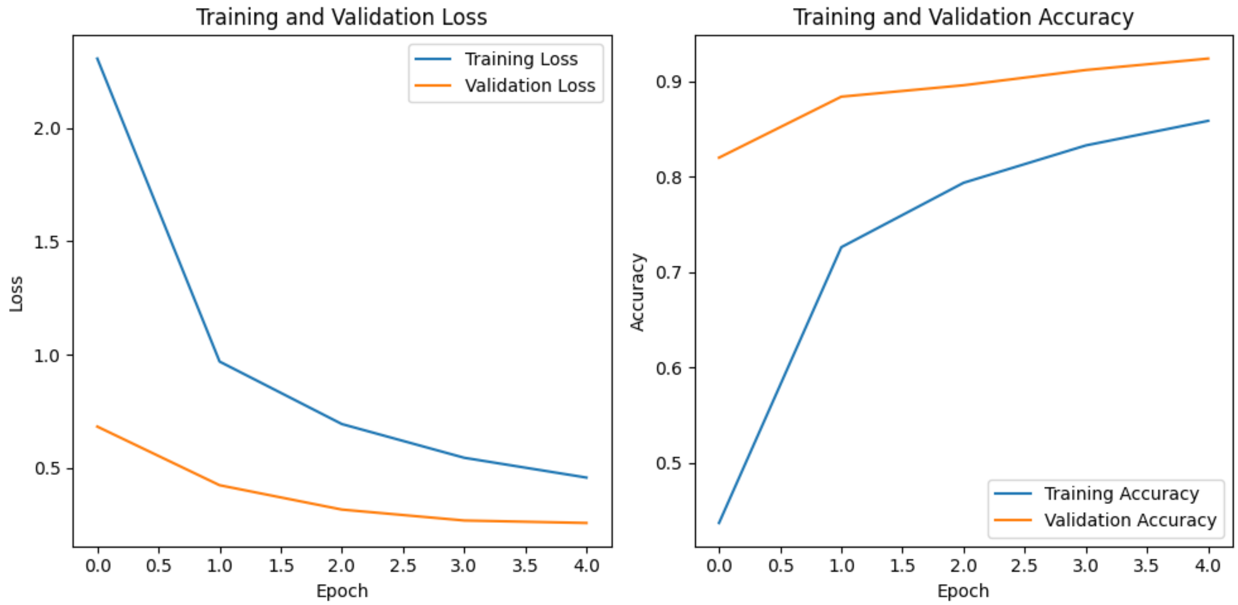
Initial Model

	loss	accuracy	f1 score	val_loss	val_accuracy	val_f1 score
Epoch 1	3.5572	0.5129	0.642	0.8464	0.77	0.8898
Epoch 2	0.7811	0.8487	0.8762	1.1074	0.78	0.8804
Epoch 3	0.7373	0.8841	0.9049	1.6709	0.778	0.879
Epoch 4	0.7152	0.8982	0.9161	1.9373	0.786	0.8907
Epoch 5	0.6789	0.9215	0.9364	3.0421	0.78	0.8807



Fine-tuned Model

	loss	accuracy	f1 score	val_loss	val_accuracy	val_f1 score
Epoch 1	2.3057	0.4367	0.5827	0.6826	0.82	0.9041
Epoch 2	0.9694	0.7261	0.7809	0.4241	0.884	0.9427
Epoch 3	0.6938	0.7937	0.8332	0.3166	0.896	0.9512
Epoch 4	0.5452	0.8329	0.8646	0.2685	0.912	0.9495
Epoch 5	0.4578	0.8587	0.8831	0.2576	0.924	0.9589



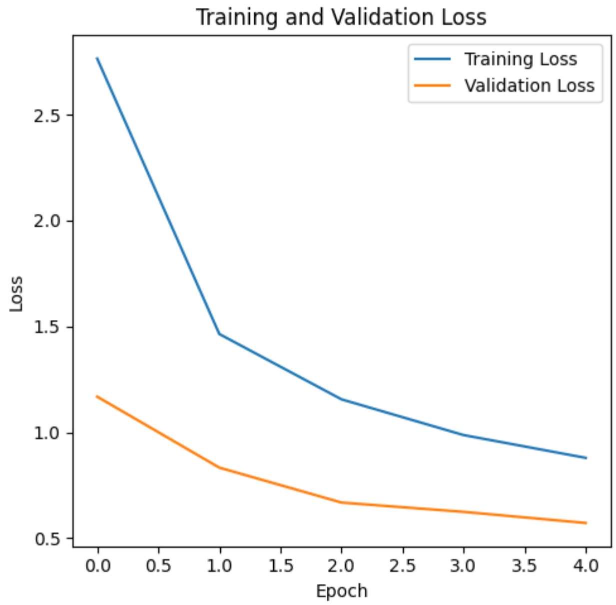
InceptionV3

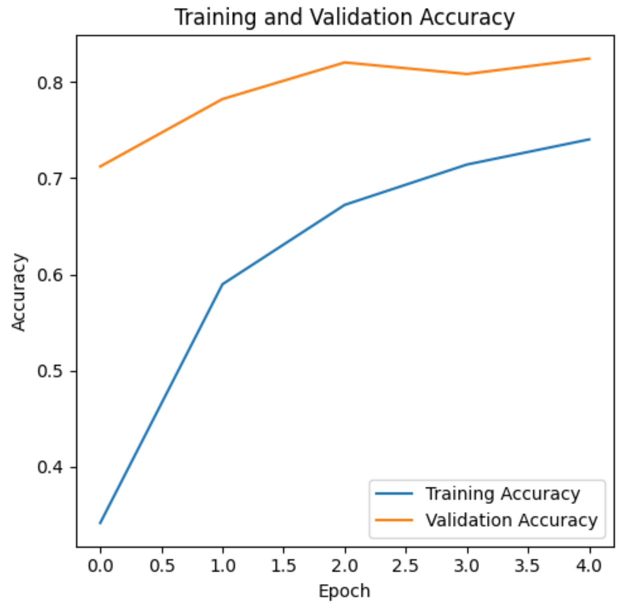
Initial Model

	loss	accuracy	f1 score	val_loss	val_accuracy	val_f1 score
Epoch 1	3.5332	0.446	0.6004	1.0017	0.75	0.862
Epoch 2	1.0671	0.7391	0.7936	0.9292	0.75	0.862
Epoch 3	0.7681	0.8128	0.8476	1.1567	0.77	0.8775
Epoch 4	0.6021	0.8574	0.8833	1.087	0.792	0.8824
Epoch 5	0.6398	0.8625	0.8831	1.4919	0.802	0.8905

Fine-tuned Model

	loss	accuracy	f1 score	val_loss	val_accuracy	val_f1 score
Epoch 1	2.7646	0.3418	0.5226	1.168	0.712	0.8457
Epoch 2	1.4638	0.5897	0.6802	0.8328	0.782	0.8821
Epoch 3	1.1554	0.6721	0.7377	0.6679	0.82	0.8936
Epoch 4	0.9866	0.714	0.7717	0.624	0.808	0.8874
Epoch 5	0.879	0.7401	0.7914	0.5715	0.824	0.8995

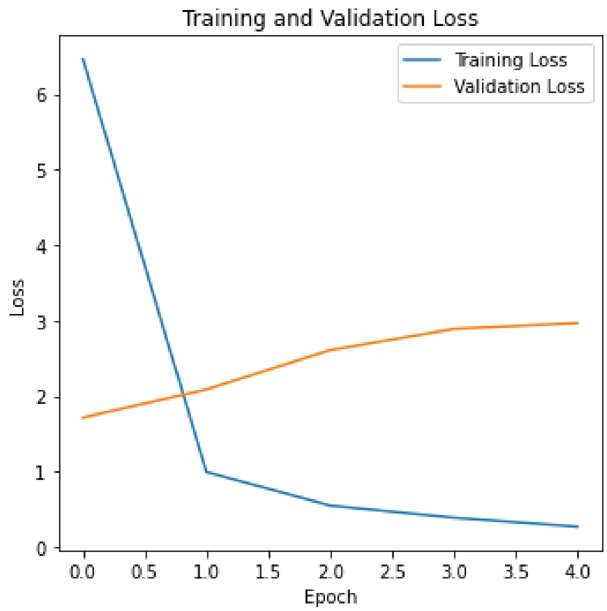


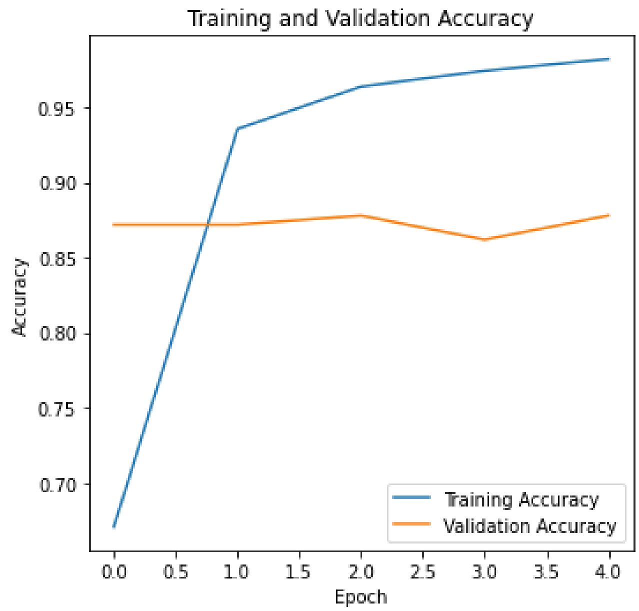


MobileNetV2

Initial Model

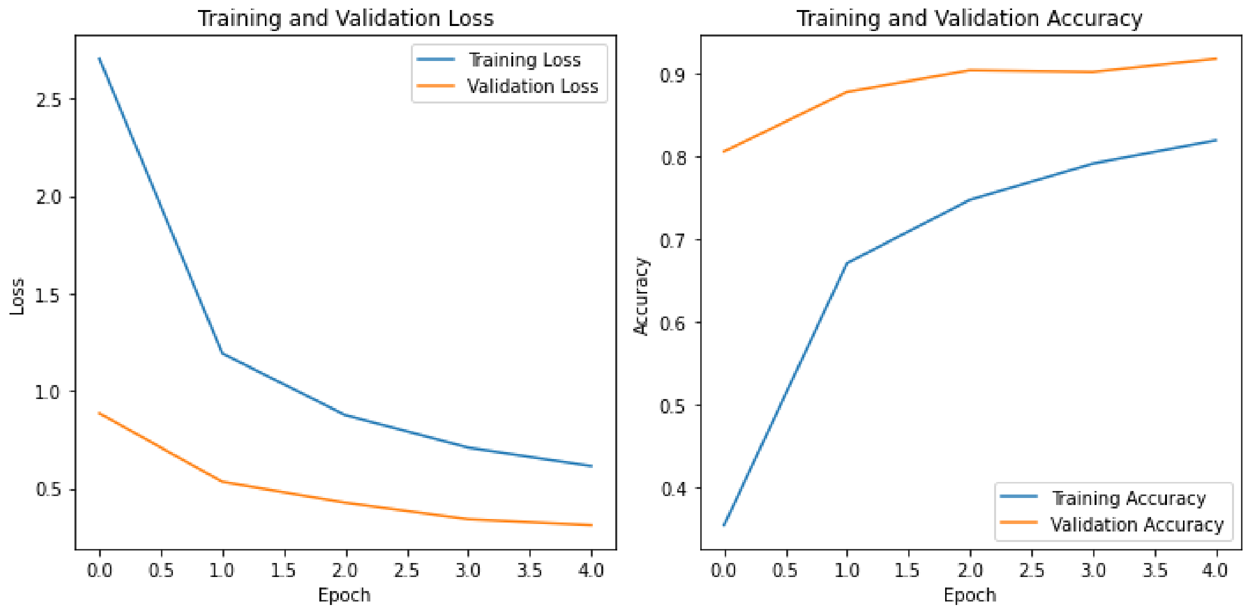
	loss	accuracy	f1 score	val_loss	val_accuracy	val_f1 score
Epoch 1	6.4651	0.6708	0.7556	1.7143	0.8720	0.9257
Epoch 2	0.9949	0.9357	0.9469	2.0874	0.8720	0.9224
Epoch 3	0.5494	0.9640	0.9704	2.6076	0.8780	0.9386
Epoch 4	0.3882	0.9744	0.9786	2.8884	0.8620	0.9261
Epoch 5	0.2690	0.9824	0.9849	2.9655	0.8780	0.9311





Fine-tuned Model

	loss	accuracy	f1 score	val_loss	val_accuracy	val_f1 score
Epoch 1	2.7046	0.3546	0.5330	0.8846	0.8060	0.8964
Epoch 2	1.1916	0.6711	0.7366	0.5327	0.8780	0.9320
Epoch 3	0.8757	0.7476	0.7969	0.4260	0.9040	0.9440
Epoch 4	0.7084	0.7914	0.8286	0.3403	0.9020	0.9424
Epoch 5	0.6141	0.8194	0.8552	0.3100	0.9180	0.9509



Qualitative & Error Analysis

ResNet

Model Performance Overview

During the initial phase of training with the ResNet model, it was observed that the model achieved a validation accuracy of 78%, a validation loss of 3.0421, and an F1 score of 0.8807 after 5 epochs. This presents a fairly strong start, with a high F1 score indicating a balanced precision and recall in the classification task. However, the relatively high validation loss indicates the model might be struggling to correctly classify some of the classes, potentially making errors in distinguishing between classes with similar features.

Following the fine-tuning process, the ResNet model demonstrated significantly improved performance, recording a validation accuracy of 92.4%, a markedly decreased validation loss of 0.2576, and an F1 score of 0.9589 after 5 epochs. These metrics indicate that the fine-tuning process has enhanced the model's ability to generalize from the training data to unseen data, leading to more accurate classifications.

Error Analysis

The error analysis highlighted a disparity in performance between the initial and fine-tuned models, evidenced by a significant reduction in validation loss from 3.0421 to 0.2576. This suggests that our initial model was suffering from overfitting the training dataset. The overfitting hypothesis is backed by the upward trend of the validation loss as the epochs progress.

Qualitative Analysis

The qualitative analysis underscored the improvements in model generalization achieved through fine-tuning, as indicated by heightened validation accuracy and F1 score.

Conclusion & Future Directions

In conclusion, the fine-tuned ResNet model exhibits substantial promise in the classification task, showcasing significant enhancements in performance metrics compared to the initial model. Future work can focus on refining the model further, potentially exploring advanced techniques to mitigate identified errors and augment accuracy. A closer inspection of the misclassifications and expanding the dataset to include a more diverse range of images might also be pivotal steps in further honing the model's performance.

InceptionV3

Model Performance Overview

The experimentation with the InceptionV3 architecture yielded promising results. The initial model posted a validation loss of 1.4919, a validation accuracy of 80.2%, and a validation F1 score of 89.05%. The fine-tuned version further enhanced performance, recording an improved validation loss of 0.5715 and an F1 score of 89.95%, alongside a slight increase in validation accuracy to 82.4%.

Error Analysis

A deeper dive into the error analysis reveals that the fine-tuned model significantly reduced the validation loss compared to the initial model, indicating a better generalization on unseen data. However, the marginal enhancement in validation accuracy suggests that there might be areas where the model is struggling to correctly classify some specific categories, potentially hinting at the existence of complex patterns or noise in the data that could not be fully captured.

Qualitative Analysis

Qualitatively, the higher F1 score in the fine-tuned model points towards a more balanced classification performance across different classes, even though the increase in accuracy is slight. It demonstrates that the fine-tuned model, with global average pooling, seems to be more adept at distinguishing between various features, thus being a more reliable choice for nuanced image recognition tasks.

Conclusion & Future Directions

In conclusion, the fine-tuned InceptionV3 model signifies a step forward in optimizing image classification tasks, demonstrating improvements in both loss reduction and balanced class prediction, as mirrored in the F1 score. Looking ahead, future endeavors may explore further architectural refinements or incorporate additional data augmentation strategies to build upon the achievements of the fine-tuned model, potentially unlocking even higher levels of accuracy and efficiency. Moreover, more comprehensive error analyses could be conducted to pinpoint specific areas where the model can be refined further.

MobileNetV2

Model Performance Overview

The initial MobileNetV2 model, which incorporated a Flatten layer, demonstrated a validation loss of 2.97 and a validation accuracy of 88%. This indicated that the model had a decent capability of categorizing the images correctly, although the relatively high validation loss suggested that the predictions were not exceedingly confident, potentially pointing to areas where the model could be fine-tuned to achieve better precision and avoid overfitting.

On the other hand, the fine-tuned MobileNetV2 model which utilized a Global Average Pooling layer exhibited superior performance with a significant reduction in validation loss to 0.31 and an improvement in validation accuracy, reaching 91.8%. This demonstrated that the fine-tuned model was not only more proficient at correctly identifying the categories of the images but also made these predictions with a higher degree of confidence.

Error Analysis

Upon dissecting the errors observed during the validation phase, several insights can be derived. The initial model, with a higher validation loss, might be experiencing a higher rate of misclassifications. This suggests the possibility of the model being sensitive to the noise and variations present in the dataset, indicating a need for further optimization to enhance its generalization capabilities.

Conversely, the fine-tuned model with a significantly lower validation loss portrays a more robust performance. The presence of the Global Average Pooling layer might be contributing to a more abstract representation of features, helping in mitigating overfitting and achieving better generalization on the validation set.

Qualitative Analysis

Qualitatively, the fine-tuned model’s higher validation accuracy indicates a better discernment of complex patterns within the image data. This suggests that the adjustments made to the architecture, including the incorporation of global average pooling, were successful in enhancing the model’s ability to differentiate between varied and nuanced categories present in our dataset.

Conclusion & Future Directions

From our analysis, it is evident that the fine-tuned model exhibits superior performance compared to the initial model, both in terms of accuracy and loss metrics. Future work should focus on further exploring optimization techniques and possibly expanding the training dataset to enhance the model's ability to generalize across a broader spectrum of data variations.

Final Evaluation and Conclusion

After a rigorous evaluation of the MobileNetV2, ResNet, and InceptionV3 models on the task of sports image recognition, it is evident that different models offer diverse strengths. Considering the metrics obtained after 5 epochs, the fine-tuned ResNet model outshines the InceptionV3 model with significant margins in terms of validation loss and accuracy, and still barely comes ahead when compared to the MobileNetV2 model. Its superior performance is further underlined by the highest F1 score, a measure that considers both the precision and the recall of the test.

The ResNet fine-tuned model demonstrated a validation loss of 0.2576, an accuracy of 92.4%, and an F1 score of 0.9589, indicating its proficient ability to generalize and make accurate predictions on unseen data. Therefore, we recommend proceeding with the ResNet fine-tuned model for further developments in the sports image recognition project. This model not only showcases high accuracy but also promises robust performance, making it a reliable choice for real-world applications. In the future, it would be beneficial to explore potential improvements and optimizations to further enhance its capabilities.

Initial	val_loss	val_accuracy	val_f1 score
ResNet	3.0421	0.78	0.8807
InceptionV3	1.4919	0.802	0.8905
MobileNet	2.9655	0.8780	0.9311

Fine Tuned	val_loss	val_accuracy	val_f1 score
ResNet	0.2576	0.924	0.9589
InceptionV3	0.5715	0.824	0.8995
MobileNet	0.3100	0.9180	0.9509

Members/Roles

In the execution of this significant project, a collaborative approach was strategically adopted to optimize the resources and strengths of each team member. The initial stage of model selection was a collective effort where all four members actively participated, bringing diverse perspectives and expertise to the table in choosing the most promising models for our task.

Following this united decision-making process, the responsibility of running the models was distributed among the team members. This division of labor was not only a step to expedite the project but also a means to utilize everyone's computational power to the fullest, ensuring an efficient and streamlined process. Each member was accountable for running a portion of the models, monitoring their performance, and gathering preliminary results.

Once the individual tasks were completed, the team reconvened in a series of meetings where detailed discussions were held to critically analyze the results yielded by each model. These collaborative sessions were vital in synthesizing the collective findings and insights, fostering a rich understanding of the models' performances.

In the final phase, the team worked in unison to craft the final report. This involved consolidating the individual summaries of each model, and meticulously composing a coherent and comprehensive document that encapsulated the entirety of our research and findings.

Throughout the project, the synergy of the team was apparent, as every member brought their unique strengths to the forefront, contributing immensely to the project's successful completion. The collaborative spirit ensured that the final report was not only a compilation of results but also a testament to the cohesive and effective teamwork that powered this project.

Time Plan