DS340 Final Project

Data Augmentation and 101 Sports Classification Neural Network

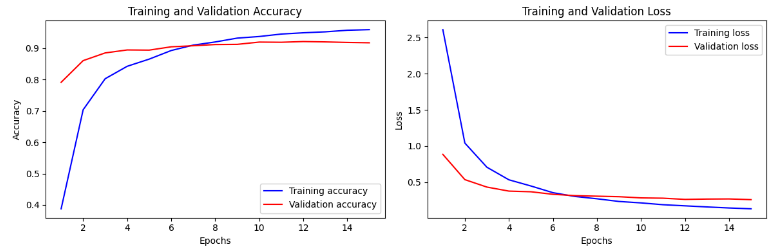
Laerk Ndreko and Sadiq Alhumood

To begin, the problem we were trying to solve with this project was to experiment with how data augmentation paired with neural networks can help classify different types of images. For this project, we chose to classify 101 different sports. We wanted to tackle this issue to demonstrate our problem-solving skills with neural networks on a topic we are both interested in: sports.

To tackle the challenge of classifying sports images, we decided to start with a convolutional neural network. However, it would take a lot of time to train a new one to identify images with additional layers due to the size of our dataset. Therefore, we decided to use leverage transfer learning with the state-of-the-art neural network ResNet50 and EfficientNetB0 models for image classification. We used the model without its top layer so that we could use our own classification groups. We also froze training so that we could leverage what the model already knows. In addition, we tried to reduce overfitting by dropping out some of the inputs in training. Using datasets from Kaggle consisting of 101 sports, we trained the model, and we kept getting low accuracies due to the imbalance of some sports in the datasets. The soccer folder was not part of the original dataset, so we had to input it manually. It had more than 2000 images in training which was very large compared to the 100-200 images in the other sports folders. Therefore, we removed all additional images and trained the model with a balanced dataset.
We also had issues with the validation accuracy being very high due to the lack of images in the validation set. So we decided to restructure the whole dataset to make it well distributed and balanced, where 80% of images were in training and 20% were split into testing and validation sets.

After training the model with both EfficientNetB0 and ResNet50, we got a higher accuracy of ~0.92 with ResNet50 compared to the ~0.91 that EfficientNetB0 gave us. We decided to stick with ResNet50 due to its efficiency and effectiveness.

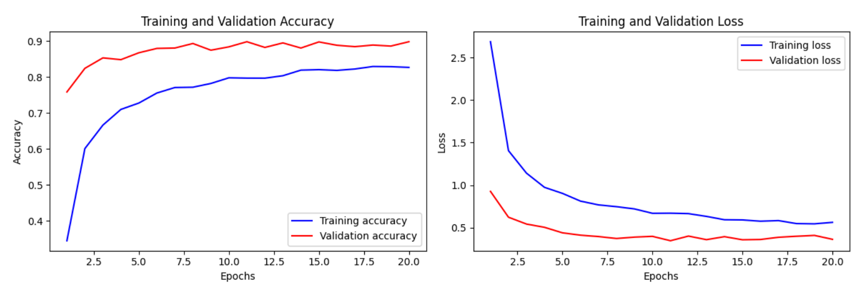
Figure 1



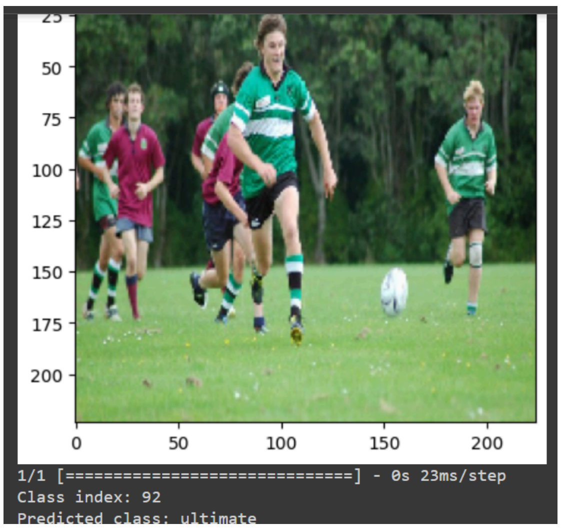
Although we had a high accuracy with our model, we soon realized that the model was overfitting as we compared validation accuracy with training accuracy as shown in Figure 1. This was our conclusion as the training accuracy began to get much higher than the validation accuracy, so we assumed that the model was memorizing data well, but not generalizing new data correctly. In addition, we also tested random pictures of sports from Google, and photos of our own. Most of them were classified incorrectly.

To fix this we decided to play with the images of the dataset, and to add variation, therefore we added a data augmentation method to rotate, zoom in, shift, or change the image in some way so that the model would be trained with a better dataset.

Figure 2



The results were much more sufficient, we got a total accuracy of ~0.90 we have seen much less overfitting as we compared validation accuracy with training accuracy again, as shown in Figure 2. When retesting the specific images that were misclassified when there was no data augmentation, the newly implemented feature of our model showed to be successful in generalizing better to data it had not seen previously. This was the case with rugby, a team sport that involved playing on a field. Team-sports and those on fields were harder for the model to distinguish, but after data augmentation, the model got the prediction correct. Here is the before and after of data augmentation:



In conclusion, this project provided valuable insights into the challenges and techniques of image classification with neural networks. We learned the importance of using a balanced dataset and the benefits of transfer learning, which allowed us to experiment with pre-trained models like ResNet50 and EfficientNetB0 for effective feature extraction. Overfitting was our biggest challenge, which we addressed through data augmentation, a powerful tool that enhanced the model's ability to generalize better to new images. Testing and refining different models with varying features was necessary for finding the best model in ResNet50. In our final remarks, this project deepened our understanding of neural network problem-solving and troubleshooting.

While we are aware that our model is not completely accurate, the process of finding the best solution will be something that we will carry with us in the future.

Datasets:

https://www.kaggle.com/datasets/gpiosenka/sports-classification/data

https://www.kaggle.com/datasets/fabrciohenrique/soccer-images