Final Project Proposal: Soccer Image Classification with Convolutional Neural Networks (CNN)

Group 4: Saloni Birthare, Rishin Tiwari, Martin Zuluaga

Introduction

Soccer is the most popular event worldwide, with an approximate 4 billion fans globally, generating a massive amount of visual content from recorded and streamed events.

Most Popular Sports Ranked By Total Fans

Rank	Sport Name	Amount of Fans	Regions
1	Football (Soccer)	3.5 Billion	Europe, Africa, Asia, America
2	Cricket	2.5 Billion	Asia, Australia, UK
3	Hockey	2 Billion	Europe, Africa, Asia, Australia
4	Tennis	1 Billion	Europe, Asia, America
5	Volleyball	900 Million	Europe, Australia, Asia, America
6	Table Tennis	850 Million	Europe, Africa, Asia, America
7	Basketball	800 Million	America, Oceania, Middle East
8	Baseball	500 Million	America, Japan
9	Rugby	475 Million	Oceania, South Africa, England
10	Golf	450 Million	America, Oceania, Europe

(Veroutsos)

This content can be used to analyze and classify different plays. Automatic classification of soccer events can have tremendous potential impact on different business activities and industries, enhancing sports broadcasting, coaching, referee assistance, fan engagement and betting experience.

Currently, classifying soccer events from broadcasts and streaming relies on manual analysis processes. These processes are not only extremely time-consuming, but they are prone to human error. This proposal aims to tackle these issues by using Convolutional Neural Networks (CNNs) and create a competitive deep learning model capable of classifying soccer images into predefined categories.

Automating the classification of soccer images, the model looks to streamline the process, reduce potential errors, and provide valuable insights and trigger business decisions for its stakeholders. The specific event categories in this model include penalty kicks, free kicks, yellow cards, red cards, substitutions, tackles and corner kicks.

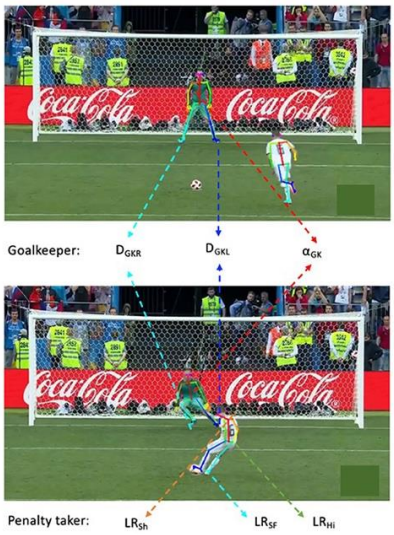
These events are extremely important in the dynamics of soccer, as many of them represent a high scoring chance, affecting the game by ejecting a player or the substitution of a player. These events

can be translated into player performance metrics, live betting odds, broadcast highlights and statistic records.

In this project, the primary business use case is, and purpose is to provide valuable player and team performance analysis. Soccer clubs and coaches can use image classification to analyze player and team performance during matches and training. By categorizing and analyzing soccer events, they can evaluate the overall player contributions, tactical decision and game strategies.

Using the model to help coaches analyze game events for training and identify individual and collective strength and weaknesses in their own teams and rivals. This application can boost player development and overall team performance. This analysis can also be used to analyze other team's players performances and plan potential signings and investments.

The following image is a notable example of how coaches can analyze the positioning, technique and overall performance of players when taking/saving a penalty kick:



Understanding the attributes of each player and personalizing training sessions can make the difference in a small margin of error situation such as a penalty shootout.

Dataset Description

The dataset used for this project consists of approximately 7,000 soccer images collected from different online sources. These images display a truly diverse range of soccer events, providing a robust dataset for training and testing the model.

Class Name	# of Train Images	# of Test Images
Corner Kick	800	200
Free Kick	800	200
Penalty Kick	800	200
Red Card	800	200
Yellow Card	800	200
Substitute	800	200
Tackle	800	200

The images were sourced from different online sources, using Python's 'bing_image_downloader' library for efficient scraping and collection from videos. This eased gathering images through automated data scraping techniques to ensure a large variety of events.

The images were collected from professional soccer competition events from multiple seasons and tournaments, ensuring a realistic set of images. Regarding geographical context, the images also vary in location and players involved, ensuring diversity on the representation of the events.

For this model, the target variable consists of the event category, with seven distinct options: red cards, yellow cards, penalty kicks, free kicks, corner kicks, tackles and substitutions.

Modeling Approach

Given the visual nature of the business use cases and the dataset, Convolutional Neural Networks (CNNs) are the selected model for this project. CNNs are recognized for their reliable ability to detect features and patterns in images, making them ideal for image classification tasks.

Data Processing and Transformation

To ensure consistent data quality, the following preprocessing techniques will be employed:

- Uniform image size and format: all images will be resized to a uniform dimension (28 x 28) to keep consistency during model training.
- Data augmentation: rotation, scaling and flipping techniques will be applied to enhance the dataset's diversity and improve the model's generalization ability.
- Data cleaning: irrelevant and duplicate images will be removed from the dataset to ensure data quality. Inadequate quality and unclear content images will be filtered out of the dataset to improve the performance of the model.

Model Selection and Fitting

The initial approach consists of training a baseline CNN model using the built dataset to set up initial performance metrics. The architecture of the CNN will be designed to capture key features from soccer images, primarily focusing on recognizing the specific events previously described.

The following steps will outline and decide the model selection and fitting process:

- Convolutional layers: the model consists of multiple convolutional layers, varying in kernel sizes and extracting features at different scales.
- Activation function: Rectified Linear Unit (ReLU) will be used as the activation function to introduce non-linearity into the neural network.
- Pooling layers: max-pooling layers will be included to reduce dimensionality and ease the retention of key features.
- Fully connected layers: after the convolutional and pooling layers, the model will have fully connected layers to classify the images into the event categories.
- Output layer: the final (output) layer will use a softmax activation function to provide the probabilities for each of the 7 event categories.

Advanced Models

After stablishing the baseline performance of the initial CNN model and performing some hyperparameter tuning and optimization techniques, the project will continue through the exploration of advanced predefined models LLAVA (Large Language and Vision Assistant) LLM model using ollama for image classification. This combined system efficiently categorizes images depicting various soccer events such as corner kicks, free kicks, penalty kicks, red cards, yellow cards, substitutes, or tackles.

Leveraging the capabilities of LLAVA, the system not only identifies the depicted soccer event but also generates detailed descriptions, enhancing the interpretability and usability of the classification results. This integrated approach synergizes the strengths of computer vision and natural language processing, providing a comprehensive solution for soccer event recognition and description in image data.

These types of models are recognized for their efficiency and high accuracy on large datasets, making them proper for further analysis and exploration for this project. The goal is to improve accuracy and efficiency when classifying the soccer images into the predefined event categories.

Evaluation Metrics

The following metrics will be considered to evaluate the performance of the model:

- Accuracy: the percentage of correctly classified images.
- Precision and recall: precision measures the proportion of true positive classifications among all positive predictions, while recall measures the proportion of the true positive classifications among all actual positive cases. These metrics will be further analyzed and explained in the specific context of the different business use cases.
- Log loss: measures the accuracy of a classification model's predicted probabilities against actual class labels, making it relevant when evaluating CNNs, providing a continuous probabilistic assessment of the classification performance.
- F1 score: a balanced performance measure of precision and recall.
- Confusion matrix: provides a visual representation of the model's performance and helps identify and understand areas of improvement. Can provide information on what classes are being commonly confused by the model.

These metrics were selected to ensure a balanced evaluation of the model's performance, focusing on the model's ability to classify specific events accurately.

Conclusion and Future Research

The project aims to develop a CNN model for classification of soccer images into different soccer event categories. The model can have a significant impact in different businesses, like sports broadcasting, coaching, sports betting, as well as other disciplines within soccer and the sports industry in general. One of the major motivations of the project is its high potential for scalability and reproduction, being able to grow and reach huge markets.

The proposed approach uses advanced neural network architectures and an original dataset of annotated soccer images to achieve a reliable performance when classifying soccer events.

Reflecting on the project's goals and objectives, we believe we are on the correct track, by obtaining original data and having an outline that will build a competitive model. However, there are other use cases and research opportunities that can be considered in the future:

- Broadcasting and game highlights: use the model to automatically generate highlights for broadcasts, enhancing the user experience.
- Sports betting and player performance tracking: provide valuable and real time data for improving betting accuracy and odds by classifying live soccer events and tracking player performance throughout the season.
- Referee assistance and injury prevention: help referees with automated analysis and finding potential injury risks to protect the players from injuries.

These business use cases offer significant opportunities for growth and scalability for future research and further development of the proposed model.

References

This project proposal is inspired by the research paper:
Karimi, A., Toosi, R., & Akhaee, M. A.
(2021). Soccer Event Detection Using Deep Learning

Online resources papers used for research and graphic sources:

The Most Popular Sports In The World

Olav Andre Nergård Rongved, Markus Stige , Steven Alexander Hicks, Vajira Lasantha Thambawita, Cise Midoglu, Evi Zouganeli 1, Dag Johansen 4, Michael Alexander Riegler and Pål Halvorsen. (2021). Automated Event Detection and Classification in Soccer

How Data Science is Changing Soccer