Zusammenfassung

Fußball stellt bekannte Herausforderungen in der KI-Forschung dar. Eine Abstraktion mit kleinerem Parameterraum liegt im Tischfußball. Bei einem halbautomatisierten Kicker wird ein Team mit Industriemotoren durch einen KI-Agenten gesteuert, während das andere Team von Menschen gespielt wird. Für eine dynamische Spielweise benötigt der KI-Agent den Zugriff auf so viele Daten wie möglich, insbesondere auf den Zustand des Spiels in echtzeit. Bei einem Kickerspiel wird der Spielzustand durch die Position und die Rotation der einzelnen Figuren sowie die Position des Balls definiert. In dieser Arbeit wird ein System konzipiert, welches die Erkennung des Spielzustands auf Basis von Computer Vision und Convolutional Neural Networks ermöglicht. Das System wird auf einem halbautomatisierten Tischkicker getestet, trainiert und evaluiert. Es wird ein End-to-End-Bildregressionsmodell verwendet, um die Position und die Rotation der einzelnen Stangen vorherzusagen. Fünf verschiedene Architekturen, nämlich ResNet18, ResNet50, MobileNetV3, EfficientNetV2 und eine eigene Architektur, werden als Feature-Extractor für das Regressionsmodell verwendet und evaluiert. Für das Training der Modelle wird ein Datensatz erstellt, welcher die Position und Rotation der einzelnen Stangen enthält. Die Daten der schwarzen, automatisierten Figuren werden direkt aus den Motoren ausgelesen. Im Gegensatz dazu müssen die Daten für die weißen Figuren manuell erhoben werden. Dazu wird die Rotation mit Beschleunigungssensoren gemessen, während die Position auf Basis des Bildmaterials berechnet wird. Das System zur Erkennung des Spielzustands wird als Prototyp entwickelt und evaluiert. Es wird gezeigt, dass der Prototyp eine hohe Genauigkeit bei der Vorhersage der Position und Rotation der Figuren erreicht, welche den Anforderungen entspricht. Speziell die ResNet-basierten Modelle erreichen vielversprechende Ergebnisse. Die geforderte Vorhersagegeschwindigkeit von 60 Frames pro Sekunde konnte das System aufgrund hoher Inferenzzeiten und sequentieller Ausführung jedoch nicht erreichen. Am Ende dieser Arbeit wird ein Überblick über mögliche Lösungen für dieses Problem präsentiert.

Abstract

The game of Football poses well-known challenges in AI research. A smaller parameter space abstraction can be found in the game of Foosball. In a semi-automated Foosball game, one team is controlled using industrial motors through an AI agent while the other team is played by humans. For dynamic gameplay, the AI agent requires access to as much data as possible, namely the real-time game state. In a Foosball game, the game state is defined by the position and rotation of the individual figures as well as the position of the ball. In this work, a concept for a real-time game state detection system based on Computer Vision and Convolutional Neural Networks is presented. The system is tested, trained and evaluated on a real-world semi-automatic Foosball table setup. An end-to-end image regression model is employed to predict the position and rotation of the individual rods. Five different architectures, namely ResNet18, ResNet50, MobileNetV3, EfficientNetV2 and a custom architecture, are utilized and evaluated as feature extractors for the regression model. For the training of the models, a custom dataset is created containing the position and rotation of the individual rods. The data for the black, automated figures is directly received from the motors. The rotation of the white, human played figures is measured by using accelerometer sensors while the position is calculated based on image data. The game state detection system is implemented as a proof-of-concept. It is demonstrated that the prototype achieves high accuracy in predicting the position and rotation of the figures, meeting the requirements. Especially the ResNet-based models show promising results. However, the system could not achieve the required prediction speed of 60 frames per seconds due to high inference times and sequential execution. An overview over possible solutions to this is presented at the end of this work.