Entwicklung eines neuronalen Netzes zur Beurteilung der Landung eines Schispringers

Thomas Hofer BSc[1], Bernhard Hollaus PhD[1]

Medizin, Gesundheits- und Sporttechnologie, Management Center Innsbruck[1], Österreich

E-Mail: hofertommy@hotmail.com

Schlüsselwörter: Neuronales Netz, Telemarklandung, Klassifizierung

Einleitung

Aufgrund der stetig ansteigenden Dichte im Schisprungsport wird der Anteil der Haltungsnoten immer wichtiger für die Athleten, damit sie sich im Spitzenfeld behaupten können. Wiederum ein großer Teil dieser Haltungsnoten fallen auf die Landebewegung und der damit verbundenen Telemarklandung. Eine genaue und regelkonforme Auswertung der Landung bekommt im Training kaum Beachtung und ist ohne Anwesenheit eines Kampfrichters kaum möglich. Die bisherigen Forschungen in diesem Gebiet belaufen sich auf die Beurteilung und Verbesserung des Absprungs des Athleten oder wurden mit Hilfe von Sensorwerten durchgeführt. Das Ziel dieser Arbeit ist es, ein neuronales Netz zu erstellen und zu trainieren, damit eine objektive Entscheidung über die Landung im Schisprungsport gefällt werden kann. Die Entscheidung beläuft sich dabei darauf, ob der Athlet eine Telemarklandung vollbringt oder nicht. Als Trainingsdaten werden 1082 Bilderserien der automatisierten Weitenmessung verwendet, welche jeweils aus insgesamt 21 Serienbilder des gesamten Landeablaufs bestehen.

Methode

Die Trainingsdaten werden vorab mit zwei vorhanden Algorithmen (Yolov5 – „You only look once“ und OpenPose) und einem selbst erstellten Codeblock bearbeitet, um am Ende Koordinaten von Keypoints (Hauptpunkte des menschlichen Skeletts) aus der Pose des Athleten zu erhalten.

Im ersten Schritt werden in den 21 Bilder der gesamten Landesequenz mit Hilfe des YoloV5 Algorithmus alle Personen lokalisiert und anhand von sogenannten Bounding Boxes eingegrenzt.

Abbildung 1: Output YoloV5 Algorithmus

Im nächsten Schritt folgt ein Codeblock (Movement Detector), der mit Hilfe der Koordinaten der Bounding Boxes jeweils den Schispringer in den Bildern erkennt. Die ursprünglichen Bilder

der Serie werden dann jeweils auf die Größe der Bounding Boxes zugeschnitten, dass nur mehr der Schispringer erhalten bleibt. Daraufhin wird im letzten Schritt der OpenPose Algorithmus angewendet, welcher eine Pose über den Athleten legt und somit Keypoints des menschlichen Skeletts markiert. Die Koordinaten (X- und Y-Koordinate) der Keypoints sowie die Wahrscheinlichkeit der Vorhersage werden für alle 21 Bilder der Serie gespeichert. Diesen 1575 somit gewonnen Daten der gesamten Bilderserie wird nun händisch ein Label für Telemark “Ja” oder “Nein” zugeordnet. Somit stehen die Trainingsdaten als Input für das neuronale Netz zur Verfügung.

Abbildung 2: Output OpenPose Algorithmus

Das neuronale Netz wird im Tensorflow Framework mit der integrierten Keras API erstellt. Genutzt wird die Umgebung von Google Colab und die verwendete Programmiersprache ist Python. Der erste Layer ist ein Dense Layer mit 1 575 Units. Durch diesen Layer werden die Daten vorweg aufgeblasen und viele trainierbare Parameter erzeugt. Danach folgt ein eindimensionaler Convolutional Layer mit insgesamt 32 Filtern und je einer Kernel Größe von 75. Diese Größe wird aufgrund der Eigenheit des Datensatzes gewählt, dass ein Bild im Datensatz aus 75 Keypoints besteht. Außerdem wird eine Schrittweite (strides) von 3 gewählt, damit der Filter immer um 3 Dateneinträge weitergeht. Diese Parametrierung erfolgt aus dem Hintergrund, dass im Datensatz jeder Keypoint aus 3 Daten besteht (X-, Y-Koordinate, Wahrscheinlichkeit). Nach diesem Layer wird eine Dropout-Layer mit 0.15 eingefügt.
Es folgt ein weiterer eindimensionaler Convolutional Layer, jedoch mit einer Kernel Größe von 25 und auch einer Schrittweite von 25. Aufgrund der Reduktion der Daten im vorherigen Ablauf sind nun nur noch 25 Features pro Bild vorhanden und am Ende des erneuten Ablaufs wurde das gesamte Datenset auf 21 Keyfeatures reduziert. Hier folgt ein weiterer Dropout Layer mit 0.15, bevor diese 21 Features in einem LSTM Layer bearbeitet werden. Ein Vorteil durch den Nutzen eines LSTM Layers hier ist, dass die Daten gleichzeitig wieder auf eine Dimension geglättet werden. Dadurch kann im Anschluss gleich der Output Layer mit Hilfe eines Dense Layers und einem Unit verwirklicht werden.

Ergebnisse

Zusammenfassend lässt sich sagen, dass eine objektive Entscheidung, ob ein Schispringer eine Telemarklandung vollbringt oder nicht, mit Hilfe von einem neuronalen Netz getroffen

werden kann. Der Prozess der Datenvorbereitung ergibt einen gesamten Verlust von 20.5%, wobei sich der größte Verlust beim selbst erstellten Codeblock des Movement Detector mit 10% einstellt. Die Auswertung des neuronalen Netzes ergibt am Ende eine Validierungsgenauigkeit von 91% und einen Verlust von 30%. Aufgrund der Annahme, dass ein Kampfrichter zu 95% richtig liegt und das neuronale Netz aktuell bei 91% Genauigkeit liegt, müssten jedoch noch einige Verbesserungen vorgenommen werden.

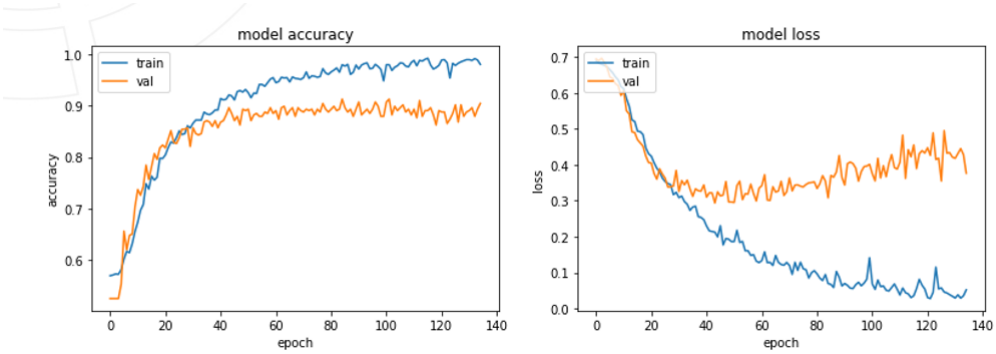

Abbildung 3: Ergebnisse des neuronalen Netzes

Solche Verbesserungen können sowohl in den verwendeten Algorithmen (YOLO, OpenPose, Movement Detector), als auch im neuronalen Netz gefunden werden.

Diskussion

Durch das Entfernen von irrelevanten Keypoints wie Gesichtskeypoints aus dem OpenPose Algorithmus konnten die Ergebnisse des neuronalen Netzes bereits erhöht werden. Dieser Schritt verbesserte die Qualität der Datengrundlage und verringerte die Anzahl von Nullwerten. In diesem Bereich können definitiv weitere Verbesserungen auch auf Basis der Kamerapositionen vorgenommen werden. Alleine durch die verschiedenen Winkel der Kamerapositionen zum Athleten ist eine große Abweichung bei der Datenqualität zu erkennen.
Unter anderem ist auch der Code des Movement Detectors noch hinsichtlich der Konsistenz zu den verschiedenen Kamerawinkeln ausbaufähig. Hier ist durch das Einsetzen von Motion Detection Algorithmen oder ein Anpassen der Berechnungslogik noch einiges an Potential, das wieder die Datenqualität erhöht.
Des Weiteren hat die Erhöhung der Quantität der Daten eine Verbesserung im Netz mit sich gebracht. Durch die Verdoppelung der Trainingsdaten verbesserte sich die Genauigkeit des neuronalen Netzes um weitere 3%.
Im Graph der Ergebnisse ist gut zu erkennen, dass ab einem gewissen Trainingsdurchlauf die Ergebnisse stagnieren. Diese Stagnation kann viele verschiedene Ursachen haben, wie zum Beispiel sogenanntes Overfitting aufgrund zu wenig Trainingsdaten.
Die Ergebnisse unterstreichen, dass die gesteckten Ziele erreicht wurden und dass diese Methode der Datenverarbeitung ein großes Potential für weitere Trainingssteuerungen im Bereich der Haltung beim Schispringen mit sich bringt.

Literatur

Dan Zecha, Christian Eggert, Moritz Einfalt, Stephan Brehm, and Rainer Lienhart. 2018. A Convolutional Sequence to Sequence Model for Multimodal Dynamics Prediction in Ski Jumps. In Proceedings of the 1st International Workshop on Multimedia Content Analysis in Sports (MMSports'18). Association for Computing Machinery, New York, NY, USA, 11–19. https://doi.org/10.1145/3265845.3265855

Lienhart, R., Fichtner, I., Zecha, D., Müller, S., & Kreibich, S. Zeitkontinuierliche Schätzung der Absprung-und Flugparameter von Skispringern in Videoaufnahmen anhand automatischer zeitkontinuierlicher Erken-nung von Körper-und Skiposen. BISp-Jahrbuch, 205.

Heike Brock, Yuji Ohgi, Kazuya Seo, Development of an Automated Motion Evaluation System from Wearable Sensor Devices for Ski Jumping, Procedia Engineering, Volume 147, 2016, Pages 694-699, ISSN 1877-7058, https://doi.org/10.1016/j.proeng.2016.06.248.