

BSc (Hons) Artificial Intelligence And
Data Science
Level 07

CM 4605
Individual Research Project

**CricXpert: A Hybrid Approach Combining Facial and
Spatio-Temporal Gait Analysis for Enhanced Fielder
Recognition with LLM-Based Statistic Generation**

PROJECT PROPOSAL

NADUN SHAMIKA SENARATHNE
IIT ID: 20210488
RGU ID: 2117538

Supervised by: Mr. PRASAN YAPA

Table of Contents

1.1 Tentative Project title	3
1.2 Background and Research Gap	3
<i>1.2.1 Background.....</i>	<i>3</i>
<i>1.2.2 Research gap</i>	<i>3</i>
1.3 Research Questions	4
1.4 Aim & Objectives	4
1.5 Data Requirement	6
1.6 Project Plan	6
1.7 Bibliography	7

1.1 Tentative Project title

CricXpert: A Hybrid Approach Combining Facial and Spatio-Temporal Gait Analysis for Enhanced Fielder Recognition with LLM-Based Statistic Generation

1.2 Background and Research Gap

1.2.1 Background

Cricket has seen considerable developments in gameplay strategies and technology integration, particularly in T20 Internationals as of late. The addition of day/night matches, complicates things more because of the varied lighting conditions, which might affect player vision and gameplay dynamics. Fielders' performances become vital, especially during the crucial final four overs of the game. Thereby, improving fielder recognition by computer vision might significantly increase spectator interest.

Current studies and technologies on player identification are primarily based on singular models such as face or gait recognition (Mahmood et al. 2015), (Haq et al. 2024), (Cai, Wang and Zhang 2023). However, these models are not as efficient concerning dynamic scenarios with partial visibility of fielders, long distance camera angles, unclear gait, and low light environmental conditions. Therefore, by combining multiple recognition technologies this limitation can be mitigated.

In addition, the rising interest in data-driven decision-making in sports has an increased demand for advanced statistical analysis tools. Although existing systems give a wide range of information, obtaining particular player data is usually time-consuming and confusing. To access important player statistics, consumers must usually traverse through several filters and sophisticated interfaces.

1.2.2 Research gap

This project identifies mainly two research gaps:

The primary research gap focuses on the integration of face recognition and hybrid spatio-temporal gait analysis into an ensemble model. This technique provides a viable route for improving identification accuracy in the more challenging conditions found in T20i cricket matches. Existing research has mainly focused on the use of facial recognition for player identification, with notable findings by (Haq et al. 2024) and (Mahmood et al. 2015) as well as a variety of gait analysis approaches such as (Kale et al. 2004) and (Arseev, Konushin, and Liutov 2018). However, these models have not used spatio-temporal features to assess gait in order to recognize a limited set of players – specifically 6 cricket players – during the critical last four overs of T20i matches. This oversight demonstrates a fundamental gap in existing technology's capacity to adapt to cricket's dynamic settings, due to various challenges identified. As mentioned above, such challenges which severely affect the effectiveness of typical identification systems are variable lighting, occlusions, and distant camera angles. Furthermore, while multimodal approaches that combine facial and gait recognition have been previously investigated (Maity, Abdel-Mottaleb and Asfour 2021), (Manssor, Sun, and Elhassan 2021), the use of spatio-temporal features to improve both accuracy and robustness in sports analytics, particularly under the fluctuating conditions of T20i matches, remains novel. Additionally, although prior studies have enhanced deep learning architectures by incorporating machine learning classifiers in contexts like remote sensing image recognition (Özyurt 2020) and brain tumor classification (Kibriya et al. 2021; Kibriya et al. 2022), the proposed ensemble model introduces this approach within the spatial model in a unique context. By applying this novel

combination specifically to player identification in T20i cricket matches, the suggested hybrid approach uniquely tackles these problems, paving the way for a creative solution in sports analytics.

The secondary research gap is pertaining to the use of large language models (LLMs) to understand and translate natural language queries inputted by the user into SQL queries, allowing for the retrieval of player statistics from a relational database. Such an approach has the potential to eliminate the time-consuming, multi-step filtering processes that are common in present systems, as detailed by (ESPN Cricinfo 2024). Previous research, especially that of (Shi, Tang, and Yang 2024), has thoroughly investigated the translation of natural language to SQL using LLMs. However, the present effort aims to improve this procedure further. By allowing users to easily enter queries and instantly receive the necessary statistics, it dramatically minimizes the need for complex filtering and interface navigation, improving the user experience tremendously. This streamlined interaction approach not only simplifies data access, but it also establishes a new standard for user-centric data interfaces in sports analytics.

By addressing these gaps, the project enhances T20i cricket analytics through improved accuracy in player identification and statistical data retrieval. This effort not only fosters greater fan engagement but also contributes to scholarly discourse by merging computer vision and natural language processing within a unique sporting context. These advancements are expected to overcome the existing technological limitations and improve the spectator experience, indicating a big step forward in the context of sports technology.

1.3 Research Questions

Research Question 1: Player Recognition

“How effectively can a computer vision ensemble model employing face recognition and gait analysis using spatio-temporal features, recognise fielders in the outfield during the last four overs of a day/night T20 International cricket match?”

Research Question 2: Statistic Generation

“How can a large language model be prompt-engineered to accurately translate natural language user-defined questions, with up to three conditions, into SQL queries for generating accurate and relevant statistics?”

1.4 Aim & Objectives

Research Question 01:

Aim:

To develop and validate a robust computer vision ensemble model that effectively employs face recognition and spatio-temporal gait analysis for accurate fielder recognition during the critical final four overs of T20 International cricket matches.

Objectives:

1. **Develop an Ensemble Model:** Develop a model that can precisely identify fielders by utilizing both hybrid spatio-temporal gait analysis and facial recognition. The goal

of this model is to capture both the static and dynamic attributes of players by integrating multiple analytical techniques.

2. **Data Collection and Preprocessing:** Compile an extensive dataset consisting of images and video feeds that depict various lighting conditions and field settings that are common in T20i matches. In order to prepare for efficient model training, this data will undergo preprocessing to normalize variations.
3. **Feature Engineering:** Determine the key elements that have the biggest influence on player identification.
4. **Model Training and Optimization:** Train the ensemble model using advanced machine learning techniques. Optimize the model for high accuracy and operational efficiency in dynamic conditions.
5. **Validation under Various Conditions:** Analyze the model's performance in various scenarios to make sure it is robust and reliable. This entails testing in various lighting conditions, with players moving in various directions, and at various field positions.

Research Question 02:

Aim:

To harness the capabilities of a prompt-engineered LLM for translating complex natural language queries into precise SQL queries, thereby enabling the generation of accurate and contextually relevant statistics for recognized players in T20 International cricket matches.

Objectives:

1. **Selection of an Optimal Large Language Model (LLM):** Choose the most suitable LLM from available options such as ChatGPT, Gemini, and Llama, based on their capabilities to understand and process natural language queries effectively.
2. **Design of a Relational Database:** Build a relational database that is both scalable and reliable for storing comprehensive player statistics. In order to facilitate the retrieval of specific player data as needed, this database should support effective SQL queries.
3. **Implementation of Prompt Engineering Techniques:** Apply and refine prompt engineering techniques to train the selected LLM on accurately interpreting natural language queries. Optimize these prompts to improve the LLM's ability to formulate SQL queries that are both syntactically correct and logically consistent with user queries.

4. **Evaluation of Input Character Length on Effectiveness of User Conditions:**
Assess how different user condition configurations and the lengths of input characters affect the relevance and accuracy of the generated statistics.

1.5 Data Requirement

Fielder Facial Data:

- **Purpose:** To train the model for facial recognition of fielders.
- **Collection Method:** A custom dataset will be created, consisting of images captured of fielders during various matches which were extracted from Google Images.

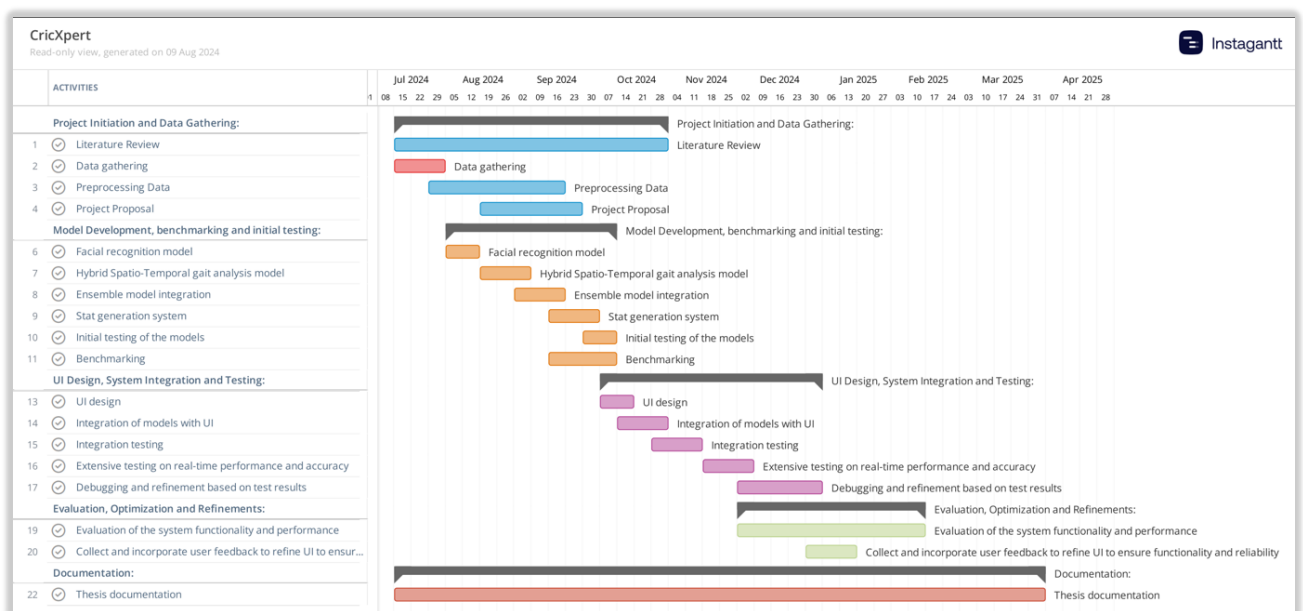
Fielder Gait Data:

- **Purpose:** To analyze and recognize fielders based on their movement patterns.
- **Collection Method:** This dataset will also be custom-built, involving the collection of spatio-temporal gait data from video footage.

Fielder Stat Data:

- **Purpose:** To provide statistical records for the recognized fielders.
- **Collection Method:** This data will be extracted from the [ESPN](https://www.espn.com) website and organized into a custom dataset.

1.6 Project Plan



1.7 Bibliography

- HAQ, M.U. et al., 2024. Automatic Player Face Detection and Recognition for Players in Cricket Games. *IEEE Access*, PP, pp. 1–1.
- MAHMOOD, Z. et al., 2015. Automatic player detection and identification for sports entertainment applications. *Pattern Analysis and Applications*, 18(4), pp. 971–982.
- BANOTH, T. et al., 2022. *A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions*.
- ZHANG, R. et al., 2020. Multi-camera Multi-player Tracking with Deep Player Identification in Sports Video. *Pattern Recognition*, 102, p. 107260.
- INSAAF, A. et al., 2020. Past, Present, and Future of Face Recognition: A Review. *Electronics*, 9, p. 1188.
- ÖZYURT, F., 2020. Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures. *The Journal of Supercomputing*, 76, pp. 1–19.
- KIBRIYA, H. et al., 2022. A Novel and Effective Brain Tumor Classification Model Using Deep Feature Fusion and Famous Machine Learning Classifiers. *Computational Intelligence and Neuroscience*, 2022(1), p. 7897669.
- KIBRIYA, H. et al., 2021. Multiclass Brain Tumor Classification Using Convolutional Neural Network and Support Vector Machine. In: *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*. 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), July 2021. pp. 1–4.
- ZHEN, T., KONG, J. and YAN, L., 2020. Hybrid Deep-Learning Framework Based on Gaussian Fusion of Multiple Spatiotemporal Networks for Walking Gait Phase Recognition. *Complexity*, 2020, pp. 1–17.
- DONG, Y. et al., 2023. HybridGait: A Benchmark for Spatial-Temporal Cloth-Changing Gait Recognition with Hybrid Explorations. Available from: <http://arxiv.org/abs/2401.00271> [Accessed 9 Aug 2024].
- MOGAN, J.N. et al., 2022. Gait-DenseNet: A Hybrid Convolutional Neural Network for Gait Recognition. | IAENG International Journal of Computer Science | EBSCOhost.
- ZHEN, T., YAN, L. and KONG, J., 2020. An Acceleration Based Fusion of Multiple Spatiotemporal Networks for Gait Phase Detection. *International Journal of Environmental Research and Public Health*, 17(16), p. 5633.
- YAO, L. et al., 2019. Robust Gait Recognition using Hybrid Descriptors based on Skeleton Gait Energy Image. *Pattern Recognition Letters*, 150.
- SINGH, J., SINGH, Dr.U. and JAIN, S., 2023. Model-based person identification in multi-gait scenario using hybrid classifier. *Multimedia Systems*, pp. 1–14.
- JUN, K. et al., 2023. Hybrid Deep Neural Network Framework Combining Skeleton and Gait Features for Pathological Gait Recognition. *Bioengineering (Basel, Switzerland)*, 10(10), p. 1133.
- LI, J. et al., 2023. Gaitcotr: Improved Spatial-Temporal Representation for Gait Recognition with a Hybrid Convolution-Transformer Framework. In: *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.

- GUL, S. et al., 2021. Multi-view gait recognition system using spatio-temporal features and deep learning. *Expert Systems with Applications*, 179, p. 115057.
- MATHIVANAN, B. and PERUMAL, P., 2022. Gait Recognition Analysis for Human Identification Analysis-A Hybrid Deep Learning Process. *Wireless Personal Communications*, 126(1), pp. 555–579.
- HUANG, X. et al., 2022. STAR: Spatio-Temporal Augmented Relation Network for Gait Recognition. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, PP, pp. 1–1.
- KHAN, M., AZAM, H. and FARID, M.S., 2023. Automatic multi-gait recognition using pedestrian's spatiotemporal features. *The Journal of Supercomputing*, 79, pp. 1–23.
- SALEEM, F. et al., 2022. Human Gait Recognition: A Single Stream Optimal Deep Learning Features Fusion. *Sensors*, 21.
- TIAN, H. et al., 2021. Skeleton-based Abnormal Gait Recognition with Spatio-temporal Attention Enhanced Gait-structural Graph Convolutional Networks. *Neurocomputing*, 473.
- SUN, X., WANG, Y. and KHAN, J., 2023. Hybrid LSTM and GAN model for action recognition and prediction of lawn tennis sport activities. *Soft Computing*, 27(23), pp. 18093–18112.
- GERATS, B. et al., 2021. *Individual Action and Group Activity Recognition in Soccer Videos from a Static Panoramic Camera*.
- KALE, A. et al., 2004. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13(9), pp. 1163–1173.
- MEHDIZADEH, S. et al., 2021. Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: a preliminary analysis with multiple trackers, viewing angles, and walking directions. *Journal of NeuroEngineering and Rehabilitation*, 18.
- HULLECK, A.A. et al., 2023. *Accuracy of Computer Vision-Based Pose Estimation Algorithms in Predicting Joint Kinematics During Gait*.
- STENUM, J., ROSSI, C. and ROEMMICH, R., 2020. *Two-dimensional video-based analysis of human gait using pose estimation*.
- SABIR, A., AL-JAWAD, N. and JASSIM, S., 2013. Gait recognition using spatio-temporal silhouette-based features. In: *Mobile Multimedia/Image Processing, Security, and Applications 2013*. Mobile Multimedia/Image Processing, Security, and Applications 2013, 28 May 2013. SPIE. pp. 194–203.
- DONG, Y. and NOH, H., 2024. *Ubiquitous Gait Analysis through Footstep-Induced Floor Vibrations*.
- SIMONI, L. et al., 2021. Quantitative and Qualitative Running Gait Analysis through an Innovative Video-Based Approach. *Sensors (Basel, Switzerland)*, 21.
- VITECKOVA, S. et al., 2020. Gait symmetry methods: Comparison of waveform-based Methods and recommendation for use. *Biomedical Signal Processing and Control*, 55, p. 101643.
- NAFEA, O. et al., 2021. Sensor-Based Human Activity Recognition with Spatio-Temporal Deep Learning. *Sensors*, 21, p. 2141.
- MAITY, S., ABDEL-MOTTALEB, M. and ASFOUR, S.S., 2021. Multimodal Low Resolution Face and Frontal Gait Recognition from Surveillance Video. *Electronics*, 10(9), p. 1013.

- MANSSOR, S.A.F., SUN, S. and ELHASSAN, M.A.M., 2021. Real-Time Human Recognition at Night via Integrated Face and Gait Recognition Technologies. *Sensors (Basel, Switzerland)*, 21(13), p. 4323.
- PRAKASH, A. et al., 2023. Multimodal Adaptive Fusion of Face and Gait Features using Keyless attention based Deep Neural Networks for Human Identification.
- CAI, M., WANG, M. and ZHANG, S., 2023. Gait Recognition by Jointing Transformer and CNN. In: W. JIA et al., eds. *Biometric Recognition*. Singapore: Springer Nature. pp. 312–321.
- FAN, C. et al., 2023. *Exploring Deep Models for Practical Gait Recognition*.
- CATRUNA, A., COSMA, A. and RADOI, E., 2024. GaitPT: Skeletons Are All You Need For Gait Recognition.
- LIAO, R. et al., 2020. A model-based gait recognition method with body pose and human prior knowledge. *Pattern Recognition*, 98, p. 107069.
- RANI, V. and KUMAR, M., 2023. Human gait recognition: A systematic review. *Multimedia Tools and Applications*, 82(24), pp. 37003–37037.
- (Lee and Grimson 2002) LEE, L. and GRIMSON, W.E.L., 2002. Gait analysis for recognition and classification. In: *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, May 2002. pp. 155–162.
- ARSEEV, S., KONUSHIN, A. and LIUTOV, V., 2018. Human Recognition by Appearance and Gait. *Programming and Computer Software*, 44(4), pp. 258–265.
- CHOPRA, A. and AZAM, R., 2024. Enhancing Natural Language Query to SQL Query Generation Through Classification-Based Table Selection. In: L. ILIADIS et al., eds. *Engineering Applications of Neural Networks*. Cham: Springer Nature Switzerland. pp. 152–165.
- ZHANG, Q. et al., 2024. Structure Guided Large Language Model for SQL Generation.
- LEE, D. et al., 2024. MCS-SQL: Leveraging Multiple Prompts and Multiple-Choice Selection For Text-to-SQL Generation.
- HONG, Z. et al., 2024. Knowledge-to-SQL: Enhancing SQL Generation with Data Expert LLM.
- GUO, C. et al., 2024. Prompting GPT-3.5 for Text-to-SQL with De-semanticization and Skeleton Retrieval. In: F. LIU et al., eds. *PRICAI 2023: Trends in Artificial Intelligence*. Singapore: Springer Nature. pp. 262–274.
- YI, J., CHEN, G. and SHEN, Z., 2024. RH-SQL: Refined Schema and Hardness Prompt for Text-to-SQL.
- NAN, L. et al., 2023. Enhancing Few-shot Text-to-SQL Capabilities of Large Language Models: A Study on Prompt Design Strategies.
- HONG, Z. et al., 2024. Next-Generation Database Interfaces: A Survey of LLM-based Text-to-SQL.
- LI, H. et al., 2024. CodeS: Towards Building Open-source Language Models for Text-to-SQL.
- TALAEI, S. et al., 2024. CHESS: Contextual Harnessing for Efficient SQL Synthesis.
- GUO, C. et al., 2023. Retrieval-augmented GPT-3.5-based Text-to-SQL Framework with Sample-aware Prompting and Dynamic Revision Chain.

RAJKUMAR, N., LI, R. and BAHDANAU, D., 2022. Evaluating the Text-to-SQL Capabilities of Large Language Models.

SHI, L., TANG, Z. and YANG, Z., 2024. A Survey on Employing Large Language Models for Text-to-SQL Tasks.