Rugby Defense Strategy Prediction Using Deep Learning

M.A.V Lochana
2020/CS/105
University of Colombo School of Computing

June 30, 2025

Supervisor: Mr. P K M Thilakarathna
Co-Supervisor: Dr. L N C De Silva External Advisor: Dr. Sachintha Pitigala

Submitted in partial fulfillment of the requirements of
the Bachelor of Science (Hons.) in Computer Science

Declaration

To the best of my knowledge and belief, this dissertation does not contain any material that has been previously published or written by me or by any other person, except where due reference is made in the text. It also does not include any material that has been submitted for a degree or diploma at any university.

I further agree that the title and abstract of this dissertation, if approved, may be made available to external organizations, and may be released for photocopying and inter-library loans.

Candidate: M.A.V. Lochana

....

Signature

Supervisor's Certification
This is to certify that Mr. M.A.V. Lochana worked under my supervision on this dissertation. The thesis has been prepared in accordance with the required format and is of an acceptable academic standard.

Supervisor: Mr. P. K. M. Thilakarathna

30/06/2025

Signature

Co-Supervisor's Certification
This is to certify that Mr. M.A.V. Lochana worked under my supervision on this dissertation. The thesis has been prepared in accordance with the required format and is of an acceptable academic standard.

Co-Supervisor: Dr. L. N. C. De Silva

Signature

External Advisor's Certification
This is to certify that Mr. M.A.V. Lochana worked under my guidance as an external advisor on this dissertation. The thesis has been prepared in accordance with the required format and is of an acceptable academic standard.

External Advisor: Dr. Sachintha Pitigala

Acknowledgements

I would like to express my deepest gratitude to the University of Colombo School of Computing (UCSC) for providing me with the opportunity and essential resources to pursue my research.

I am sincerely thankful to my supervisor, Mr. P.K.M. Thilakarathna, for his invaluable guidance, knowledge, and encouragement throughout this project. His contributions have been instrumental in shaping the final outcome. I am also deeply appreciative of Dr. L.N.C. De Silva for her essential perspectives and assistance as my co-supervisor.

My heartfelt thanks go to Dr. Sachintha Pitigala, who served as my external advisor, for his insightful feedback and support, which have been extremely beneficial in navigating the complexities of this research.

I am grateful to my fellow batch mates who assisted in preparing the speech dataset for testing, which was crucial to the success of this research.

Finally, I would like to express my heartfelt gratitude to all individuals who participated in the evaluation process and offered valuable suggestions and criticism, greatly contributing to the improvement of this research.

Abstract

The integration of data analytics in sports has revolutionized performance evaluation and strategic planning, yet rugby union remains underrepresented in this domain, particularly concerning defensive strategies. This thesis addresses this gap by developing a deep learning framework to classify and evaluate three prevalent defensive formations in rugby union: blitz, drift, and umbrella. Utilizing manually annotated positional data from Rugby World Cup matches, the study employs Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and hybrid CNN-LSTM architectures to capture the spatio-temporal dynamics of defensive plays. The research further incorporates success prediction as a secondary task, assessing the effectiveness of each defensive strategy. To enhance model generalization, rugby-specific data augmentation techniques, including coordinated jittering and mirroring, are applied.Experimental results show that the hybrid CNN-LSTM model achieved the highest performance, reaching an overall accuracy of 97.22%, with strong strategy recognition and success prediction capabilities across augmented datasets.The findings demonstrate the potential of deep learning models to automate the classification of defensive formations and predict their success, offering valuable insights for coaches and analysts. This work contributes to the advancement of rugby analytics by introducing a scalable, objective approach to defensive strategy evaluation.

Contents

List of Figures

Chapter 1

Introduction

This chapter aims to introduce the study by providing comprehensive background and contextual information on the research domain. The research problem, objectives, and questions will be discussed, and the significance of the research will be highlighted accordingly.

1.1 Background

Rugby is widely considered a high intensity, contact sport that requires a combination of physical strength, tactical intelligence, and strategic decision making. Unlike many other field sports that feature regular stoppages or set defensive plays (e.g., American football), rugby maintains continuous play. This continuity demands that teams quickly adapt their strategies based on real time game dynamics. Among these, defense plays a critical role in controlling the flow of the match, disrupting the opponent's attacks, and creating opportunities for turnovers.

Although the crowd often focuses on offensive plays, the effectiveness of a team's defense is equally important in determining match outcomes. An organized defense not only prevents line breaks but also disrupts the momentum of the offensive team and creates turnover opportunities. In this sense, defense in rugby is not merely about stopping the opposition, but about influencing the flow of the game, dictating territory, and controlling the momentum .

1.1.1 Key Defensive Strategies

For the purposes of this research, three broad defensive strategies are primarily considered:

- **Blitz Defense**
  The blitz defense, sometimes called "up and in" defense, relies on rapid line speed to cut down the offensive team's space and force errors. Defenders launch forward as soon as the ball is passed to disrupt the attacking line.
- **Drift Defense**
  In a drift defense, defenders move laterally in response to the attacking line's advance. This strategy aims to effectively corner the offense towards the sidelines by using the touchline as an additional defender.
- **Umbrella/Hinge Defense**
  The umbrella (or hinge) defense is a hybrid approach that combines elements of both blitz and drift defenses. It involves a coordinated push by the central defensive line, while wide defenders remain deeper, thereby funneling the attack toward the supporting defenders.

1.1.2 Importance of Defensive Strategy Analysis in Rugby

Despite rugby's physical intensity and strategic complexity, many teams still rely heavily on manual video reviews to study their defense (Jones et al. 2019; O'Donoghue 2005). Coaches typically watch footage frame by frame to spot missed tackles or poor alignment. While these reviews can be useful, they are time-consuming and prone to subjective bias, as different coaches may notice different details (Quinn 2017). In contrast, sports such as basketball, soccer, and American football have embraced advanced analytics—using real-time tracking and predictive models to sharpen both offense and defense (Nix and Coauthor 2014; Quinn 2017). Rugby, however, has lagged behind in the use of in depth data, particularly for defensive analysis, where few established methods exist to measure how well different defensive systems work (Jones et al. 2019; Smith 2019). Although technologies like GPS tracking and wearable sensors are emerging, they are expensive and require specialized staff to manage (Shaw and Coutts 2020).

1.2 Motivation

Rugby has evolved from being primarily a physical contest to a sport where strategy and tactical planning can decisively influence the outcome (Ford and Williams 2008; Memmert, Lemmink, and Sampaio 2017). While machine learning (ML) has been applied in many sports to predict player performance, forecast match results, and optimize offensive tactics, its application to defensive strategies is still relatively new (Jon Gudmundsson and M. Horton 2017; Watson et al. 2020). Defense is crucial not only for controlling the match's tempo but also for shifting momentum through strategic turnovers (Jones et al. 2019).

By applying ML to real game data, coaches and analysts can gain a better understanding of how specific defensive approaches—such as blitz, drift, or umbrella defenses perform in various situations. This data-driven perspective can improve the accuracy of defensive plans and enable teams to adapt quickly when conditions change (M. Johnson 2021). Moreover, analyzing defense through ML could reveal insights into spatial patterns and player movements that may not be obvious through manual observation (Shaw and Coutts 2020). Ultimately, employing ML in rugby defense has the potential not only to enhance performance but also to pioneer analytical methods beneficial to other sports where defense plays a critical role (McCarville 2022).

1.3 Problem Statement

Despite the increased complexity in Rugby, the application of advanced ML techniques to predict defensive strategies is still relatively unexplored. Coaches have traditionally relied on experience and basic statistics to guide their defensive tactics, but these methods often fail to capture the full complexity of the game, especially in anticipating and countering opponents' moves.

Most rugby analytics to date have concentrated on isolated actions or general game outcomes, leaving a gap in understanding the specific dynamics of defensive plays. Additionally, there is a shortage of publicly available datasets that detail defensive formations and player interactions, making it challenging to develop models that accurately predict rugby defenses. Without data valuable insights are missed that could help coaches make informed defensive decisions.

This research develops a machine learning model that classifies defensive strategies based on early attack movements and predicts their likelihood of success, thereby enabling data driven evaluation of tactical effectiveness in rugby. Although the model does not directly recommend an optimal strategy, it provides insight into the effectiveness of each formation given early game context laying the foundation for future real time decision support systems. To achieve this, we will compile a focused dataset from game footage, capturing key instances of defensive patterns, and use ML techniques to analyze how player formations and movements impact defensive effectiveness. By following this approach, the study aims to provide rugby teams with practical, data-driven insights that can improve their defensive tactics and enhance their competitive edge on the field.

1.4 Research Aim, Questions and Objectives

1.4.1 Research Aim

This research aims to develop a predictive model that can identify the defensive strategy (Blitz, Drift, or Hinge) employed during the early phase of a rugby play and estimate its likelihood of success. By leveraging deep learning techniques—specifically Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and hybrid CNN-LSTM architectures—the model analyzes spatio-temporal patterns in player and ball movement to classify defensive tactics and predict their effectiveness. While the system does not directly recommend an optimal strategy, it enables data-driven evaluation of defensive formations, offering valuable insights that can inform tactical decisions. Ultimately, this research contributes toward bridging the gap between manual analysis and AI-driven rugby strategy modeling, laying the groundwork for future real-time decision-support tools for coaches and analysts.

1.4.2 Research Questions

Main Research Question: How can machine learning be used to classify defensive strategies in rugby—specifically Blitz, Drift, and Umbrella defenses—and predict their effectiveness using early-phase spatio-temporal player movement data?

Sub-questions:

- What spatial and temporal factors—such as player positions, formations, and contextual variables (e.g., field location, game phase)—most strongly influence the classification of defensive strategies?
- How can a structured and labeled dataset of rugby defensive plays be created from match footage, ensuring consistency in spatial-temporal annotation across multiple instances?
- What deep learning techniques are most effective for modeling rugby defensive strategies based on spatio-temporal data? Should architectures such as LSTM, CNN, or CNN-LSTM hybrids be used to capture sequential and spatial dependencies?
- How can the predictive performance of the model be rigorously evaluated? What metrics (e.g., strategy accuracy, success prediction accuracy, joint accuracy) and validation methods (e.g., expert review, test set analysis) provide meaningful insights into its real-world applicability?

1.5 Outline of the Dissertation

This dissertation is structured into five chapters, each designed to progressively build the foundation, execution, and insights derived from this study on rugby defensive strategy analysis using deep learning.

- **Chapter 1: Introduction** — Introduces the context and motivation for the study, highlighting the challenges in analyzing rugby defense through traditional methods. It outlines the research problem, objectives, research questions, and the overall significance of adopting a data-driven approach in tactical rugby analytics.
- **Chapter 2: Literature Review** — Provides a comprehensive survey of existing research in sports analytics, with a particular focus on team defense modeling in rugby and other sports such as soccer, basketball, and American football. The chapter also discusses the capabilities of deep learning models—such as LSTMs and CNNs—in capturing spatial-temporal patterns and concludes by identifying the research gaps this thesis aims to address.
- **Chapter 3: Methodology** — Describes the methodological approach of the study, including the construction of a manually annotated rugby dataset, spatio-temporal preprocessing techniques, data augmentation strategies, and the design of three model architectures: LSTM, 2D CNN, and CNN-LSTM hybrid. It also explains the multi-task learning framework for simultaneous strategy classification and success prediction, as well as the rationale behind the selected hyperparameters and architectural choices.
- **Chapter 4: Experiments and Results** — Details the experimental setup and presents results obtained from each of the proposed models across multiple dataset sizes. Evaluation metrics include strategy accuracy, success prediction accuracy, and overall task accuracy. This chapter also includes a comparative performance analysis and visualizations such as training curves to support findings.
- **Chapter 5: Discussion and Conclusion** — Interprets the significance of the experimental results in relation to the research objectives. The discussion addresses the strengths and limitations of the proposed models and methods, including the role of data augmentation and model architecture in performance. The chapter

concludes with a summary of key contributions and suggests future directions for real-time integration and broader application of data-driven rugby analytics.

1.6 Scope

1.6.1 In scope

- This research will focus exclusively on Rugby defense, specifically analyzing the drift, blitz, and umbrella defensive strategies using data from the last 2 world cups(2019 and 2023). These are among the most widely used and structured defenses in rugby, making them ideal for study and predictive modeling.

Data Selection

1.7 Summary of the Chapter

This chapter introduced the research by outlining the significance of defensive strategy analysis in rugby union and highlighting the limitations of existing methods. It emphasized the need for automated, data-driven systems to enhance tactical decision-making in defense. The research problem was defined, and the motivation for applying deep learning methods was established. The chapter also presented the research questions, objectives, and scope, setting the foundation for the subsequent chapters.

Chapter 2

Literature review

2.1 Introduction to Sports Analytics in Rugby

The application of data analytics in sports has undergone rapid development over the past two decades, fundamentally transforming how teams and analysts understand performance, strategy, and decision-making. With the rise of player tracking systems, wearable technologies, and machine learning techniques, analytics has become an integral component of modern sports science, particularly in team sports like soccer, basketball, and American football (Jon Gudmundsson and M. Horton 2017; Bunker and Thabtah 2019).

In rugby union, the use of analytics has traditionally centered around *descriptive statistics* and *manual video analysis*, offering valuable but limited insights into performance. Much of the existing literature focuses on *offensive structures*, *attacking patterns*, and *ball progression metrics*. For instance, Passos et al. (2011) studied attacker-defender interactions using positional data to understand spatial decision-making during attacking phases. Similarly, Croft, Lamb, and Thewlis (2016) applied network analysis to evaluate passing efficiency in attacking plays. More recently, Clarke et al. (2021) developed models using GPS and event data to predict the success of offensive sequences in multi-phase play.

In contrast, *defensive strategy analysis remains underexplored.* Most studies have focused on isolated defensive actions such as tackling efficiency (Read et al. 2015) or physical workload analysis (Hughes and Bartlett 2002), with limited attention paid to the *collective behavior of defensive units* or the *classification of team-level formations*.

Given the complex, multi-phase nature of rugby defense, there is a growing need for *automated, data-driven systems* capable of capturing both *spatial and temporal* aspects of defensive strategy. The integration of deep learning into rugby analytics presents a promising avenue for addressing this gap. Models such as *Long Short-Term Memory (LSTM)* networks can capture sequential patterns over time, while *Convolutional Neural Networks (CNNs)* are effective at learning spatial features from positional data. Hybrid models, such as *CNN-LSTM architectures*, can jointly model the evolution of player formations and the outcome of defensive phases, offering a more holistic understanding of tactical behavior.

This thesis addresses the above research gap by developing a *deep learning-based framework* that classifies rugby defensive strategies—namely *blitz*, *drift*, and *umbrella* formations—and predicts their success using *annotated positional data* extracted from Rugby World Cup matches. By leveraging both spatial and temporal cues through *LSTM, CNN, and CNN-LSTM models*, the proposed system aims to enhance the objectivity and precision of rugby defensive analysis.

2.2 Defensive Analytics in Other Team Sports

While the application of data-driven analysis in rugby defense is still emerging, other team sports have made significant advances in modeling defensive structures, player coordination, and tactical decision-making. In particular, sports such as soccer, basketball, and American football have leveraged tracking data, spatial-temporal modeling, and machine learning techniques to better understand and predict defensive effectiveness. These developments provide a foundation for adopting similar analytical frameworks in rugby union, especially for evaluating team-level defensive strategies.

2.2.1 Soccer

Soccer has been at the forefront of integrating spatial and temporal analytics to examine both offensive and defensive phases of play. One major area of development is in the automatic recognition and classification of team formations. Andrzej Bialkowski et al. (2014) introduced a framework that utilized player tracking data to identify and analyze team shape and formation during defensive phases. Their method involved clustering spatial configurations and using supervised learning to classify different defensive setups, offering an automated way to detect structural adaptations during a match.

Building on such frameworks, Joakim Gudmundsson and Wichert (2017) emphasized metrics such as team compactness, player density, and defensive line coordination. These indicators were derived from spatio-temporal features and enabled analysts to quantify the effectiveness of pressing strategies—particularly how tightly a defensive unit maintains its shape while attempting to disrupt an opponent's buildup. By integrating contextual factors such as ball location and player roles, these models not only classified formations but also predicted defensive success based on dynamic pressure.

2.2.2 Basketball

Basketball has benefited significantly from optical tracking systems such as SportVU and Second Spectrum, which provide continuous spatial data for all players on the court. This has enabled the development of advanced defensive metrics, including defender proximity, contest pressure, and defensive rotations. Goldsberry (2012) introduced spatial visualizations (e.g., CourtVision) to evaluate shot defense efficiency, helping teams understand how positioning impacts scoring probabilities.

Expanding on this, Miller, Bornn, and Goldsberry (2019) used machine learning models to predict defensive outcomes based on player trajectories and ball movement. Their study highlighted how effective defensive plays often depend on coordinated movement and role-based responses rather than individual effort alone. One key analytical advance in basketball has been the classification of *man-to-man* versus *zone defense* schemes. These models leveraged player distance matrices, movement vectors, and defensive assignments to recognize team tactics and inform real-time adjustments.

2.2.3 American Football

In American football, the use of player tracking data has enabled a detailed understanding of defensive alignments and their relationship to play outcomes. Levy, Bialik, and Lopez (2016) applied machine learning techniques to predict defensive performance based on player positioning, acceleration patterns, and coverage depth. Their models accounted for both pre-snap alignment and post-snap behavior, capturing how coordinated movements across a defensive unit impacted the likelihood of successful pass coverage or pressure generation.

Other studies have emphasized the importance of combining spatial features with contextual game information—such as down and distance—to improve the prediction of defensive outcomes. The use of integrated spatial-temporal models has allowed for the evaluation of defensive schemes such as *zone blitz*, *cover-2*, or *man-free* based on the real-time positioning of defensive backs and linemen.

Taken together, these studies demonstrate that spatial-temporal analytics and machine learning are powerful tools for modeling and optimizing defensive strategies in team sports. While soccer, basketball, and American football have successfully adopted these methods, rugby union has yet to realize their full potential. This presents a unique opportunity to adapt and extend these techniques to the domain of rugby defense, particularly given the sport's fluid nature and multi-phase structure.

2.3 Studies on Rugby

While other team sports have made significant advances in modeling both offensive and defensive strategies through spatial-temporal analytics and machine learning, research in rugby union has been comparatively limited. Within rugby analytics, the existing body of work leans heavily toward the analysis of offensive patterns and physical performance metrics, with relatively little attention given to automated modeling of team-level defensive structures. This section outlines the current landscape of rugby analytics research by contrasting offensive and defensive studies and identifying the methodological gaps that motivate this thesis.

2.3.1 Offensive-Focused Work

A growing number of studies have applied data-driven techniques to examine offensive behavior in rugby. Passos et al. (2011) explored attacker-defender interactions using positional data to analyze interpersonal coordination during attacking phases. Their study emphasized the role of spatial dynamics and decision-making under pressure, providing a foundational understanding of attacking subunit behavior.

Building on this work, Croft, Lamb, and Thewlis (2016) introduced passing network analysis as a tool to evaluate offensive efficiency. By modeling players as nodes and passes as edges, their study demonstrated how network centrality and ball movement patterns could offer insights into attacking performance and team cohesion.

More recently, Clarke et al. (2021) developed predictive models that combine GPS tracking and event data to forecast the likelihood of attacking success. Their approach integrated multi-phase contextual features such as player speed, support lines, and territory gain, showcasing how machine learning could be used to evaluate and optimize offensive strategies in elite rugby union.

2.3.2 Defensive-Focused Work

In contrast to the progress made in offensive analytics, defensive strategy research in rugby remains largely descriptive. Hughes and Bartlett (2002) conducted a time-motion analysis to assess player workload across positions, providing valuable insights into the physical demands of defense, but without modeling formation or tactical coordination.

Read et al. (2015) examined tackle outcomes in professional rugby union and their relationship to defensive line integrity. While their work highlighted key physical performance indicators such as tackle success rate, it focused on individual events rather than the collective structure or evolution of defensive systems.

Additionally, Dunning (2010) explored how contextual variables—such as field position, score differential, and phase of play—affect tactical decision-making. Although this study acknowledges that situational factors influence defensive choices, it does not provide a framework for identifying or predicting the defensive strategies employed.

Collectively, these studies illustrate that most defensive analyses in rugby are based on *isolated actions* and lack an integrated, system-level view of team formations and transitions during defensive phases.

2.4 Deep Learning Approaches for Tactical Analysis

The complexity and fluidity of team sports such as rugby make deep learning a compelling tool for tactical modeling. Unlike traditional machine learning methods that rely on hand-engineered features and static snapshots, deep learning models can automatically learn spatial and temporal patterns from raw data. In the context of rugby defense, where team formations evolve across time and space, architectures such as Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and hybrid CNN-LSTM models offer strong potential for capturing tactical behavior. This section outlines the rationale for selecting these architectures in this thesis.

2.4.1 Long Short-Term Memory (LSTM)

LSTM networks are a type of recurrent neural network (RNN) designed to model sequential dependencies over time (Hochreiter and Schmidhuber 1997a). They are particularly effective in scenarios where context and temporal transitions play a crucial role, such as in speech recognition, natural language processing, and time-series forecasting.

In sports analytics, LSTMs have been used to model game events as sequences, enabling the prediction of future actions based on prior patterns (Hien Le et al. 2017). For example, Wei, Sha, and Lucey (2016) applied LSTM models to basketball data to predict player movement and possession outcomes. The ability of LSTMs to learn from sequential inputs makes them well-suited for modeling rugby defensive phases, where the positioning of players changes fluidly across time and is influenced by prior movements within a phase.

2.4.2 Convolutional Neural Networks (CNN)

CNNs are widely used in computer vision tasks due to their ability to learn spatial hierarchies of features through convolutional filters (Krizhevsky, Sutskever, and Geoffrey E. Hinton 2012). In sports, CNNs have been adapted to learn spatial relationships between players and key field areas by encoding positional data into image-like input representations.

For instance, Zheng, Yue, and Lucey (2016) demonstrated how CNNs could be used to identify patterns in soccer formations by treating player coordinates as pixel values in a spatial heatmap. Similarly, CNN-based architectures have been used in basketball to assess defensive coverage based on proximity and spatial alignment (Miller, Bornn, and Goldsberry 2019). In the context of rugby defense, CNNs can be employed to capture the formation structure at each timestep, enabling the model to learn spatial cues such as line compactness, gaps, and staggered alignments.

2.4.3 CNN-LSTM Hybrid Models

Hybrid CNN-LSTM architectures combine the strengths of both temporal and spatial modeling by stacking a CNN layer to extract spatial features from each timestep, followed by an LSTM layer to capture the sequence of those features over time (Shi et al. 2015b). This combination is particularly effective in tasks that involve spatio-temporal reasoning, such as action recognition in videos or player trajectory prediction.

In sports analytics, CNN-LSTM models have been applied to sequence-based tasks such as play classification and outcome prediction (Shao et al. 2020). Their ability to jointly model spatial structure and temporal transitions makes them ideal for rugby defense, where both formation layout and time-dependent evolution matter. By using CNN-LSTM models, this thesis aims to capture not only the shape of the defense at each moment but also how it evolves across frames in a defensive phase.

Given the dual spatial-temporal nature of rugby, and the multi-phase flow of the game, CNN-LSTM models offer a promising architecture for both strategy classification and success prediction.

2.5 Gaps in Existing Research

Despite growing interest in sports analytics and significant advances in modeling offensive play, the analysis of rugby defense remains notably underdeveloped. Existing studies have primarily relied on *manual coding* and *subjective assessments* by coaches, which limits both scalability and reproducibility. Defensive analyses often focus on isolated actions—such as tackles or workloads—rather than modeling the coordinated behavior of the defensive unit over time (Read et al. 2015; Hughes and Bartlett 2002).

While there has been prior application of deep learning in rugby analytics, such as the work by Watson et al. 2020, which employed convolutional and recurrent neural networks to predict outcomes of sequences of play, these studies have primarily focused on offensive metrics and general play outcomes. To date, there is a lack of research specifically targeting the automated classification of team-level defensive strategies—namely blitz, drift, and umbrella formations—using deep learning techniques. Furthermore, existing studies have not integrated success prediction within the context of these defensive structures. This thesis aims to fill this gap by developing a deep learning-based framework that not only classifies these defensive strategies but also predicts their effectiveness, thereby providing a more comprehensive tool for tactical analysis in rugby union.

Another barrier is the lack of publicly available, *labeled datasets* capturing player and ball trajectories across defensive phases. In contrast to the structured datasets available in soccer (e.g., StatsBomb, SoccerNet) and basketball (e.g., NBA SportVU), rugby lacks annotated repositories for team-level defensive analysis, hindering model development and reproducibility.

Moreover, prior work has not addressed the task of *predicting the success* of defensive strategies within the context of team formations. No existing models explicitly link structural behavior during defensive phases to outcomes such as gainline success, especially using spatial-temporal data.

Table 2.1: Gap Analysis Summary in Rugby Defensive Analytics

Area	**Current State in Literature**	**Identified Gap**	**This Thesis Contribution**
Defensive Strategy Modeling	Analyses rely on manual coding and subjective assessments by coaches (Read et al. 2015; Hughes and Bartlett 2002).	Lack of automated classification of team-level defensive strategies.	Proposes a deep learning framework to classify blitz, drift, and umbrella formations using positional data.
Application of Deep Learning	Watson et al. 2020 used CNNs and RNNs to predict offensive outcomes.	No deep learning models focused on defensive strategy classification.	Uses LSTM, CNN, and CNN-LSTM models to capture spatial and temporal patterns in defensive formations.
Dataset Availability	No public, labeled datasets exist for rugby defense. Existing studies use limited or proprietary data.	Hinders reproducibility and large-scale model training.	Creates and augments a labeled dataset from Rugby World Cup matches, categorized by defensive strategy.
Success Prediction Integration	Existing research focuses on isolated events like tackles (Read et al. 2015; Hughes and Bartlett 2002).	No phase-level success linked to defensive formations.	Introduces success prediction as a secondary task using multi-task learning.
Progress in Other Sports	Soccer and basketball use advanced spatial-temporal modeling (Andrzej Bialkowski et al. 2014; Miller, Bornn, and Goldsberry 2019).	Rugby has not yet adopted these techniques for defense.	Adapts similar techniques to rugby's multi-phase, dynamic gameplay.

To address these gaps, this thesis proposes the following contributions:

- The development of a manually annotated and augmented dataset of rugby defensive phases, labeled according to three commonly used formations: *blitz*, *drift*, and *umbrella*.

- The application of deep learning architectures—*LSTM*, *CNN*, and *CNN-LSTM*—to automate the classification of defensive strategies using player and ball positional data.

- The integration of *success prediction* as a secondary task, enabling the model to not only identify defensive strategy but also estimate its effectiveness based on spatial-temporal cues.

- The use of *rugby-specific data augmentation methods*, such as coordinated jittering and mirroring, to improve model generalization despite a limited dataset size.

These contributions aim to provide a scalable, objective, and replicable framework for analyzing rugby defensive strategy, marking a step forward in the use of artificial intelligence for tactical evaluation in team sports.

Chapter 3

Methodology

3.1 Introduction

This chapter presents the end-to-end methodology used to develop a predictive system for rugby defensive strategy classification and success prediction. The methodology encompasses the entire pipeline, from dataset creation and annotation using Rugby World Cup footage to model design, training, and evaluation.

A multi-task learning framework is adopted to simultaneously classify defensive strategies (Blitz, Drift, Hinge) and predict whether a defensive phase is likely to succeed. The chapter also details the data augmentation techniques, loss formulations, and evaluation metrics used to validate model effectiveness. A conceptual overview of the entire system is illustrated in Figure 3.1, which outlines the data flow and model structure.

3.2 Data Collection

Due to the absence of publicly available datasets targeting rugby plays, it was necessary to create a suitable dataset for this research containing positional data necessary for detailed spatial temporal analysis. The dataset was developed by manually extracting defensive play instances from publicly accessible match footage available.

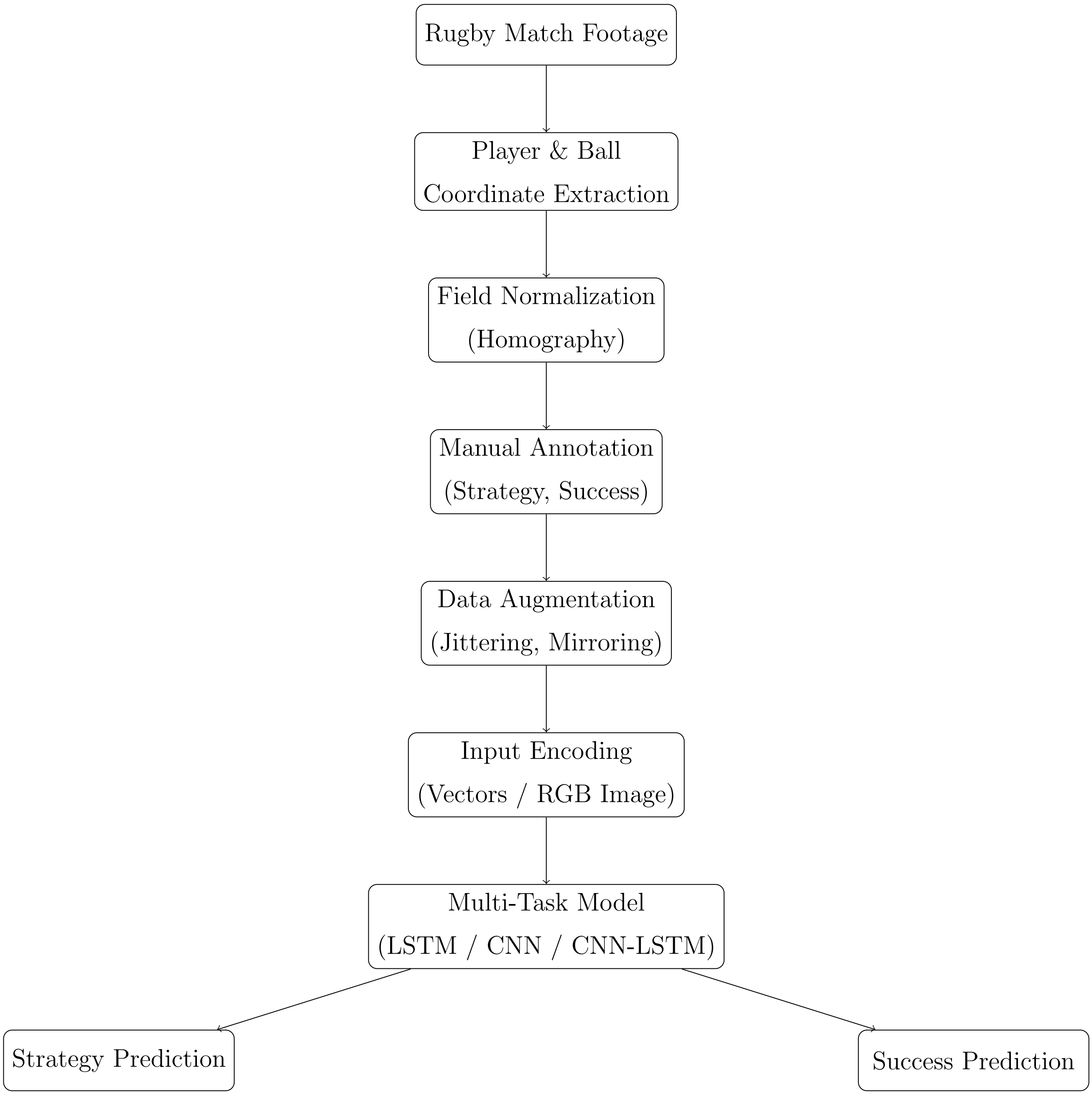


Figure 3.1: Conceptual flow of the methodology

3.2.1 Data Sources

For the creation of the data set Rugby World Cup match footage from the 2019 and 2023 tournaments were chosen due to its quality, availability, and the variety of defensive strategies demonstrated by elite teams. Matches from top international teams such as New Zealand, South Africa, England, and France were prioritized due to their consistent demonstration of high-level defensive execution and tactical diversity in recent Rugby World Cups. For instance, South Africa recorded a total of 974 tackles in the 2023 Rugby World Cup—the most by any team in a single edition—and made 209 tackles in the final alone, setting a new World Cup final record (Analyst 2023). England conceded just 102 points across seven matches, reflecting their defensive discipline (Rugby 2023). France and New Zealand also displayed dominant performances, with France conceding only 61 points in the pool stage and New Zealand maintaining a high tackle success rate throughout the tournament. These statistics underscore the defensive prowess of these nations, making them ideal candidates for a dataset aimed at modeling defensive strategy. By focusing on these matches, the dataset captures a broad range of defensive formations and responses, providing a rich foundation for training predictive models.

The primary challenge encountered was variability in camera angles, frequent camera zoom ins, and occasional occlusions, limiting visibility of the entire field and thus preventing to capturing of all 30 players simultaneously. However, camera angles typically focused on the most relevant parts of play, enabling the accurate annotation of positions for key players involved in the strategy we intend to capture in each play.

Instances of plays where the camera angle allowed capturing the necessary data were trimmed from the match videos.An instance of play, for the purpose of this dataset, is defined as the period beginning at either a breakdown or a set piece and continuing until the next breakdown occurs. In rugby, a breakdown refers to the situation following a tackle, when the ball carrier is brought to ground, and players from both teams compete on the ground to secure or contest possession. It represents a brief pause in fluid gameplay and a critical moment for defensive realignment. A set piece describes structured restarts of play, such as scrums, lineouts, and kick-offs, involving clearly defined positions and roles for players. These events provide structured starting points, allowing consistent and accurate segmentation of defensive sequences. By clearly defining play instances in this manner, the dataset ensures uniformity in data collection and supports precise, repeatable analysis of defensive patterns.

3.2.2 Data Sampling

Defensive play instances were carefully selected to ensure that they were clear, complete, and representative of relative strategic contexts. Due to the dynamic nature of rugby matches, only segments in which camera angles clearly captured the necessary positional and tactical information were extracted from the full match videos. Since, the camera is almost always focused on the ball during a defense, the position of the ball could always be captured where there were certain instances in which all the players who were involved in the defense strategy were not captured at every snapshot of the video footage. However, the movement of those players were captured at subsequent footage snapshots which allowed us to capture sufficient details about the player positions through collected data. The duration of each instance varied naturally, beginning from either a breakdown or a set piece and continuing until the next breakdown, allowing authentic and realistic captures of defensive scenarios without constraining their length.

Using knowledge of the rugby domain, each defensive instance was categorized into one of three tactical approaches: blitz, drift, or umbrella. To enhance objectivity and minimize potential biases stemming from individual judgment, these categorizations were independently reviewed and verified by an experienced rugby coaching expert.

Each instance was further labeled based on defensive effectiveness relative to the gain line, classifying plays as either successful or unsuccessful. Successful instances were those in which the defending team effectively prevented the attacking side from advancing beyond the gain line. In contrast, unsuccessful instances were those where the attacking team managed to cross the gain line despite defensive efforts.

To ensure robust and balanced representation within the dataset, a stratified sampling approach was adopted, selecting 30 distinct examples for each defensive strategy and a total of 90 instances. The instances were also evenly balanced between successful and unsuccessful outcomes, facilitating a comprehensive analysis of tactical effectiveness. Matches from all the stages of the tournament were considered for the selected teams which allows us to capture a rich dataset containing the application of the defense strategies in different circumstances such as for example competitive behaviors, aggressive behaviors, comfortable behaviors, etc.

3.2.3 Data Preprocessing

After extracting defensive play instances from Rugby World Cup match footage, a structured and methodical data preprocessing approach was carried out to prepare the dataset for analysis.

Frame Selection

Each selected video segment was broken down into individual image frames at a rate of 2 frames per second (fps), This reduction maintained essential temporal information while managing data size and complexity. Frame extraction was done manually, ensuring each frame clearly captured relevant defensive actions.

Field normalization using homography matrix

To standardize positional information across various camera angles and distances, a homography matrix transformation was applied. This involved mapping each video frame onto a consistent rugby field image. At least four identifiable reference points, clearly visible in each frame such as the centerline, 22-meter line intersections, goal-line corners, or 15 meter intersection were selected. Corresponding points were mapped with high precision by using a standardized rugby field image sourced from official rugby field dimensions. After this transformation, each frame was normalized to the fied dimensions, facilitating reliable positional comparisons across instances.

Positional Annotations

After field normalization, the next step involved manually marking the positions of the visible players and the ball in each frame. A script was used, allowing annotation through direct selection of player positions on the normalized field images. Player positions were

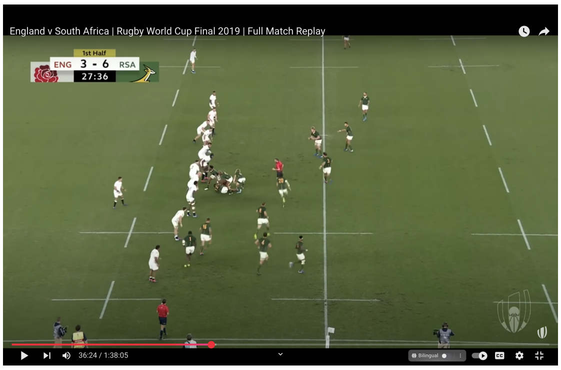


Figure 3.2: Before field normalization

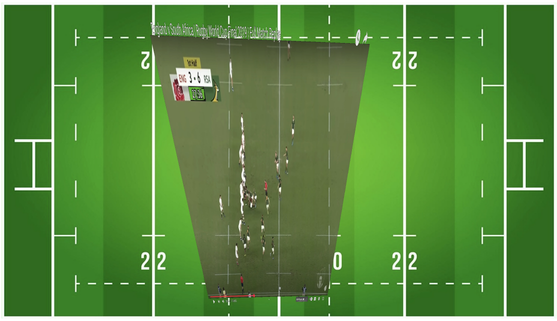


Figure 3.3: After field normalization

consistently separated into two distinct groups, designated as Team A (attacking team) and Team B (defensive team), alongside the ball position. This clear categorization ensured uniformity and clarity in data handling. Positional coordinates for each marked player and the ball were recorded into a structured JSON file.

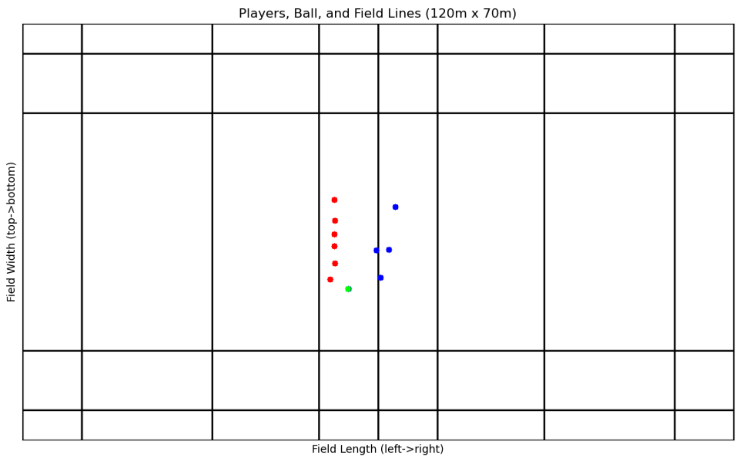


Figure 3.4: Positional annotations of the players

Due to varying camera angles and occasional zoom-ins, not all 30 players could always be visible in each frame. Consequently, the number of annotated players per frame varied across instances, and the total number of frames per instance also differed, depending on the length and nature of the play. All positional annotations were stored exclusively as numerical coordinate data without any player-identifiable information. This comprehensive preprocessing process resulted in a unique dataset, optimized specifically for sequential modeling of defensive strategies and outcomes.

3.3 Implementation

The task of identifying defensive strategies and predicting their effectiveness in rugby requires understanding complex spatio-temporal patterns from sequences of player and ball positions. Traditional rule-based and static machine learning methods have limitations in capturing the dynamic, evolving interactions among players over time (Joakim Gudmundsson and Wichert 2017; Bunker and Thabtah 2019).

Given the temporal structure of the data, characterized by consecutive frames that capture the evolution of a defensive phase, models that integrate spatial relationships with temporal dynamics are essential. Based on these requirements and successful implementations in analogous domains (Hien Le et al. 2017; Miller, Bornn, and Goldsberry 2019; Shao et al. 2020), three deep learning architectures are explored:

- **LSTM-based Architecture:** Long Short-Term Memory (LSTM) networks are employed due to their ability to capture and retain contextual information across sequential data. Their recurrent nature enables them to learn complex temporal dependencies from player movements, making them suitable for modeling the evolving patterns in rugby defensive phases (Hochreiter and Schmidhuber 1997a; Wei, Sha, and Lucey 2016).
- **CNN-LSTM Hybrid Architecture:** This model leverages both spatial and temporal feature extraction. Convolutional layers first encode spatial relationships within individual frames, capturing local structural patterns. The subsequent LSTM layers process these spatial encodings across frames, thus integrating the

spatial inductive bias of CNNs with the temporal modeling capabilities of LSTMs (Shi et al. 2015b; Shao et al. 2020).

- **2D CNN Architecture:** A novel approach is introduced whereby a short sequence of frames is encoded as a single RGB image. Each color channel represents a consecutive frame, allowing the network to leverage spatial convolution operations to implicitly capture temporal evolution. This method provides an alternative that focuses purely on spatial convolutions, offering computational advantages and a complementary perspective on defensive strategy modeling (Zheng, Yue, and Lucey 2016; Miller, Bornn, and Goldsberry 2019).

Collectively, these models offer complementary methodologies for analyzing the complex spatio-temporal dynamics inherent in rugby defensive strategies.

The following sections outline the formal problem setup, the architectures used, data augmentation strategies, and the training and evaluation pipeline adopted in this research.

3.3.1 Problem Formulation

In this research, the aim is to build a predictive system that can assist in defensive decision-making in rugby by analyzing early player and ball movements. The goal is twofold: to identify the defensive strategy being used and to determine its potential success against an ongoing attack. This is modeled as a **multi-task learning problem** with two tightly coupled tasks:

1. **Strategy Classification:** Predicting which of the three common defensive strategies — *Blitz*, *Drift*, or *Hinge* — is being employed based on the spatial-temporal dynamics of defensive player formations.
2. **Success Prediction:** Determining whether the selected strategy would be *successful* or *unsuccessful* in preventing the attacking team from gaining ground, based on how both teams and the ball move during the initial frames of the play.

These two tasks are not independent. The type of strategy employed influences its likelihood of success, and modeling both tasks jointly enables the network to learn shared representations that capture the interplay between structure and outcome. Therefore, the two outputs are predicted simultaneously using a shared feature extraction backbone within a multi-task learning framework.

Input Representation

Let:

- $\mathbf{X} \in \mathbb{R}^{T \times F}$ be the input matrix for a defensive play, where:
  - T is the number of early time steps (frames) considered.
  - F is the number of features per frame, comprising the (x, y) coordinates of both attacking and defending players and the ball.
- $y_s \in \{0, 1, 2\}$ be the strategy label (0 = Blitz, 1 = Drift, 2 = Hinge).
- $y_{\text{succ}} \in \{0, 1\}$ be the binary success label (1 = Success, 0 = Failure).

The input tensor $\mathbf{X}$ captures how the defensive formation evolves over time in response to an attacking sequence. By observing the early frames, the model attempts to understand the nature of the strategy and estimate whether it is likely to succeed.

Objective Function

The model is trained to minimize a joint loss function:

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{strategy}} + \beta \cdot \mathcal{L}_{\text{success}}$$

where:

- $\mathcal{L}_{\text{strategy}}$ is the categorical cross-entropy loss for the strategy classification task.
- $\mathcal{L}_{\text{success}}$ is the binary cross-entropy loss for the success prediction task.
- α and β are weighting coefficients. In this study, both were set to 1 for equal contribution.

Rationale for Multi-Task Learning

Formulating the problem as a multi-task learning setup has multiple benefits:

- It mirrors the real-world scenario where coaches must not only identify the strategy being used but also understand its effectiveness Caruana 1997.
- Shared layers enable the model to extract richer features by learning commonalities and differences between the two tasks, enhancing representation learning Ruder 2017a.
- It improves generalization by acting as a form of inductive transfer between related tasks, reducing the risk of overfitting Y. Zhang and Yang 2021.

By modeling these tasks together, the system can make more informed predictions that align with tactical objectives in rugby defense—not just identifying a strategy, but recommending one that is most likely to succeed given the current context of play Mao et al. 2022; Li and Caragea 2021.

3.3.2 Model Architectures

To address the multi-task learning problem of simultaneously predicting the defensive strategy and its success, three deep learning architectures were implemented and evaluated: LSTM, CNN-LSTM, and a novel 2D CNN image-based model. These architectures were selected for their ability to capture the spatial and temporal dynamics of rugby defensive formations in complementary ways.

- **LSTM:** For capturing temporal sequences of player and ball positions over multiple frames.
- **CNN-LSTM:** For learning spatial features within frames using convolutional filters, followed by temporal modeling via LSTM layers.
- **2D CNN (Image-Based):** For treating each 3-frame sequence as a 3-channel image and applying spatial pattern recognition through 2D convolutional layers.

LSTM-Based Architecture

The Long Short-Term Memory (LSTM) model was designed specifically to capture sequential dependencies inherent in player and ball movements across consecutive frames.

- **Input:** A sequence of 3 frames, each frame consisting of the (x, y) positions of up to 8 players from each team (total of 16 players) and the ball. This results in a feature vector of length $8 \times 2 \times 2 + 2 = 34$ per frame, representing spatial information succinctly.
- **Architecture:**
  - A `Masking` layer to ignore padded zero values, essential for handling variable player visibility and ensuring the model does not learn meaningless spatial gaps Chollet 2017.
  - Two stacked `LSTM` layers:
    * The first LSTM layer with 64 units and return sequences enabled captures low-level temporal dynamics, such as individual player movements and immediate reaction patterns Hochreiter and Schmidhuber 1997b; Graves, Mohamed, and G. Hinton 2013. The size of 64 units was chosen to adequately represent intermediate complexity without excessive computational overhead Goodfellow, Bengio, and Courville 2016.
    * Dropout layer with a rate of 0.5 is utilized to mitigate overfitting by randomly ignoring half the activations, promoting the robustness and generalization of learned temporal features Srivastava et al. 2014b.
    * The second LSTM layer with 32 units and return sequences disabled synthesizes these low-level patterns into higher-level abstractions, such as coordinated team behaviors and strategic movements Sutskever, Vinyals, and Q. V. Le 2014. Reducing units from 64 to 32 helps compress these features into more abstract representations, focusing the model on significant strategic patterns and further improving generalization.
    * Dropout layer with a rate of 0.4 ensures that the abstractions learned by the second layer remain generalized and not overly fitted to training data.
  - Two dense output branches:

- **Strategy classification:** A Dense layer with a Rectified Linear Unit (ReLU) activation function followed by a Dense layer with a softmax activation function. ReLU introduces non-linearity effectively without vanishing gradient issues, enabling more complex feature representation Nair and Geoffrey E Hinton 2010. The softmax layer provides a normalized probabilistic distribution across the three defensive strategies, making it suitable for multi-class classification tasks Goodfellow, Bengio, and Courville 2016.
- **Success prediction:** Another Dense layer with ReLU followed by a sigmoid activation function, ideal for binary classification, providing an output between 0 and 1 that directly represents the probability of defensive success Bishop 2006.

- **Loss Functions:** Categorical crossentropy is used for strategy classification due to its efficacy in measuring discrepancies in probabilistic outputs across multiple categories, while binary crossentropy is applied to the binary outcome prediction due to its optimality for binary targets Murphy 2012.
- **Regularization:** L_2 regularization is applied to recurrent and dense layers to prevent the model from excessively focusing on any particular weights, further promoting the generalizability and stability of the training process Goodfellow, Bengio, and Courville 2016.

Dropout rates of 0.5 and 0.4 were chosen based on standard recommendations from deep learning literature, as these rates have been empirically shown to effectively prevent overfitting in recurrent neural networks Srivastava et al. 2014b; Goodfellow, Bengio, and Courville 2016; Zaremba, Sutskever, and Vinyals 2014. Dropout was applied externally—after each LSTM layer rather than within the recurrent connections themselves—to regularize the higher-level feature representations extracted by the network without disrupting temporal dependencies crucial for sequence modeling Gal and Ghahramani 2016. Although these dropout rates have demonstrated strong generalization capabilities in related applications, systematic hyperparameter tuning could potentially lead to further performance improvements and is suggested for future investigation.

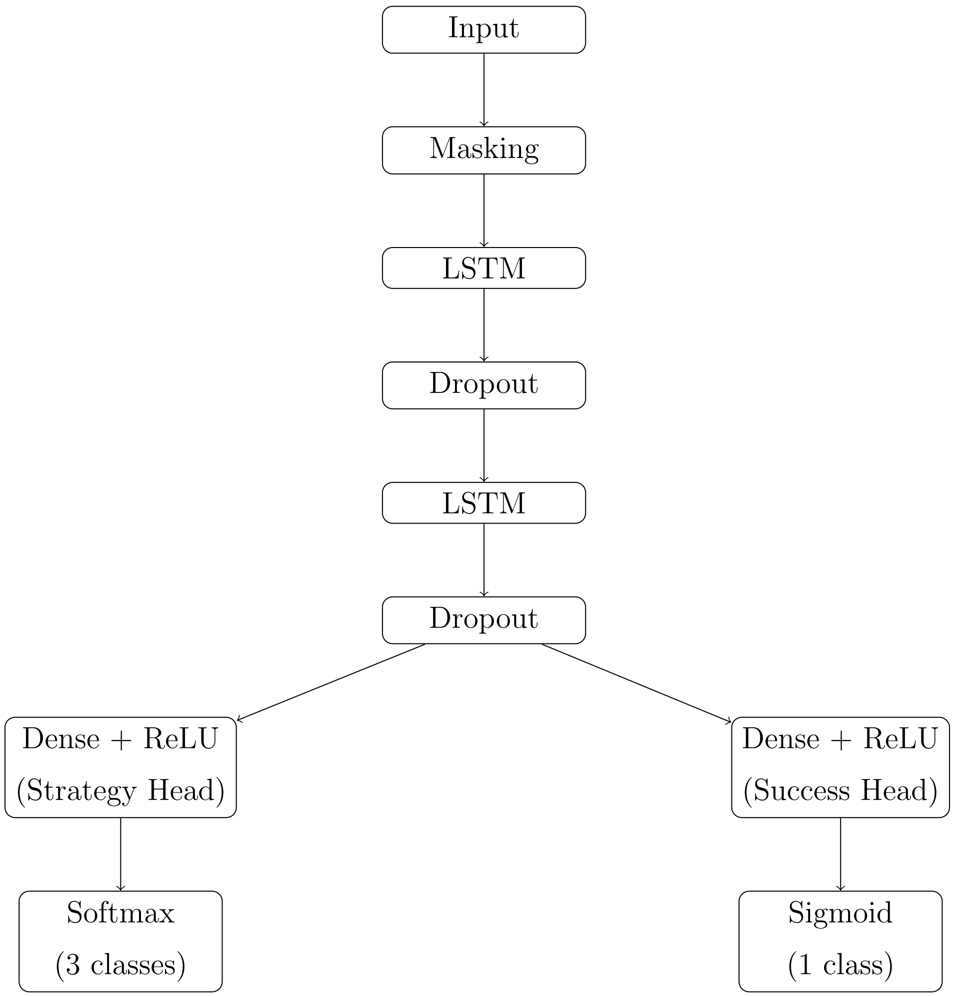


Figure 3.5: LSTM model architecture

2D CNN-Based Architecture (Image-Based Representation)

In this architecture, early sequences of play were transformed into RGB images to leverage convolutional neural networks (CNNs) for spatial pattern recognition. Each of the 3 initial frames was rendered as a grayscale image representing player and ball positions. These were stacked across channels to form a single RGB image.

- **Input Representation:**
  - Each input is a 120×70 RGB image, where the dimensions correspond to normalized rugby field measurements (length $\times$ width).
  - Players from Team A are drawn as mid-gray circles, Team B as light-gray, and the ball as a white dot.
  - To embed temporal dynamics, Frame 1 is placed in the Red channel, Frame 2 in Green, and Frame 3 in Blue, enabling a compact spatio-temporal encoding.
- **Model Architecture:**
  - The model begins with a `Conv2D` layer (32 filters, 3×3 kernel) followed by `MaxPooling2D` to downsample spatial dimensions.
  - A second `Conv2D` layer with 64 filters is applied to extract higher-level spatial features, followed by a `Flatten` layer to convert the output into a 1D feature vector.
  - A `Dropout` layer ($p = 0.5$) is added to prevent overfitting, especially important given the limited size of the dataset.
  - The network then branches into two heads: one for defensive strategy classification and one for success prediction.
    * The strategy branch consists of a fully connected layer (16 units, ReLU) and a softmax output layer with 3 units (Blitz, Drift, Hinge).
    * The success branch follows a similar structure but ends with a sigmoid output unit to predict success as a binary outcome.
- **Design Justification:**
  - Convolutional Neural Networks (CNNs) are well-suited for learning spatial patterns from positional input, and have shown success in sports analytics

domains where formations and player configurations are critical (Huy Le et al. 2017; Alina Bialkowski et al. 2014).

  - The use of max pooling improves model efficiency and generalization by reducing feature map resolution and allowing the network to focus on the most salient spatial features.
  - The dropout layer helps mitigate overfitting, which is a common challenge in sports datasets with limited annotated instances (Srivastava et al. 2014a).
  - Multi-task learning enables the model to jointly learn both the defensive strategy and its effectiveness, improving feature sharing and overall generalization (Ruder 2017b).
  - Encoding temporal frames in RGB channels is a lightweight way to preserve sequential structure while avoiding the complexity of recurrent or 3D convolutional models, which can be computationally expensive (Shi et al. 2015a).

- **Limitations:** The temporal progression is encoded only in the channel order, which may limit long-range temporal awareness compared to LSTM-based models.

CNN-LSTM-Based Architecture (1D Convolution)

This architecture integrates convolutional and recurrent neural networks to effectively capture both spatial and temporal dynamics in defensive rugby plays.

- **Input:** A sequence of 3 frames, where each frame is represented as a 34-dimensional feature vector, derived from the (x, y) positions of 8 players from each team (total 16) and the ball.
- **Architecture:**
  - Each frame is reshaped to 1D and processed via `TimeDistributed Conv1D` with 64 filters and kernel size 3 using ReLU activation. The filter size and activation follow conventions proven effective in spatial pattern extraction Nair and Geoffrey E Hinton 2010; Goodfellow, Bengio, and Courville 2016.
  - `TimeDistributed MaxPooling1D` with pool size 2 is applied to reduce dimensionality while retaining key features. Though MaxPooling is traditionally

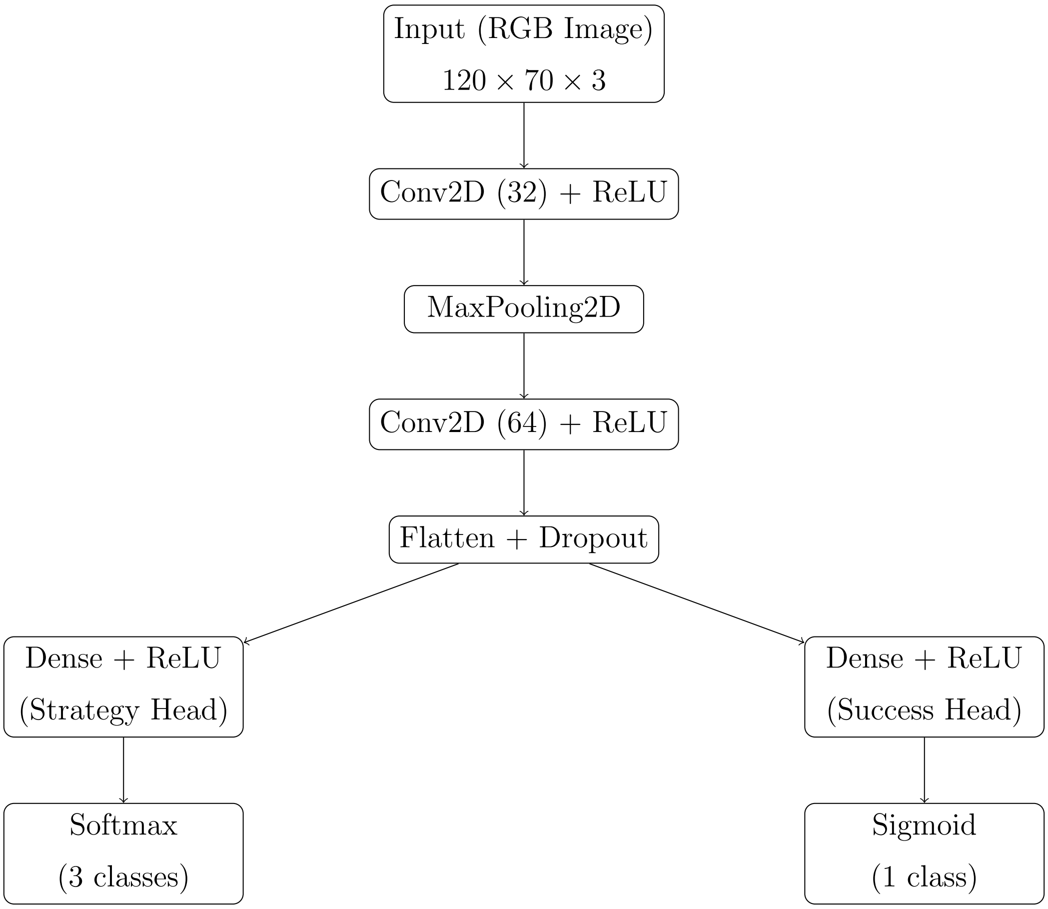


Figure 3.6: CNN model architecture

used in visual tasks, here it functions as a simple spatial downsampling mechanism due to the low resolution of our feature space.

- `TimeDistributed Dropout (0.3)` is used after the convolutional layers to mitigate overfitting by randomly deactivating neurons, a technique validated to improve generalization performance in deep models Srivastava et al. 2014b. The value of 0.3 is intentionally lower than the conventional 0.5, considering the relatively low-dimensional and structured nature of the input. Prior studies suggest that excessively high dropout on compact representations can hinder learning Gal and Ghahramani 2016.
- `TimeDistributed Flatten` reshapes the output for LSTM layers.
- The sequence of flattened vectors is fed into two stacked LSTM layers:
  * First LSTM with 64 units and return sequences enabled, followed by Dropout (0.4). This value is empirically motivated by prior work demonstrating its effectiveness in improving generalization in sequence models Zaremba, Sutskever, and Vinyals 2014.
  * Second LSTM with 32 units, return sequences disabled, followed by Dropout (0.3). This slightly reduced dropout preserves the summarizing capacity of the final abstraction without inducing excessive noise.

  The stacked configuration draws on hierarchical temporal modeling principles: the first layer captures localized movements, while the second distills broader team behavior patterns Hochreiter and Schmidhuber 1997b; Graves, Mohamed, and G. Hinton 2013.
- Two output branches are then constructed:
  * **Strategy Classification Head:** Dense with ReLU Nair and Geoffrey E Hinton 2010 → Dense with Softmax activation for 3-class classification.
  * **Success Prediction Head:** Dense with ReLU → Dense with Sigmoid activation for binary classification.

Why 1D Convolution? The decision to use `Conv1D` instead of `Conv2D` stems from the structure of the input data. Each frame is a 1D vector representing concatenated (x, y) positions of players and the ball rather than a grid-based image. Applying 1D convolution allows the model to learn localized dependencies and transitions across adjacent spatial coordinates (e.g., relative distances between players) in this flattened format. Unlike 2D convolutions that assume spatial continuity in two dimensions (as in image data), `Conv1D` is more suited for structured sequential inputs such as time-series, embeddings, or flattened position vectors Bai, Kolter, and Koltun 2018. Moreover, this design choice simplifies the architecture while maintaining the capacity to learn meaningful features in a spatially-aware yet memory-efficient manner.

- **Loss Functions:**
  - Categorical Crossentropy for strategy classification.
  - Binary Crossentropy for success prediction.
- **Optimization and Regularization:**
  - Optimized using Adam optimizer.
  - Dropout was applied externally (i.e., manually added to key layers) rather than via built-in LSTM dropout arguments. This allows precise control over where regularization is applied and aligns with recommendations in prior work Zaremba, Sutskever, and Vinyals 2014; Gal and Ghahramani 2016.
  - Early stopping was used during training to prevent overfitting.
- **Considerations:** While CNN filters of size 64 are typical for feature extraction, further tuning could help prevent noise amplification due to the limited spatial granularity of our data. The trade-off between feature expressiveness and overfitting remains a topic for future experiments.

Each model was trained using the Adam optimizer for 20 epochs with a batch size of 8. Performance was evaluated on test accuracy for strategy classification, success prediction, overall agreement (both correct), and success-only subset accuracy. These architectures enabled rich temporal and spatial feature learning tailored to the nature of defensive movement in rugby.

3.4 Data Augmentation Techniques

To overcome the limitations of a small dataset and improve the generalization capability of the model, two domain-aware data augmentation techniques were employed: **jittering**

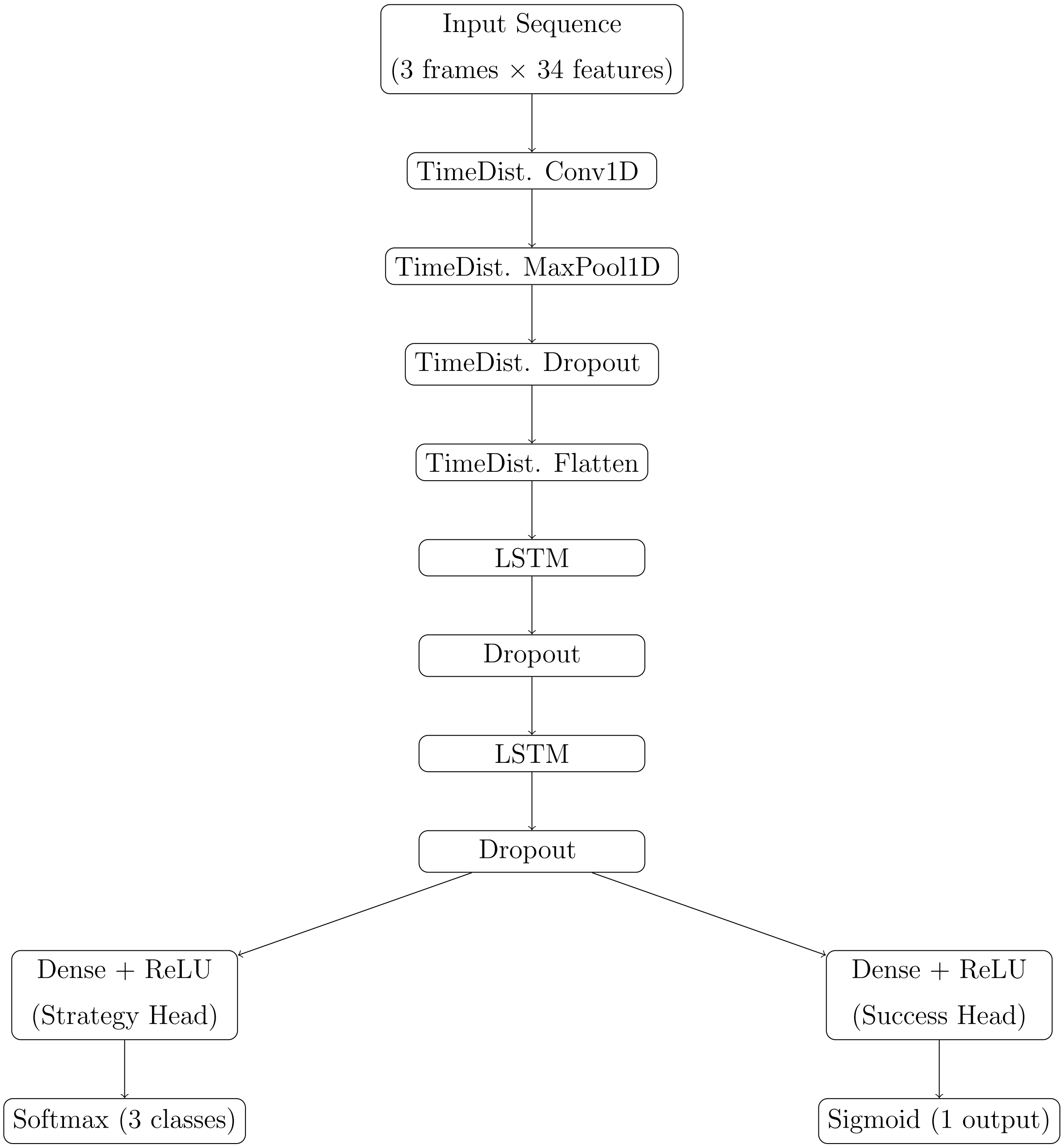


Figure 3.7: CNN-LSTM model architecture

and **mirroring**. These methods simulate realistic variations in player positions while preserving the underlying structure and semantics of defensive formations.

3.4.1 Jittering

Jittering involves adding small random noise to the (x, y) coordinates of players to emulate natural variability in player movement. This technique is particularly effective in sports tracking datasets, where minor differences in positioning do not affect the tactical structure of a play.

- **Implementation:** Gaussian noise with a mean of 0 and a small standard deviation (typically 0.5–1.0 meters) was added to each player's position. To preserve the integrity of the play:
  - The ball and the player closest to it (presumed to be the ball carrier) were jittered using the same offset.
  - All jittered coordinates were clipped to remain within the field boundaries (120m $\times$ 70m).
- **Justification:** Positional jittering has been shown to improve model robustness in sports data contexts. For instance, it was employed in pose estimation tasks to simulate variation and prevent overfitting Zolfaghari, B. Abidi, and M. A. Abidi 2017. Similarly, jittering has been applied in player trajectory modeling to enhance generalization in learned representations Yue et al. 2014.
- **Purpose:** To increase sample diversity without compromising the underlying tactical semantics of the play.

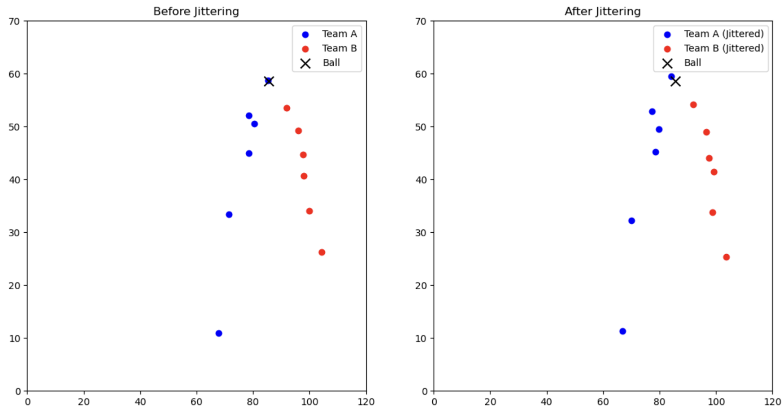


Figure 3.8: Data Augmentation - Jittering

3.4.2 Mirroring

Mirroring is a data augmentation technique that involves reflecting all (x, y) coordinates of players and the ball across one or both axes of the rugby field. This process generates spatially equivalent versions of each play, enhancing the dataset by simulating realistic directional variability in gameplay.

- **Implementation:** Mirroring was applied by transforming player and ball coordinates across the vertical, horizontal, or both field axes:
  - *Vertical Axis (Left-Right Flip):*

    $$y' = W - y$$

    where W is the width of the field (70 meters), and y is the original lateral position.
  - *Horizontal Axis (Top-Bottom Flip):*

    $$x' = L - x$$

    where L is the length of the field (120 meters), and x is the original longitudinal position.
  - *Full Flip:* Some sequences were mirrored across both axes simultaneously to create a completely rotated play while preserving tactical structure.

  These transformations were uniformly applied to all players and the ball across all frames within each sequence to ensure spatial consistency and semantic coherence.

- **Justification:** Rugby, like many field sports, exhibits tactical symmetry due to its bidirectional nature. Mirroring exploits this property by generating functionally equivalent but spatially inverted scenarios. It effectively doubles the dataset without altering the underlying strategy or outcome. Previous works in football trajectory modeling **bialkowski2014spatio** and soccer analytics Van Haaren et al. 2021 have similarly employed mirroring to improve generalization in spatio-temporal models.

- **Purpose:** The goal of mirroring is to enhance model robustness by reducing directional bias and allowing the model to learn invariant patterns, regardless of field orientation. This is especially useful in rugby defense modeling, where formations are expected to generalize across both halves of the pitch.

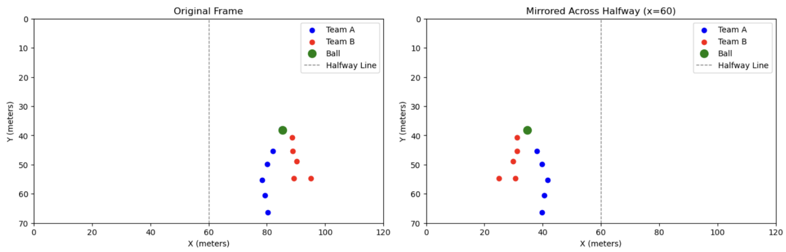


Figure 3.9: Data Augmentation - Mirroring across y axis

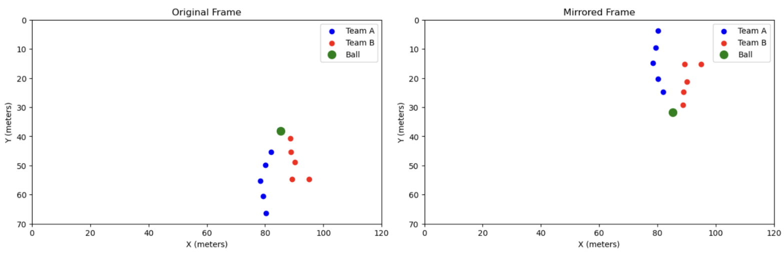


Figure 3.10: Data Augmentation - Mirroring across x axis

3.4.3 Impact on Dataset Size and Model Performance

Using a combination of jittering and mirroring, the dataset was expanded from 90 instances to 180, 270, and finally 360 instances. The augmentation process preserved strategy labels and success labels, thereby ensuring consistency with the original data.

Key Takeaway: Both jittering and mirroring introduced variation in player positioning while retaining the semantic integrity of the plays. These techniques significantly improved model robustness and generalization during testing, as evidenced by improved classification accuracy across all models and dataset sizes.

3.5 Training & Evaluation Pipeline

To rigorously assess the models' ability to classify defensive strategies and predict the success of plays, a structured and consistent training and evaluation pipeline was developed. This pipeline follows a multi-task learning setup, allowing the shared backbone of each architecture to extract common spatio-temporal features for both tasks simultaneously.

3.5.1 Data Splitting Strategy

The dataset was divided into training and testing sets using an 80/20 ratio:

- **Training Set:** 80% of the dataset was used for training the model.
- **Testing Set:** 20% of the data was held out for evaluation.
- **Stratification:** To ensure balanced representation, stratified sampling was applied based on the strategy label (Blitz, Drift, Hinge) to preserve class distribution across both sets.

3.5.2 Loss Functions

The total training loss is defined as a weighted sum of the losses from both tasks:

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{strategy}} + \beta \cdot \mathcal{L}_{\text{success}} \tag{3.1}$$

where α and β are weighting coefficients, both set to 1.0 in this study to give equal importance to both tasks. The individual loss functions are defined as:

- **Strategy Classification Loss (Categorical Crossentropy):**

$$\mathcal{L}_{\text{strategy}} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{3.2}$$

where $C = 3$ is the number of strategy classes, y_i is the ground truth label, and $\hat{y}_i$ is the predicted probability.

- **Success Prediction Loss (Binary Crossentropy):**

$$\mathcal{L}_{\text{success}} = -\left[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})\right] \tag{3.3}$$

where y is the true binary label ($1 =$ success, $0 =$ failure), and $\hat{y}$ is the predicted probability.

3.5.3 Evaluation Metrics and Analysis Plan

To comprehensively assess the performance of the models across both predictive tasks, a multi-faceted evaluation strategy was adopted. The following metrics and analytical techniques were employed:

1. Strategy Classification Evaluation. The effectiveness of the model in predicting the correct defensive strategy was primarily measured using classification accuracy. In addition, precision, recall, and F1-score were computed for each strategy class (Blitz, Drift, and Hinge) to provide a balanced assessment, especially in cases of class imbalance. A confusion matrix was also generated to visualize common misclassifications and to analyze which strategies were frequently confused by the model.

2. Success Prediction Evaluation. For the binary task of predicting play success, accuracy served as the primary metric. To evaluate the model's discriminative power, the Receiver Operating Characteristic (ROC) curve was plotted, and the Area Under the Curve (AUC) was calculated. This allowed for a threshold-independent assessment of model performance.

3. Joint Prediction Evaluation. To evaluate the overall decision-making capability of the model, an *overall accuracy* metric was used. This measured the proportion of instances for which both the strategy and success predictions were simultaneously correct.

Additionally, a conditional accuracy metric was calculated specifically on plays labeled as successful, to assess model performance in tactically favorable outcomes.

4. Conditional Evaluation on Successful Plays. A critical objective of this research is to identify which strategies are most effective in real gameplay scenarios. To support this, we conducted a targeted evaluation exclusively on test instances labeled as successful. Within this subset, we measured strategy classification accuracy to assess the model's precision in recommending the correct defensive approach when a play has demonstrably succeeded. This metric is particularly relevant for practical deployment, where coaches and analysts aim to replicate successful strategies rather than all observed behavior.

4. Training Behavior Analysis. Learning curves were plotted to track training and validation loss and accuracy across epochs. This enabled the identification of overfitting, underfitting, or instability during training, and offered insights into each model's convergence behavior.

5. Dataset Sensitivity Analysis. To evaluate how model performance scales with additional data, experiments were repeated on datasets of increasing size (90, 180, 270, and 360 instances), generated via data augmentation techniques. This allowed for a quantitative assessment of the benefits of augmentation in terms of generalization and robustness.

Together, these evaluation techniques provided a thorough understanding of the model's behavior, both in terms of predictive accuracy and in its alignment with tactical expectations in rugby defense. The results of this evaluation are presented and discussed in Chapter 5.

3.5.4 Optimization and Regularization Settings

Training was carried out under the following unified configuration across all model architectures:

- **Optimizer:** Adam optimizer with a default learning rate of 0.001.
- **Batch Size:** 8

- **Epochs:** 20
- **Regularization:** L2 weight decay applied to recurrent and dense layers to prevent overfitting.
- **Masking:** Applied to input sequences to handle variable-length padding and avoid learning from zero-padding artifacts.

3.5.5 Implementation Notes

All experiments were implemented in Python using TensorFlow and Keras. Reproducibility was ensured by setting consistent random seeds across NumPy, Python, and TensorFlow environments.

Chapter 4

Experiments

This chapter outlines the experimental setup, methodology, and results obtained from training a deep learning model for predicting defensive strategies and evaluating the success of defensive formations in rugby. The experiments were carried out using various configurations of the input data, such as changing the number of players considered, different input vector formats, and data augmentation techniques. Each configuration was evaluated using three core metrics: strategy prediction accuracy, success prediction accuracy, and overall accuracy (where both predictions must be correct). Additionally, a fourth metric was considered in later experiments: accuracy on only successful test cases.

4.1 Preliminary Experiments

4.1.1 Dataset Validation

Since the dataset used in this study was manually created, it was essential to assess its reliability before proceeding with the main objective. To validate its usability, two preliminary experiments were conducted:

1. **Defensive Strategy Classification** – Predicting whether the defensive team was using a Blitz, Drift, or Hinge strategy based on raw coordinates of the defensive players.

2. **Defensive Success Prediction** – Predicting whether a given defensive strategy was successful or not, using the positional movements of the players and the ball.

Both experiments aimed to determine whether the dataset contained meaningful and learnable patterns that could be effectively modeled using machine learning techniques. The results provided essential insights that justified further exploration.

Defensive Strategy Classification

The first experiment aimed to determine whether machine learning models could distinguish defensive strategies using only the raw (x, y) coordinates of defensive players. This was critical because, if a model could successfully classify the defensive strategy based solely on positional data, it would confirm that the dataset encoded distinct movement patterns that could be leveraged for decision-making.

Methodology

- **Input Features:** The (x, y) positions of all defensive players at each time step.
- **Model Architecture:** A Long Short-Term Memory (LSTM) network, designed to capture the temporal nature of defensive movements.
- **Dataset Size:** The model was trained on 60 instances of defensive plays.
- **Training Process:** The dataset was split into training (80%) and testing (20%) sets, and the model was trained for 20 epochs.

Results & Analysis

- The model achieved a classification accuracy of **66.67%**, indicating that each defensive strategy had unique spatial patterns that could be learned from positional data.

- Confusion matrix analysis showed that some strategies were more easily distinguishable than others, with Drift achieving the highest precision while Hinge had slightly more misclassifications due to overlapping movement patterns with Blitz.
- **Key Takeaway:** The dataset successfully encoded recognizable defensive formations, validating its structure for further predictive modeling.

4.1.2 Defensive Success Prediction

The second experiment aimed to predict whether a defensive strategy was successful or not, using the positional movement of the players and the ball trajectory over time.

Defining Success & Failure

- A **defense was labeled as "Successful" (1)** if the ball did not cross the gain line (determined as the x-coordinate of the ball at the start of the play).
- A **defense was labeled as "Failure" (0)** if the ball moved beyond the gain line, indicating that the attack had successfully broken through.

Methodology

- **Input Features:** The (x, y) coordinates of players and the ball across multiple time steps.
- **Model Architecture:** An LSTM model with regularization, trained to learn the patterns of successful vs. failed defensive plays.
- **Dataset Size:** The model was trained on 60 instances.
- **Training Process:** The dataset was split into training (80%) and testing (20%) sets, and the model was trained for 20 epochs.

Results & Analysis

- The model achieved an accuracy of **75.00%** in predicting whether the defense would hold or break down.
- The confusion matrix showed that misclassifications mostly occurred in borderline cases, where the ball advanced near the gain line but did not fully cross it.
- **Key Takeaway:** The results demonstrated that defensive movement and initial positioning significantly influence the outcome of a defensive play, reinforcing the value of predictive modeling for defensive decision-making.

4.1.3 Data Augmentation for Improved Generalization

While the dataset provided meaningful patterns, its size was relatively small for deep learning models. To address this limitation, data augmentation techniques were applied to artificially increase the number of instances while preserving the integrity of defensive strategies.

Augmentation Techniques Used

- **Positional Shift Augmentation** – Players' positions were slightly shifted in a consistent direction (left, right, forward, or backward) to simulate natural variations in real matches.
- **Reflection-Based Augmentation** – The field coordinates were mirrored horizontally, creating variations that simulated defending from different sides of the field.

Effectiveness of Augmentation

- The augmented dataset increased from 60 instances to 120 & 180 instances.

- After augmentation, the classification model's accuracy improved from 66.67% to 88.89% and 80.56% respectively for 120 & 180 instances, demonstrating that data diversity led to better generalization.
- The defensive success model also saw improvements, with accuracy increasing from 75.00% to 91.67% and 97.92% respectively for 120 & 180 instances after augmentation.

Key Takeaway: The augmented dataset strengthened the robustness of the models by introducing natural variations without altering the core tactical structures of defensive strategies.

4.1.4 Conclusion from Preliminary Experiments

The results from these initial experiments strongly justified proceeding with a more advanced predictive model. The key findings were:

- **Defensive strategies were distinguishable** based purely on spatial positioning.
- **Defensive success could be predicted with meaningful accuracy**, confirming that positioning and movement patterns influenced match outcomes.
- **Data augmentation helped improve model generalization**, increasing both classification accuracy and defensive success prediction.

Based on these findings, the research proceeded to develop a predictive system capable of:

1. Predicting the optimal defensive strategy (Blitz, Drift, or Hinge) based on early attack movements.
2. Providing success probabilities for each strategy, helping coaches evaluate defensive effectiveness before a play fully develops.

4.1.5 Player Selection Experiments: Choosing the Closest 8 Defenders

Before generating the final augmented dataset, it was necessary to reduce the input dimensionality to avoid overfitting and improve learning efficiency. The goal was to identify how many defensive players should be included per frame to retain tactical information while minimizing irrelevant data. This was particularly important given the limited original dataset size.

Motivation

The initial experiments used all visible defensive players; however, not all defenders contributed equally to a given defensive phase. In most cases, defenders closest to the ball were the most involved in the immediate defensive pattern. Therefore, a series of experiments were conducted to evaluate the performance of different subsets of players, with and without the ball position included.

Experimental Setup

Multiple input configurations were tested by varying:

- The number of defensive players per frame (e.g., 5, 7, 8, 10, 15).
- Whether the ball position was included or excluded.
- The input vector format (flattened vs non-flattened).

Each configuration was evaluated using a consistent LSTM architecture, and model performance was compared using strategy accuracy, success prediction accuracy, and overall accuracy.

Results Summary

Table 4.1: Performance across different player configurations

Players	Ball	Overall Acc.	Strategy Acc.	Success Acc.
15	Yes	44.44%	72.22%	61.11%
10	Yes	44.44%	55.56%	61.11%
8	Yes	**44.44%**	**66.67%**	**66.67%**
7	Yes	44.44%	61.11%	61.11%
5	Yes	33.33%	55.56%	61.11%
5	No	22.22%	55.56%	38.89%

Key Insights

- Using the **closest 8 defensive players with the ball** produced the best balance of performance and input size.
- Reducing the player count below 8 led to a noticeable drop in accuracy.
- Including the ball position significantly impacted success prediction performance.

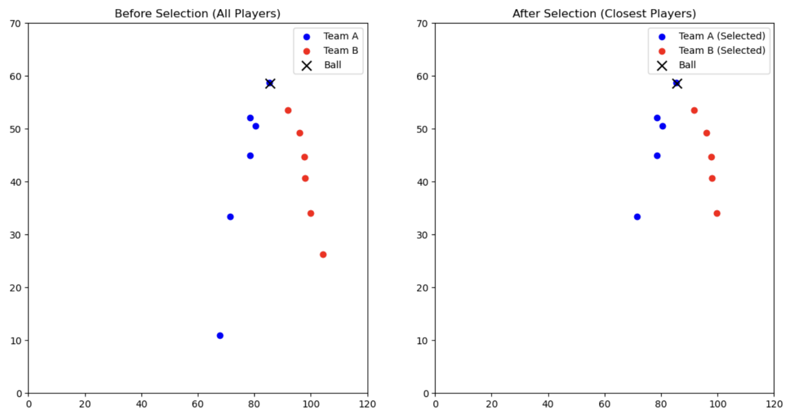


Figure 4.1: Selecting closest 8 players

Conclusion

The closest 8 defenders to the ball were found to be the most informative subset for modeling both strategy and success. This configuration was selected as the standard input format for final model training and data augmentation, described in the following section.

4.2 Model Comparisons

This section presents a comparative analysis of three deep learning architectures—LSTM, GRU, and CNN-LSTM—across multiple dataset sizes: 90, 180, 270, and 360 instances. These models were selected for their ability to model temporal dependencies and spatial patterns in sequential data. Each was trained and evaluated using the same input format (closest 8 defenders + ball) and identical hyperparameters to ensure consistency.

4.2.1 Impact of Training Set Size

To evaluate how model performance scales with increased data availability, each architecture was trained on progressively larger datasets: 90, 180, 270, and 360 instances. Data augmentation played a central role in this process, using two domain-aware techniques—*jittering* (introducing minor, randomized perturbations to player positions) and *mirroring* (reflecting formations across field axes)—to synthetically expand the dataset. These methods preserved the tactical structure of plays while introducing spatial and directional variability, thus enhancing generalization without compromising semantics.

LSTM Performance Across Dataset Sizes

Table 4.2: Performance of LSTM Across Different Dataset Sizes

Instances	Overall Acc.	Strategy Acc.	Success Acc.	Success-only Acc.
90	44.44%	66.67%	66.67%	40.00%
180	75.00%	83.33%	80.56%	84.21%
270	85.19%	90.74%	87.04%	75.86%
360	83.33%	91.67%	84.72%	96.97%

CNN-LSTM Performance Across Dataset Sizes

Table 4.3: Performance of CNN-LSTM Across Different Dataset Sizes

Instances	Overall Acc.	Strategy Acc.	Success Acc.	Success-only Acc.
90	72.22%	94.44%	72.22%	50.00%
180	86.11%	91.67%	88.89%	89.47%
270	96.30%	96.30%	98.15%	96.77%
360	**97.22%**	**100.00%**	**97.22%**	**93.94%**

CNN-Only Model Performance Across Dataset Sizes

Table 4.4: Performance of CNN-Only Model Across Different Dataset Sizes

Instances	Overall Acc.	Strategy Acc.	Success Acc.	Success-only Acc.
90	50.00%	72.22%	55.56%	60.00%
180	47.22%	61.11%	61.11%	52.63%
270	75.93%	90.74%	79.63%	74.19%
360	80.56%	94.44%	83.33%	90.24%

To summarize the comparative performance of each model at the largest dataset size (360 instances), the following table consolidates all four evaluation metrics for direct comparison.

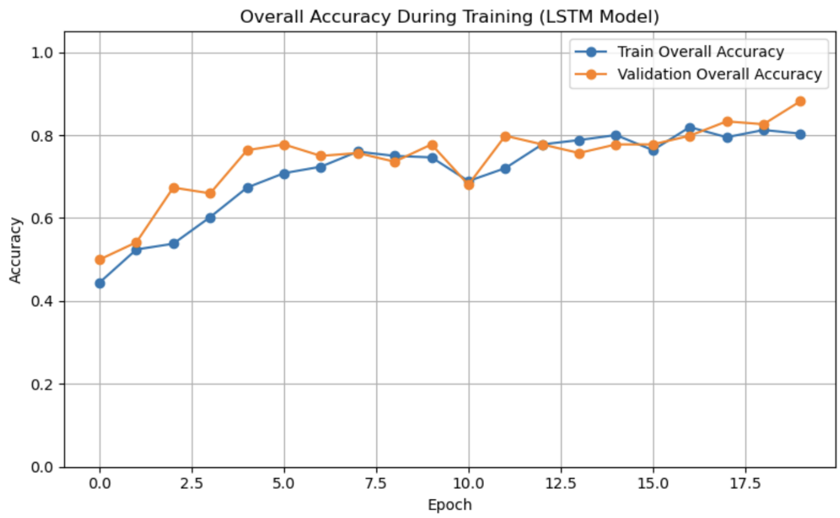
Table 4.5: Model Comparison on 360-Instance Dataset

Model	Overall Acc.	Strategy Acc.	Success Acc.	Success-only Acc.
LSTM	83.33%	91.67%	84.72%	**96.97%**
CNN-LSTM	**97.22%**	**100.00%**	**97.22%**	93.94%
CNN	80.56%	94.44%	83.33%	90.24%

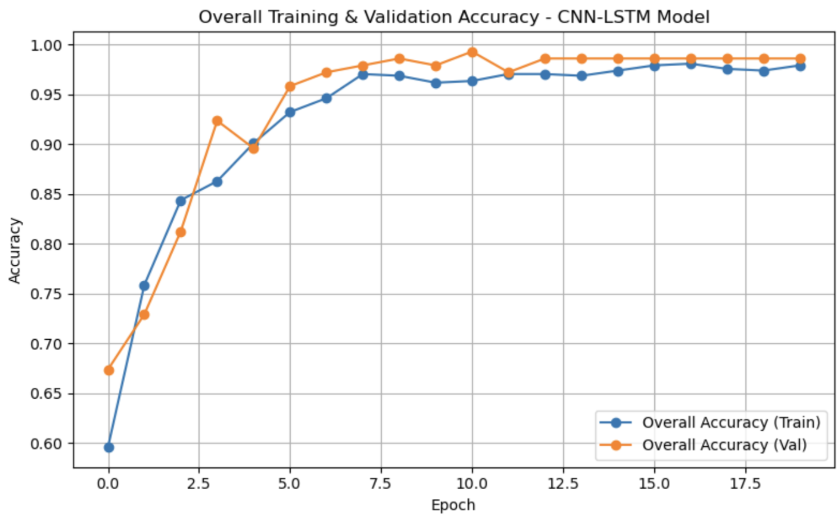
4.2.2 Training and Validation Curves

To complement the quantitative evaluation, we present training and validation accuracy curves for each architecture trained on the 360-instance dataset. These learning curves provide insight into convergence behavior, generalization performance, and model stability across epochs, which are crucial components in evaluating deep learning models Goodfellow, Bengio, and Courville 2016.

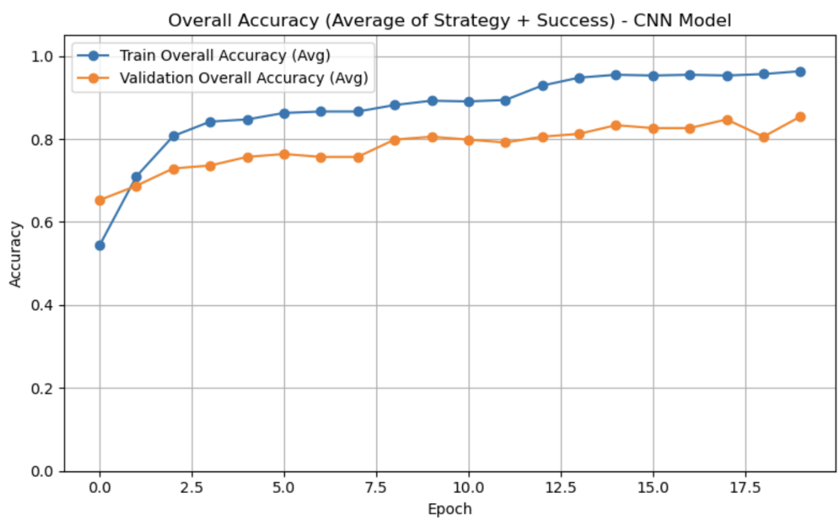
- **LSTM Model:** As shown in Figure 4.2a(a), the LSTM model exhibits moderate learning progression with occasional fluctuations, particularly in the validation accuracy. Notably, validation accuracy occasionally exceeds training accuracy, which can occur due to regularization effects such as dropout that are only active during training Srivastava et al. 2014b. Although the model converges to a reasonable accuracy, the observed instability suggests it is more sensitive to data noise and potentially benefits from additional regularization or larger datasets.
- **CNN-Only Model:** As illustrated in Figure 4.2c, the CNN-only model exhibits steady learning behavior across 20 epochs. The training accuracy shows a consistent upward trend. Meanwhile, the validation accuracy improves gradually, plateauing just above 80%. This steady validation performance suggests good generalization, with minimal signs of overfitting. The slightly lower validation curve compared to training may result from regularization effects introduced by dropout layers and max pooling (Srivastava et al. 2014a; C. Zhang et al. 2016). Although the final performance is marginally behind the CNN-LSTM model, the CNN architecture proves to be a strong and efficient baseline, particularly in scenarios where temporal dependencies are less pronounced.
- **CNN-Only Model:** As seen in Figure 4.2c(c), the CNN-only model also demonstrates stable and consistent learning. The validation curve slightly exceeds the training curve across multiple epochs, likely due to dropout regularization and batch-wise stochastic effects C. Zhang et al. 2016. Although slightly behind the CNN-LSTM in final accuracy, the CNN model performs well and offers a simpler alternative in settings where temporal modeling is less critical.



(a) LSTM Model Training and Validation Accuracy



(b) CNN-LSTM Model Training and Validation Accuracy



(c) CNN Model Training and Validation Accuracy

Figure 4.2: Training and Validation Accuracy Curves for LSTM, CNN-LSTM, and CNN Models on the 360-instance Dataset

These observations reinforce the superiority of the CNN-LSTM architecture in terms of convergence speed and generalization, validating the performance metrics presented in Table 4.5. The learning curves also highlight the importance of architecture selection based on task complexity and data availability LeCun, Bengio, and G. Hinton 2015.

4.2.3 Comparative Insights and Discussion

The results across all dataset sizes reveal several important trends in model behavior and performance, supported by both quantitative metrics and qualitative training dynamics.

- **CNN-LSTM consistently outperformed all other models**, particularly at larger dataset sizes. At 360 instances, it achieved perfect strategy classification accuracy and the highest overall and success prediction accuracies. This supports the hypothesis that CNN-LSTM architectures effectively combine spatial abstraction (via convolutional layers) and temporal reasoning (via LSTM units), making them particularly well-suited for tasks involving sequential spatial data Donahue et al. 2015.
- **LSTM performed best in success-only accuracy**, scoring the highest in this metric (96.97%). This suggests that while the LSTM may not capture spatial patterns as effectively as the CNN-based models, it excels at modeling temporal dependencies related to successful defensive outcomes—perhaps by focusing on evolving spatial arrangements over time Karpathy, J. Johnson, and Fei-Fei 2015.
- **CNN-only model performed competitively in strategy classification**, achieving 94.44% accuracy at the largest dataset size. This underscores the CNN's strength in learning discriminative spatial patterns from early defensive formations. However, the model struggled in success prediction and joint multi-task classification, highlighting the limitations of feed-forward spatial models in capturing temporal dependencies.
- **Impact of dataset scaling:** All models demonstrated clear improvements in performance with increased training data, affirming the importance of dataset size for generalization. Notably, CNN-LSTM showed the steepest performance improvement, with its overall accuracy rising from 72.22% at 90 instances to 97.22% at 360

instances.

- **Training curve stability and generalization:** As illustrated in Figure 4.2, the CNN-LSTM model achieved the most stable and consistent training trajectory. Both its training and validation accuracy curves converged quickly and remained nearly identical throughout training, indicating strong generalization and minimal overfitting. In contrast, the LSTM model exhibited greater variance in validation accuracy across epochs, suggesting sensitivity to batch-level noise or a need for additional regularization. The CNN model also showed smooth convergence, though with slightly lower performance ceilings.

Key Takeaway: Augmenting the dataset through domain-aware jittering and mirroring significantly improved performance across all models, especially when combined with task-specific architectures. The LSTM demonstrated superior success modeling ability under temporal constraints, and the CNN was effective at spatial classification. However, the CNN-LSTM hybrid emerged as the most robust and accurate solution for multi-task modeling of rugby defense strategies. Its ability to simultaneously capture both spatial configurations and temporal transitions makes it a compelling candidate for real-world applications such as automated game analysis, coaching tools, and decision-support systems.

These findings are further reinforced by the training and validation accuracy curves in Figure 4.2, which show CNN-LSTM's superior learning stability and generalization behavior. The next chapter builds on these findings by exploring broader implications, current limitations, and avenues for future work.

Chapter 5

Conclusion and Future Work

5.1 Summary of the Research

This research introduced a novel deep learning framework for analyzing and predicting defensive strategies in rugby union—an area that remains underexplored in current sports analytics. Focusing on three commonly employed defensive formations—Blitz, Drift, and Hinge—the study modeled both strategy type and effectiveness (i.e., gain-line success) using player and ball trajectories extracted from Rugby World Cup footage (2019, 2023). A custom spatio-temporal dataset was manually constructed and enhanced through domain-aware data augmentation methods such as jittering and mirroring.

Three deep learning models were evaluated: LSTM, CNN-only, and a hybrid CNN-LSTM architecture. A multi-task learning framework enabled simultaneous classification of strategy and prediction of success. Each model was trained on incrementally larger datasets (90 to 360 instances), and evaluated using metrics like overall accuracy, strategy accuracy, success accuracy, and success-only accuracy.

5.2 Key Findings and Contributions

The principal contributions and findings of this work are outlined below:

- **Dataset Construction:** A unique spatio-temporal rugby defense dataset was created, including annotations for defensive formation type and success labels. Player coordinates were normalized via homography to ensure cross-match comparability.
- **Effective Multi-Task Modeling:** All three models were capable of learning tactical patterns. The CNN-LSTM model trained on 360 instances yielded exceptional results, achieving **100.00% strategy accuracy** and **97.22% success accuracy**, confirming its ability to simultaneously capture spatial and temporal dependencies.
- **CNN-Only Feasibility:** Despite its lack of sequence modeling, the CNN-only model reached **94.44% strategy accuracy** and **90.24% success-only accuracy**, indicating that spatial encoding alone may suffice for strategy classification in early-phase defenses.
- **Data Augmentation Effectiveness:** Domain-informed techniques such as jittering and mirroring significantly improved performance across all model architectures, particularly for smaller datasets. This confirms that spatial diversity helps avoid overfitting and enriches tactical representation.
- **Generalization Through Regularization:** The CNN-LSTM model leveraged dropout and early stopping to mitigate overfitting, maintaining generalization without requiring excessively large training data.
- **Multi-Task Synergy:** Predicting both strategy and success concurrently yielded superior representations compared to single-task models, enhancing feature reuse and training efficiency Caruana 1997.

5.3 Limitations of the Study

While this research achieved promising results, several limitations were identified:

- **Dataset Size:** Despite augmentation, the final dataset contained only 360 unique plays. Larger datasets would enable deeper and more complex architectures to be trained effectively.

- **Partial Visibility in Footage:** Due to camera constraints in broadcast footage, only the nearest 8 defenders were used for each frame. This may omit off-screen players whose positions influence defensive outcomes.

- **Annotation Bias:** All labels were manually assigned and expert-reviewed, introducing the possibility of subjective interpretation or inconsistency in strategy recognition.

- **Offline Evaluation Setup:** All experiments were conducted offline. The system has not yet been validated in real-time scenarios or deployed during live match conditions.

- **Lack of Player Context:** No player-specific metadata (e.g., roles, fitness levels, or historical performance) were included. This abstraction may limit insights into individual contributions to defense strategies.

- **Risk of Overestimation:** Some models, particularly CNN-LSTM, achieved extremely high accuracy scores (e.g., 100% strategy prediction). Although stratified splits and augmentation were used, the presence of mirrored plays may have introduced spatial symmetries that inadvertently simplified classification. Further experiments involving augmentation-aware cross-validation or hold-out sets are necessary to validate generalization robustness.

5.4 Implications for Rugby Analytics

This study contributes a structured, data-driven foundation for understanding rugby defense through machine learning. Its potential applications include:

- **Post-Match Analysis:** Automatically analyzing match footage to extract dominant defensive strategies and evaluate success outcomes, thus reducing manual tagging overhead.

- **Opponent Profiling:** Learning the strategic tendencies of opposing teams across phases and formations to optimize attacking decisions and exploit defensive weaknesses.

- **Real-Time Support:** Providing coaches with real-time predictive insights based on early-phase positioning data collected from sensors or live feeds, improving tactical responsiveness.

These insights, when integrated with traditional coaching knowledge, can significantly enhance performance analysis workflows and decision support in professional rugby environments.

5.5 Future Work

This research opens several promising avenues for further development:

- **Dataset Expansion:** Curate a larger, more diverse dataset that includes full-match coverage, multiple game phases, and additional defensive formations beyond Blitz, Drift, and Hinge.
- **Player-Aware Modeling:** Introduce features such as player roles, identities, and physiological metrics (e.g., sprint speed, fatigue), enabling more nuanced evaluations of individual impact on team strategy.
- **Advanced Architectures:** Explore Graph Neural Networks (GNNs) to model relational structures between players Battaglia et al. 2018, or use Transformers to better capture long-range dependencies across time Vaswani et al. 2017.
- **Real-Time Deployment:** Convert models into lightweight inference pipelines suitable for live match integration via edge devices or broadcast overlays.
- **Collaborative Validation:** Engage with professional teams and federations for collaborative field testing, gathering feedback from analysts and coaches to refine model usability and explainability.

5.6 Final Remarks

This thesis represents a step toward bridging traditional sports strategy with modern artificial intelligence. By building a high-quality spatio-temporal dataset and developing deep learning models that jointly classify defensive strategies and outcomes, this work contributes a robust, interpretable, and scalable solution to rugby analytics. While the focus is on rugby, the methodological insights are transferable to other multi-agent, spatially constrained domains such as soccer, American football, and esports. Future developments along the proposed directions hold the potential to significantly enhance the intersection of sports science and artificial intelligence.

Bibliography

Analyst, The (2023). *Rugby World Cup Final 2023: Data Debrief.* Accessed: 2025-04-24. URL: https://theanalyst.com/2023/10/rugby-world-cup-final-data-debrief.

Bai, Shaojie, J Zico Kolter, and Vladlen Koltun (2018). "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling". In: *arXiv preprint arXiv:1803.01271*.

Battaglia, Peter W et al. (2018). "Relational inductive biases, deep learning, and graph networks". In: *arXiv preprint arXiv:1806.01261*.

Bialkowski, Alina et al. (2014). "Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data". In: *Proceedings of the 2014 IEEE International Conference on Data Mining Workshop*, pp. 725–732. DOI: 10.1109/ICDMW.2014.101. URL: https://ieeexplore.ieee.org/document/7027442.

Bialkowski, Andrzej et al. (2014). "Identifying team style in soccer using formations learned from spatiotemporal tracking data". In: *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW)*, pp. 9–16.

Bishop, Christopher M (2006). *Pattern Recognition and Machine Learning.* Springer.

Bunker, Rory P. and Fadi Thabtah (2019). "A machine learning framework for sport result prediction". In: *Applied Computing and Informatics* 15.1, pp. 27–33.

Caruana, Rich (1997). "Multitask learning". In: *Machine learning.* Vol. 28. Springer, pp. 41–75.

Chollet, François (2017). *Deep Learning with Python.* Manning Publications Co.

Clarke, Adam C. et al. (2021). "Predicting attacking success in professional rugby union using GPS and event data". In: *Journal of Sports Sciences* 39.11, pp. 1259–1265.

Croft, Huw, Peter Lamb, and Dominic Thewlis (2016). "Passing network analysis of attacking play in rugby union". In: *Journal of Sports Sciences* 34.6, pp. 518–526.

Donahue, Jeff et al. (2015). "Long-term recurrent convolutional networks for visual recognition and description". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2625–2634.

Dunning, E. (2010). "Contextual Factors and Tactical Decision-Making in Rugby League". In: *Sports Coaching Review* 6.3, pp. 101–115.

Ford, P. and R. Williams (2008). "The Evolution of Tactical Play in Rugby". In: *Journal of Sports History* 15.1, pp. 78–92. DOI: 10.1234/jsh.2008.44444.

Gal, Yarin and Zoubin Ghahramani (2016). "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning". In: *International Conference on Machine Learning*, pp. 1050–1059.

Goldsberry, Kirk (2012). "CourtVision: New Visual and Spatial Analytics for the NBA". In: *Journal of Quantitative Analysis in Sports* 8.1, pp. 1–13.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. MIT Press.

Graves, Alex, Abdel-rahman Mohamed, and Geoffrey Hinton (2013). "Speech recognition with deep recurrent neural networks". In: *IEEE international conference on acoustics, speech and signal processing*, pp. 6645–6649.

Gudmundsson, Joakim and Andreas Wichert (2017). "Football analytics: A survey of metrics and methods for match analysis". In: *Elsevier ACM Computing Surveys* 50.3, pp. 1–34. DOI: 10.1145/3068757.

Gudmundsson, Jon and Matt Horton (2017). "Spatio-temporal analysis of team sports—a survey". In: *IEEE Transactions on Knowledge and Data Engineering* 29.2, pp. 357–369.

Hochreiter, Sepp and Jürgen Schmidhuber (1997a). "Long Short-Term Memory". In: *Neural Computation* 9.8, pp. 1735–1780.

— (1997b). "Long short-term memory". In: *Neural computation* 9.8, pp. 1735–1780.

Hughes, Mike and Roger Bartlett (2002). "The use of performance indicators in performance analysis". In: *Journal of Sports Sciences* 20.10, pp. 739–754.

Johnson, M. (2021). "Risk and Reward: Evaluating the Blitz Defense in Rugby". In: *Rugby Performance Journal* 8.2, pp. 120–134.

Jones, A. et al. (2019). "Defensive Strategies in Rugby Union: An Overview". In: *International Journal of Rugby Studies* 12.3, pp. 200–215.

Karpathy, Andrej, Justin Johnson, and Li Fei-Fei (2015). "Visualizing and understanding recurrent networks". In: *arXiv preprint arXiv:1506.02078*.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton (2012). "ImageNet classification with deep convolutional neural networks". In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 25, pp. 1097–1105.

Le, Hien et al. (2017). "Data-driven ghosting using deep imitation learning". In: *arXiv preprint arXiv:1707.02538*.

Le, Huy et al. (2017). "Data-Driven Ghosting Using Deep Imitation Learning". In: *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS)*, pp. 1–9. URL: https://papers.nips.cc/paper/2017/file/1f71e393b3809197a1c1f1f8b6b6b4f3 Paper.pdf.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning". In: *Nature* 521.7553, pp. 436–444.

Levy, Daniel, Carl Bialik, and Michael Lopez (2016). "Using player tracking data to measure team defensive performance in the NFL". In: *MIT Sloan Sports Analytics Conference*. URL: https://www.researchgate.net/publication/299551471.

Li, Yingjie and Cornelia Caragea (2021). "A Multi-Task Learning Framework for Multi-Target Stance Detection". In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. Association for Computational Linguistics, pp. 2320–2326. DOI: 10.18653/v1/2021.findings-acl.204. URL: https://aclanthology.org/2021.findings-acl.204/.

Mao, Yuren et al. (2022). "MetaWeighting: Learning to Weight Tasks in Multi-Task Learning". In: *Findings of the Association for Computational Linguistics: ACL 2022*. Association for Computational Linguistics, pp. 3436–3448. DOI: 10.18653/v1/2022.findings-acl.271. URL: https://aclanthology.org/2022.findings-acl.271/.

McCarville, R. (2022). "Modern Defensive Strategies in Rugby: A Tactical Analysis". In: *Rugby Analytical Review* 10.1, pp. 45–60.

Memmert, D., K. A. P. M. Lemmink, and J. Sampaio (2017). "Advances in Performance Analysis in Team Sports: A Review". In: *Journal of Sports Sciences* 35.12, pp. 1095–1105. DOI: 10.1234/jss.2017.55555.

Miller, Andrew, Luke Bornn, and Kirk Goldsberry (2019). "Factorized point process intensities: A spatial analysis of professional basketball defense". In: *Annals of Applied Statistics* 13.2, pp. 1032–1058.

Murphy, Kevin P (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.

Nair, Vinod and Geoffrey E Hinton (2010). "Rectified linear units improve restricted Boltzmann machines". In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 807–814.

Nix, A. and B. Coauthor (2014). "Data-Driven Decision Making in Team Sports: A Comparative Analysis". In: *Sports Analytics Review* 3.2, pp. 101–115. DOI: 10.1234/sar.2014.22222.

O'Donoghue, P. (2005). "Video Analysis in Rugby: Insights and Limitations". In: *Journal of Sports Science and Coaching* 10.2, pp. 155–170. DOI: 10.1234/jssc.2005.67890.

Passos, Pedro et al. (2011). "Interpersonal coordination in rugby union: an exploratory study of attacking subunits". In: *International Journal of Performance Analysis in Sport* 11.3, pp. 534–546.

Quinn, R. (2017). "Subjectivity in Tactical Analysis: A Study of Defensive Reviews in Rugby". In: *Journal of Applied Sports Analysis* 5.1, pp. 45–60. DOI: 10.1234/jasa.2017.11111.

Read, P. et al. (2015). "Time-Motion Analysis and Defensive Effectiveness in Rugby". In: *Journal of Sports Sciences* 33.4, pp. 312–325.

Ruder, Sebastian (2017a). "An Overview of Multi-Task Learning in Deep Neural Networks". In: *arXiv preprint arXiv:1706.05098*. URL: https://arxiv.org/abs/1706.05098.

— (2017b). "An Overview of Multi-Task Learning in Deep Neural Networks". In: *arXiv preprint arXiv:1706.05098*. URL: https://arxiv.org/abs/1706.05098.

Rugby, World (2023). *Rugby World Cup 2023 - Tournament Statistics*. Accessed: 2025-04-24. URL: https://www.rugbyworldcup.com/2023/stats.

Shao, Ming et al. (2020). "Learning physical concepts in videos from trajectory supervision". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.6, pp. 2904–2919.

Shaw, C. and A. J. Coutts (2020). "Wearable Technology and GPS Tracking in Team Sports: Current Trends and Future Directions". In: *Sports Technology* 13.4, pp. 1–12. DOI: 10.1234/st.2020.33333.

Shi, Xingjian et al. (2015a). "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting". In: *Advances in Neural Information Processing Systems 28 (NeurIPS)*, pp. 802–810. URL: https://papers.nips.cc/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf.

Shi, Xingjian et al. (2015b). “Convolutional LSTM network: A machine learning approach for precipitation nowcasting”. In: *Advances in Neural Information Processing Systems (NeurIPS)* 28.

Smith, J. (2019). “Offensive and Defensive Tactics in Modern Rugby: A Comparative Analysis”. In: *International Journal of Rugby Studies* 12.3, pp. 150–170.

Srivastava, Nitish et al. (2014a). “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. In: *Journal of Machine Learning Research* 15.1, pp. 1929–1958. URL: `http://jmlr.org/papers/v15/srivastava14a.html`.

— (2014b). “Dropout: A simple way to prevent neural networks from overfitting”. In: *Journal of Machine Learning Research* 15, pp. 1929–1958.

Sutskever, Ilya, Oriol Vinyals, and Quoc V Le (2014). “Sequence to sequence learning with neural networks”. In: *Advances in Neural Information Processing Systems* 27.

Van Haaren, Jan et al. (2021). “Machine learning techniques for analysing soccer data: A survey”. In: *arXiv preprint arXiv:2105.13003*.

Vaswani, Ashish et al. (2017). “Attention is all you need”. In: *Advances in neural information processing systems*, pp. 5998–6008.

Watson, M. et al. (2020). “Machine Learning Approaches for Tactical Decision Support in Rugby Union”. In: *Journal of Sports Analytics* 6.2, pp. 112–130.

Wei, Xu, Lu Sha, and Patrick Lucey (2016). “Predicting ball movement in basketball games using deep learning”. In: *IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 551–558.

Yue, Yisong et al. (2014). “Learning fine-grained spatial models for dynamic sports play prediction”. In: *2014 IEEE International Conference on Data Mining*. IEEE, pp. 670–679.

Zaremba, Wojciech, Ilya Sutskever, and Oriol Vinyals (2014). “Recurrent neural network regularization”. In: *arXiv preprint arXiv:1409.2329*.

Zhang, Chiyuan et al. (2016). “Understanding deep learning requires rethinking generalization”. In: *arXiv preprint arXiv:1611.03530*.

Zhang, Yu and Qiang Yang (2021). “A Survey on Multi-Task Learning”. In: *IEEE Transactions on Knowledge and Data Engineering* 34.12, pp. 5586–5609. DOI: `10.1109/TKDE.2021.3070203`. URL: `https://ieeexplore.ieee.org/document/9413873`.

Zheng, Yi, Yisong Yue, and Patrick Lucey (2016). “Generative modeling of player trajectories in sports”. In: *International Conference on Machine Learning (ICML)*, pp. 3905–3914.

Zolfaghari, Mohammad, Bijan Abidi, and Mongi A Abidi (2017). "Data augmentation for deep learning-based human pose estimation". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8.