A Topic Model Approach to Represent and Classify American Football Plays

Jagannadan Varadarajan[1]
vjagan@adsc.com.sg
Indriyati Atmosukarto[1]
indria@adsc.com.sg
Shaunak Ahuja[1]
shaunak@adsc.com.sg
Bernard Ghanem[12]
bernard.ghanem@kaust.edu.sa
Narendra Ahuja[13]
n-ahuja@illinois.edu

[1] Advanced Digital Sciences Center of Illinois, Singapore
[2] King Abdullah University of Science and Technology, Saudi Arabia
[3] University of Illinois at Urbana-Champaign, IL USA

Automated understanding of group activities is an important area of research with applications in many domains such as surveillance, retail, health care, and sports. Despite active research in automated activity analysis, the sports domain is extremely under-served. One particularly difficult but significant sports application is the automated labeling of plays in American Football ('football'), a sport in which multiple players interact on a playing field during structured time segments called plays. Football plays vary according to the initial formation and trajectories of players - making the task of play classification quite challenging.

We define a play as a unique combination of offense players' trajectories as shown in Fig. 1(e). This dynamic group activity (play) is labeled by play type (run or pass) and direction (left, middle, or right). A *run* play is simply an attempt to advance the ball via running with the ball in left/middle/right directions, while a *pass* play is defined as an attempt to advance the ball by throwing it to the left/middle/right third of the field. In this paper, we address the combined problem of automatic representation and classification of football plays as summarized in Fig. 1. Such analysis will allow broadcasters to analyze large collections of videos, index and retrieve them more efficiently. Furthermore, this will also help coaches to infer patterns and tendencies of opponents more efficiently, resulting in better strategy planning in a game. For instance, Team A's coach can input Team B's videos and quickly understand commonly chosen play types and underlying strategies of Team B.

Problem definition. Given a play (short video clip) consisting trajectories of six important offense players: quarterback (QB), running back (RB), wide receiver left (WR-L), wide receiver right (WR-R), tight end left (TE-L) and tight end right (TE-R), we consider two problems: 1) *representation*, *i.e.*, derive canonical templates (common routes taken by players) for each play type, and 2) *classification*, *i.e.*, predict a label from one of the six play types for a given play, as shown in Fig. 1(e).

More formally, we are given D plays (or clips) denoted by $\{v_d\}, 1 \leq d \leq D$, where a play v_d is represented as a set of trajectories from R players: $\{X_r^d(t) : 0 \leq t \leq T, 1 \leq r \leq R\}$, T is the maximum possible duration of a trajectory, R is the maximum number of tracked players, and each play, indexed by d, is also associated with a unique class label $y_d \in \mathcal{C} = \{1,2,\cdots,C\}$ called a play type. Our objective is to learn a template τ_y corresponding to each class y, and a classifier function $F(v_d)$ that predicts a label $\hat{y}_d$ with high classification accuracy.

Approach. Our approach to modeling play types from trajectories builds on recent success of supervised topic models (STMs). Topic models in general, are used to capture dominant co-occurrent patterns, called "topics", in large collections of data. Intuitively, such patterns correspond to unique player interactions. For instance, run plays usually involve an intersection of QB and RB. STMs are effective in learning such salient co-occurring actions while achieving good classification performance. A particular example of STMs is the MedLDA model [2]. The MedLDA model learns salient topics with high classification accuracy by optimizing an objective function that combines the dual goals of maximizing inter-class margins and data log-likelihood simultaneously.

For our task of generating play type templates and classifying play types, we apply the MedLDA model on plays represented as a bag-of-words. In our case, a word in our vocabulary is defined by three different aspects of players' trajectories namely: motion direction, time and player role. More precisely, we use the following procedure to create words from plays: **(1)** motion directions coming from consequent trajectory observations are quantized into 18 equal sized directional bins of 20 degrees each; **(2)** temporal information is captured by grouping observations from every consecutive 10 frames into single time instance. Thus, each observation is also associated with a time-stamp; **(3)** a player's trajectory is associated with the player role, *e.g.*, QB, RB, etc. In our experiments, we use trajectories of six different players in creating our vocabulary. Therefore, a play with six different player roles occurring for a maximum duration of 250 frames can result in a vocabulary of $18 \times 25 \times 6 = 2700$ words.

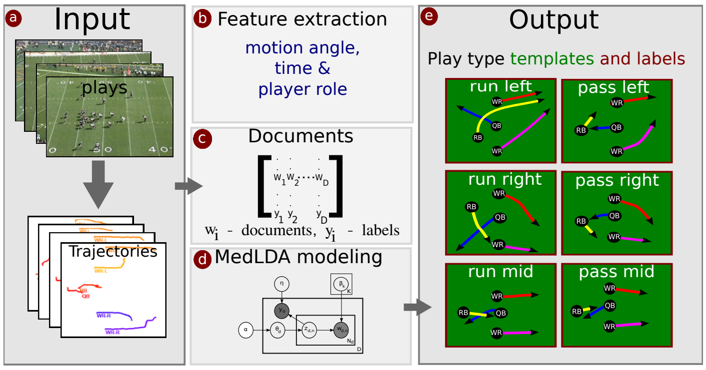

Figure 1: The goal of our work is to represent and classify play types in football videos. Our framework consists of a) player tracking, b) feature extraction, c) document creation and d) modeling using MedLDA [2]. Our outputs are: play type templates (e) and labels. Templates show are typical trajectories of four players (QB, RB, WR-L and WR-R).

Results. By applying MedLDA model on play clips, we obtain topics as well as a classifier that can be used to predict labels for test play clips. To qualitatively demonstrate that topics indeed capture salient features of play types, we associate each play type to a topic and generate trajectories from the respective topics. Quantitatively, we evaluate classification performance of our method by comparing against three other baselines on 6 class and 2 class (run vs pass) classification tasks for varying model sizes. We show that our approach based on MedLDA achieves 70% accuracy on 6-class and 88% on 2-classes, outperforming all three baselines for any number of topics. We also compare our results with the labels obtained from a sports expert who was asked to label plays based on only trajectory videos. By comparing the confusions of the expert with MedLDA, we show that MedLDA classifies marginally better (70%) than the expert while mostly agreeing on the classification performance for 4 out of 6 classes.

In addition to classifying plays, we extend our study to investigate which of the player roles is most discriminative in general and for a given play type in particular. From this analysis, we show that QB and RB give the highest classification accuracy as their actions largely determine if the play is a run or pass, and that WRs play an important role in pass plays than run plays. We validate our method using experiments on a large dataset comprising 271 play clips from real-world football games, which will be publicly available for future comparisons from [1].

[1] http://publish.illinois.edu/videoadsc/publications/football-trajectory-dataset/.

[2] Jun Zhu, Amr Ahmed, and Eric P. Xing. Medlda: Maximum margin supervised topic models. *JMLR*, 13:2237–2278, 2012.