CrossMark

ELSEVIER

Procedia Computer Science

An Exploratory Sentiment and Facial Expressions Analysis of Data from Photo-sharing on Social Media: The Case of Football Violence

Vasiliy Boychuk[1], Kirill Sukharev[1], Daniil Voloshin[1]
and Vladislav Karbovskii[1]

[1]*ITMO University, Saint-Petersburg, Russia*
`pphator@gmail.com, kvsuharev@gmail.com,`
`achoched@gmail.com, vladislav.k.work@gmail.com`

Abstract

In this article we propose the possibility to increase the level of security during football matches due to analysis of data that are placed on the social networks of these events visitors. We considered different ways to recognize emotions from photographs to trace the changes in the level of emotions in the photos depending on whether these pictures were taken during the game with fights in the stands or during normal games. We tested this assumption and our hypothesis has been partially confirmed. The software solution for emotion recognition from Microsoft Oxford showed that the level of emotional anger is noticeably higher in the photographs taken during the match with fights. In addition, other curious results were obtained, including an analysis of the key of the comments left by events visitors' photos.

Keywords: sentiment analysis, facial expression recognition, social media, football violence, data analysis

1 Introduction

According to research of Statistic Brain Research Institute, football is the most popular sport in the world. About 3.5 billion people follow this type of competition[1]. Each week football matches are held at various levels around the world. The average attendance at football games is relatively high compared to other sporting events – for example, around 40 thousand spectators came to every Bundesliga game in Germany in 2015 [2]. In the United States, during the same period, the audience at football matches exceeded 20 thousand visitors [3]. In terms of match attendance, Russia has one of the lowest numbers in Europe, but the number of spectators at the events of the Russian Major League exceeded 10 thousand people [2].

Football culture and violence have been inseparable since the introduction of professional football in England in the late 19th century [4]. Constant skirmishes between fans occur both outside the stadium

and in the stands. Violent clashes during a match only contribute to the increasing tensions between rival fans. In the entire history of football in the world there have been many notable incidents are further the most important ones described. On November 6, 1955 during the match "Napoli" - "Bologna" in Naples, there was unrest in the stands, injuring 152 people [5]. October 31, 1976 in Cameroon was marked by fights among the fans during a match of the local national team against Congo, which resulted in the deaths of two people [6].

Similar events continue to occur today. The biggest tragedy in recent years took place in Egypt. The match between the "Al-Masri" and "Al-Ahly" clubs, in which the hosts triumphed with a score of 3:1, ended in a massive attack on a group of spectators supporting the visiting team [7].

To avoid such cases, the authorities develop special regulations, imposing more control over the fans or increasing the number of police officers in the stands. However, the quantity of law enforcers present during the event has very little to do with the accurate prediction of potential confrontations between the fans or the infliction of material damage, since the intentions of football fans to start a clash are hidden from police.

This study has been carried out in an attempt to pursue the ultimate goal – assist the anticipation and timely prevention of future fights in stadium stands. Moreover, this is closely related to the market of mobile devices and social networks.

Today there are around 2.6 billion smart phones in use worldwide [8]. Analysts predict that by 2020 the number of devices will grow up to 6.1 billion [8]. Along with the increasing popularity and availability of smartphones, the social media market continues to evolve, too. As of August 2015, there were more than 2.2 billion active social network users in the world [9]. Today, social networks are not just a platform for interaction, communication, and cooperation. Services like Facebook, Twitter, Instagram also facilitate the collective participatory formation of an aggregate reflective view of the most important events of the present as well as an outlook to the past and future. Growth of the market of mobile devices and social networks is motivating the use of these funds for the solution of contemporary problems of society.

These facts lead us to assume that conclusions about the prevailing mood as a proxy for violent intentions during a soccer match can be drawn from the assessment of the "tone" of the facial expressions found on the photos taken by spectators and comments related to them. In this research, we gave preference to Instagram as the source of both textual and visual data. This social network has 400 million monthly active users [10], is developing dynamically (its growth doubled in 2015), is mainly accessed from smart phones, stores user geo-location data linked to entries and is widely distributed in Europe, North and South America [11]. Instagram is a convenient source of data – it provides researchers not just visual data (photos and short videos), but text (comments to posts). And last but not least, Instagram provides developers with a convenient API to retrieve data from servers.

2 Related Works

Sentiment analysis is a part of appraisal theory, which is focused on automatically identifying text topics or semantic which influence on real life action. Many studies focus on the semantic analyses to understand "positive", "negative" or "neutral" sentiment polarity and ones weights [12]. Sentiment analysis in online communication between people successfully predicts product consumption preferences and the trust among friends. The experiment results show that informal and short posts can predict preferences without any information history about preferences before [13].

Methods of emotion recognition are divided into three categories, based on the methods of face recognition: template-based methods (associated with holistic approaches to face recognition), feature-based methods (for the analytic-following approach to recognition of faces), and hybrid ones [17].

The Active Appearance Model (AAM), Labeled Graph Method (LGM), Point Distribution Model (PDM), Random Block Eigenvectors (RBE) belong to the first category, while Facial Characteristics