British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 43-48)
www.abjournals.org

NOLLYWOOD MOVIE SEQUENCES SUMMARIZATION DATASET FOR MACHINE LEARNING APPLICATIONS

Ofut Ogar Tumenayu[1] and Awuna Kile[2]

[1]Department of Software Engineering, University of Cross River State.

[2]Computer Science, University of Cross River State

Cite this article:

Ofut O. T., Awuna K. (2024), Nollywood Movie Sequences Summarization Dataset for Machine Learning Applications. British Journal of Computer, Networking and Information Technology 7(2), 43-48. DOI: 10.52589/BJCNIT-CVZ6HABH

Manuscript History

Received: 2 Oct 2024

Accepted: 26 Dec 2023

Published: 24 Jun 2024

ABSTRACT: *Summarization in recent times has become one of the most exploited areas of natural language processing tasks, due to significant increase in data volume, which is more than terabytes and petabytes. In this article, we present a custom dataset for Nollywood movie sequences, an extensive collection of 13 Nollywood movies representing various genres, including drama, comedy and action. This custom dataset was impendent into pre-train benchmark dataset TVSum following the same file format, frame size and others in order to suit our purpose. Fast Forward Moving Picture Expert Group (FFMPEG) was used for video encoding during pre-processing and final skimming. The proposed framework is effective for creating movie summaries with a high user satisfaction rate and having reasonably low computing requirements.*

KEYWORDS: Summarization, Dataset, Training, Data Pre-Processing, Nollywood custom dataset, models, movie sequences, Machine Learning.

Article DOI: 10.52589/BJCNIT-CVZ6HABH
DOI URL: https://doi.org/10.52589/BJCNIT-CVZ6HABH

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 43-48)

www.abjournals.org

INTRODUCTION

One of the most important components of any application using natural language is data. The more data you can use to train your model, the more advanced and precise the outcomes will be. Generally, sufficient benchmark datasets are needed for research on automatic movie summarizing. The utilization of existing datasets in this field is frequently severely constrained by challenges: for instance, Title-based Video Summarization (TVSum) is a benchmark dataset that is employed for the proposed system movie sequences validation summarization techniques (Yale et al., 2015). TV series and movies can be automatically generated if the application of video summarization is tailored for specific content-based scenarios (Sharghi et al., 2017).

There is no publicly available dataset suitable for Nollywood movie sequences summarization. We collected a TVSum50 dataset that contains 50 videos and their shot-level importance scores obtained via crowdsourcing. The video and annotation data permits an automatic evaluation of various video summarization techniques, without having to conduct (expensive) user study.

We selected 10 categories from the TRECVid Multimedia Event Detection (MED) task (Smeaton & Kraaij, 2006) and collected 13 Nollywood movies (5, 4, 4 each category) from YouTube and IROKO TV using the category name as a search query term. From the search results, we chose videos using the following criteria: (i) under the Creative Commons license; (ii) duration is 2 to 10 minutes; (iii) contains more than a single shot; (iv) its title is descriptive of the visual topic in the video. We collected videos representing various genres, including drama, comedy and action. This composition, which gave us a custom dataset, was impendent into the pre-train TVSum dataset following the same file format, frame size and others in order to suit our purpose. The proposed framework produced good results on the two benchmark datasets, which shows that the experimentation of the new model is properly executed (Ofut et al. 2022).

DATA PRE-PROCESSING

Real-world data tends to be dirty, incomplete, and inconsistent. Data preprocessing techniques can improve data quality, thereby helping to improve the accuracy and efficiency of the subsequent mining process (Iliou et al., 2015). A preliminary processing of data to prepare it for the primary processing or for further analysis is carried out at this stage by utilizing existing tools and methodologies. Firstly, a movie is automatically divided into a set of shots and scenes, where high-level visual and sound features are annotated manually for each shot. Some movie actors, events and semantic concepts are considered to be high-level visual features of a movie. The high-level visual features, especially the semantic concepts, are chosen by considering their frequency of occurrences within the domain of the movie and their ease of detection using a machine learning classifier. Visual features such as semantic events and actors are annotated manually.

Speech transcripts of each shot are selected using FFMPEG. Movie shots contain spoken content as audio along with their starting and ending timings in the corresponding movie. Since a movie is segmented into a set of shots based on the visual properties of the movie, audio sound in movie shorts have the same starting and ending timings with the shots. As such, speech transcripts of each shot are synchronized with the video frames.

Article DOI: 10.52589/BJCNIT-CVZ6HABH
DOI URL: https://doi.org/10.52589/BJCNIT-CVZ6HABH

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 43-48)
www.abjournals.org

Steps for Creating a Nollywood Custom Dataset

1). Collect data: Collect Nollywood movie files of different genres from YouTube, Netflix, and IROKO TV.

2) Prepare data: Take sample shots from each video file collected as shown in Figure 4.2 as well as Figure 4.3 showing the genre of each sample shot.

3) Annotate or label data: Annotate the movie data in different ways of annotating video data based on interest, highlighting objects of interest in the video frame and labeling them, as shown in Figure 4.4.

4) Export the annotated data into a file format (e.g., JSON, TSV, CSV or txt etc).

5) Create an ML model.

6) Use the prepared dataset to train the model.

7) Validate and test model.

Mathematical Model for the Dataset Engineering

An enhanced mathematical model for creating a custom dataset automatic movie sequences summarization was formulated as:

Let D_1 denote dataset 1. Note that D_1 is a function of u_i ($i = 1,\ldots,50$) foreign video files. Then,

$$D_1 = D_1(u_i) \qquad (1)$$

Let D_2 denote dataset 2. Note that D_2 is a function of v_i ($i = 1,\ldots,25$) foreign video files. Then,

$$D_2 = D_2(v_i) \qquad (2)$$

Let D_3 denote dataset 3. Note that D_3 is a function of l_i ($i = 1,\ldots,13$) local video files. Then,

$$D_3 = D_3(l_i) \qquad (3)$$

Let Z_1 be a function which depends on u_i ($i = 1,\ldots,50$) foreign video files through D_1. Then,

$$Z_1 = Z_1\{D_1(u_i)\} \qquad (4)$$

Let Z_2 be a function which depends on v_i ($i = 1,\ldots,25$) foreign video files through D_1. Then,

$$Z_2 = Z_2\{D_2(v_i)\} \qquad (5)$$

Article DOI: 10.52589/BJCNIT-CVZ6HABH
DOI URL: https://doi.org/10.52589/BJCNIT-CVZ6HABH

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 43-48)
www.abjournals.org

Let Z_* be a function which depends on D_1 and D_3 respectively through u_i ($i = 1, \ldots, 50$) foreign videos files and l_i ($i = 1, \ldots, 13$) local video files. Then,

$$Z_* = Z_* \{D_1(u_i) D_3(l_i)\} \qquad (6)$$

Let Z^* be a function which depends on D_2 and D_3 respectively through v_i ($i = 1, \ldots, 25$) foreign video files and l_i ($i = 1, \ldots, 13$) local video files. Then,

$$Z^* = Z^* \{D_2(v_i) D_3(l_i)\} \qquad (7)$$

Let Z be a two-path function which can assume the path of Z_* or Z^*. Then,

$$Z = \begin{cases} Z_* = Z_* \{D_2(u_i) D_3(l_i)\} \text{ as defined in (vi)} \\ Z^* = Z^* \{D_2(v_i) D_3(l_i)\} \text{ as defined in (vii)} \end{cases} \qquad (8)$$

Now, if Z assumes the path of Z_*, then the required model in this instance will be:

$$\frac{dZ}{du_{*i}} = \frac{dZ_*}{du_{*i}} = D_1(u_i) \frac{dD_3(l_i)}{du_{*i}} + D_3(l_i) \frac{dD_1(u_i)}{du_{*i}} \qquad (9)$$

Now, if Z assumes the path of Z^*, then the required model in this instance will be:

$$\frac{dZ}{dv_{*i}} = \frac{dZ^*}{dv_{*i}} = D_2(v_i) \frac{dD_3(l_i)}{dv_{*i}} + D_3(l_i) \frac{dD_2(v_i)}{dv_{*i}} \qquad (10)$$

where D_1, D_2, and D_3 denote TVsum, SumMe, and Nollywood custom dataset, D_1 is a function of u_i ($i = 1, \ldots, 50$) foreign videos files, D_2 is a function of v_i ($i = 1, \ldots, 25$) foreign videos files, D_3 is a function of l_i ($i = 1, \ldots, 13$) local video files. Z is a two-path function which can assume the path of Z* or Z_*.

Experiments and Results

Two tested and trusted benchmark datasets are used for training the model. TVSum dataset, as used in this research work (see Table 5.1), contains 50 videos of genres such as news, documentaries, vlogs that collected from YouTube with videos having lengths from 2 to 10 minutes, and annotated by 20 people for each video with shot-level importance scores. A collection of 13 Nollywood movies (5 Drama, 4 Action and 4 Comedy categories) from YouTube, Netflix and IROKO TV for our custom dataset creation. SumMe dataset, as shown in Table 5.1, contains 25 videos having lengths from 1 to 6 minutes, where each of the videos are annotated by 15-18 persons.

Article DOI: 10.52589/BJCNIT-CVZ6HABH
DOI URL: https://doi.org/10.52589/BJCNIT-CVZ6HABH

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 43-48)
www.abjournals.org

Table 1: Details of Dataset

Dataset	No. of videos	Length(Min)	No. of Annotators
TVsum	50 (+13 Nollywood video)	2-10	20
SumMe	25 (+13 Nollywood video)	1-6	15-18

Nollywood Custom Dataset

Figure 1: Sample Data for Nollywood Custom Dataset

category	video_id	title	url	length
AC	video_1	#In My Country	https://www.9ijarocks.com	5:54
AC	video_2	#In My Country	https://www.9ijarocks.com	5:03
AC	video_3	#10 Days in suncity	https://www.NetNaija	5:01
AC	video_4	#10 Days in suncity	https://www.NetNaija	5:01
CM	video_5	#30 days in atlanta	https://www.youtube.com/watch?v=XzYM3PfTM4w	4:58
CM	video_6	#Meet The Inlaws"	https://www.9ijarocks.com	4:54
CM	video_7	#10 Days in suncity	https://www.NetNaija	4:55
CM	video_8	#10 Days in suncity	https://www.NetNaija	5:01
DR	video_9	#Lying game	https://www.9ijarocks.com	5:00
DR	video_10	#Lying game	https://www.9ijarocks.com	5:08
DR	video_11	#Meet The Inlaws"	https://www.9ijarocks.com	5:01
DR	video_12	#In My Country	https://www.9ijarocks.com	4:59
DR	video_13	#The Hustle	https://www.9ijarocks.com	4:59

Figure 2: Nollywood Custom Dataset Description (Ofut et al., 2022)

Article DOI: 10.52589/BJCNIT-CVZ6HABH
DOI URL: https://doi.org/10.52589/BJCNIT-CVZ6HABH

Figure 3: Nollywood Custom Dataset Annotation

REFERENCES

Iliou, T.; Christos-Nikolaos, A.; Nerantzaki, M.; Anastassopoulos, G. A Novel (2015). Machine Learning Data Preprocessing Method for Enhancing Classification Algorithms Performance. In Proceedings of the 16th International Conference on Engineering Applications of Neural Networks, Rhodes Island, Greece, 25–28 September 2015; Association for Computing Machinery: NewYork, NY, USA, 2015; pp. 1–5.

Ofut O. Tumenayu, F. S. Bakpo, D. U. Ebem (2022). Nollywood Movie Sequences Summarization Using a Recurrent Neural Model. Quest Journal of Software Engineering and Simulation Volume 8 ~ Issue 1 (2022) pp: 06-1ISSN(Online) :2321-3795 ISSN (Print):2321-3809.

Ofut Ogar Tumenayu, Francis Sunday Bakpo, Hyacinth Agozie Eneh and Umoh Enoima Essien (2022). “Unified Framework for Summarization of Nollywood Movie Sequences and Audio Sound Data”. International Journal of Computer Applications (0975 – 8887) Volume 184– No.38, December 2022.

Sharghi, A., Laurel, J.S., Gong, B., 2017. Query-focused video summarization: Dataset, evaluation, and a memory network-based approach, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4788–4797.

Smeaton, A., P. Over, & W. Kraaij (2006). Evaluation campaigns and TRECVid. In MIR, 2006.

Yale Song, Jordi Vallmitjana, Amanda Stent, & Alejandro Jaimes (2015). *TVSum: summarizing web videos using titles*, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 5179–5187.

Article DOI: 10.52589/BJCNIT-CVZ6HABH
DOI URL: https://doi.org/10.52589/BJCNIT-CVZ6HABH