MOOCS RECOMMENDER BASED ON LEARNING STYLES

Analyze and classify complex types of MOOC video styles

Research Thesis
Project ID: 19-089

Author: P.H.P.S.L. Pathirana
IT 16 0043 82

Bachelor of Science Special (Honors) in Information Technology
Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

August 2019

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. Also, I hereby grant to Sri Lanka Institute of Information Technology the non-exclusive right to reproduce and distribute my dissertation in whole or part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as article or books).

Name	ID	Signature
P.H.P.S.L. Pathirana	IT16004382	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Supervisor: Mr. Nuwan Kodagoda

Co-supervisor: Ms. Kushnara Suriyawansa

Signature Date:

Signature of the supervisor: Date:

Signature of the co-supervisor: Date:

ABSTRACT

Massive Open Online Courses (MOOCs) has become one of the major sources of learning nowadays and number of MOOCS available increasing rapidly over the years. MOOCS include multiple modalities such as lecture videos, audio transcriptions, slides, textbooks, forum discussions and clickstream log data. Among them, lecture videos are the most fundamental component of MOOCs. Generally, the types of MOOC videos include: talking head, slides, coding, animations, writing (khan academy) etc. Some videos may contain multiple video production styles. Identification of each video production style in a video and calculation of the composition of each video production style using Image-based classification approach will be discussed in this paper. For this pre-trained VGG16 model has been used and Transfer learning mechanism has been used to train the model.

Keywords: MOOCs, Video Classification, Video Styles, Transfer learning

ACKNOWLEDGEMENT

This academic Thesis and work described in it were conducted as a part of 4th year research module (Comprehensive Design Analysis Project) under the degree program Bsc (Hons) in Software Engineering in Sri Lanka Institute of Information Technology. The fruition of this research and the associated software project are results of dedication of the 4 group members and the immense support given by a supervising panel.

Thus, we would like to express our gratitude towards Dr. Nuwan Kodagoda and Ms. Kushanara Suriyawansa, who are our supervisors as well as lecturers of Sri Lanka Institute of Information Technology. Completion of this academic work would not have been possible without their guidance and insight. We are extremely grateful to Dr. Malitha Wijesundara and Dr. Dasuni Nawinna, who reviewed and approved our research topic and provided valuable suggestions. Furthermore, we wish to express our gratitude to Mr. Jayantha Amaraarachchi, the Head of SLIIT Centre for Research who provided us with immense knowledge and guidance throughout.

Also, we would like to express our gratitude to all our colleagues and companions who helped us in testing and in various other ways to make our research come to fruition. At the same time, we would like to thank everyone who might not have been mentioned but has given their support and encouragement to us.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

Table 1:List of Abbreviation

MOOC	Massive Open Online Course
CNN	Convolutional Neural Network
VGG	Visual Geometry Group
RAM	Random Access Memory
CPU	Central Processing Unit

1. INTRODUCTION

1.1. Background Literature

MOOC resources include multiple modalities such as lecture videos, audio transcriptions, slides, textbooks, forum discussions and clickstream log data. Among them, lecture video is arguably the central and omnipresent component for knowledge transfer, to which other data modalities support [1]. There are countless MOOC courses in the Internet with countless videos associated with them. People have emphasis on retrieving those videos with specific category and it is infeasible for user to go through tremendous number of videos and find the video of interest. Generally, the types of content in MOOC videos include: talking head (instructor explaining the content), slide (PowerPoint presentation with voice-over), coding (screen casting of writing codes), animations, writing (khan academy) etc. In production of some MOOC videos, there are lot of transitions between visual views, e.g.: switching from a talking head to a slide or switching from a slide to an animation. The ability to automatically classify individual frame of a video into a production style and determining the composition of each production style can help to collect data more quickly and perform analysis of MOOCs in more granular level.

1.2. Research Gap

Even though MOOCREC V1 has introduced a set of key features that were not present in any other tool a year ago, through literature review we identified a set of key areas that can be improved or extended along with a new feature that is not present in MOOCREC V1. Also, during the phase of literature review, there were no other tools that matched the feature set offered by MOOCREC V1; which led us to compare MOOCREC V1 and our solution, which is referred to as MOOCREC V2. In areas such as processing, using resources optimally is an area that MOOCREC V1 has not covered. Also, it was clear that newer video styles can be introduced to broaden the number of recommendations that can be provided across a range of learning dimensions. There are only few numbers of studies that have been carried out about video styles used in MOOCs.

1.3. Research Problem

It is an important to analyze and classify the MOOC videos when recommending MOOCs to different learning characteristics and requirements. Furthermore, MOOCs belong to a wide range of video styles in which they are delivered such as animations, presentation slides, conversations etc. This present a dilemma to a consumer as to which MOOC to choose since some video styles may appeal to a consumer more than the other.

1.4. Research Objectives

The main objective of this component is to classify different video production styles that are not covered by MOOCREC1 V1 such as animations, khan academy styles and MOOCs with multiple video styles.

2. METHODOLOGY

2.1. Methodology

MOOCREC V1 has classified only coding, slides and talking head styles. In MOOCREC V2, we are classifying all 5 types of video production styles (coding, slides, talking head, animation and khan academy writing). For this, image-based classification approach is used where a single video is split in to frames and each frame get classified into a video production style. OpenCV [2] library in python has been used to split video into frames. Then image frames are classified into video styles and composition of each style is calculated. Each frame of a video is classified into one of the video production styles. The composition of each video style is obtained by calculating the average of each style in a single video and then the average values for each MOOC course is calculated.

Among Inception V3, VGG16 and VGG19 pre-trained deep CNN models, VGG16 model has been used because it has shown the highest accuracy of 94.0% and lowest loss of 0.24. Comparison between accuracy and loss of training of different CNN models are given in the table.

Table II: Comparison of CNN models

CNN Model	Validation Accuracy	Validation Loss
Inception V3	74.5%	0.42
VGG16	94.0%	0.24
VGG19	92.5%	0.35

We used 1000 images (200 images per each video style) as the training dataset and for test dataset 200 images (40 images per each video style) have been used. Those datasets were generated manually with OpenCV [2] as the datasets were not available publicly.

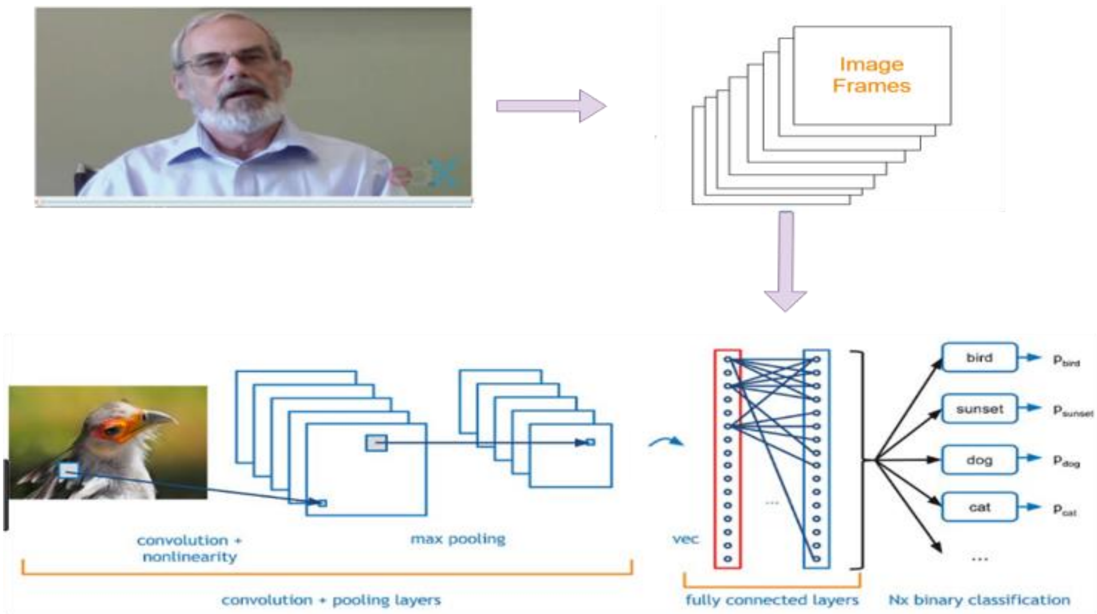

Figure 1 Fragmentation and classification of a video

If n is the total number of frames split by the classifier and a is the number of frames classified as animations, then the composition of animation in a single video is given by:

$$animation = a / n * 100$$

Finally, total composition of the video of the video styles of a course is calculated by calculating the average of each video production style.

2.2. Testing & Implementation

After passing the video into the pipeline, images get fragmented and its output is given in the figure.

```
Creating..../images/img_1.jpg
Creating..../images/img_2.jpg
Creating..../images/img_3.jpg
Creating..../images/img_4.jpg
Creating..../images/img_5.jpg
Creating..../images/img_6.jpg
Creating..../images/img_7.jpg
Creating..../images/img_8.jpg
Creating..../images/img_9.jpg
Creating..../images/img_10.jpg
Creating..../images/img_11.jpg
Creating..../images/img_12.jpg
Creating..../images/img_13.jpg
Creating..../images/img_14.jpg
Creating..../images/img_15.jpg
```

Figure 2: Creation of image frames from a video

After creating the frames it'll predict the video production style of each frame and returns the average composition of the video as follows.

```
Talking Head: 0.0
Code: 0.0
Slide: 3.33
Animation: 96.67
Writing: 0.0
```

Figure 3: Output of the video classifier

3. RESULTS & DISCUSSION

3.1. Results

Among InceptionV3, VGG16 [3] and VGG19 [4] deep convolutional neural networks, VGG16 model was selected because it has shown the highest accuracy of 94% and lowest loss of 0.24. Figures 4 to 6 shows the validation accuracy and validation loss curves of the Inception V3, VGG16 and VGG19 models respectively generated by matplotlib library in python.

Figure 4: Accuracy and loss of Inception V3

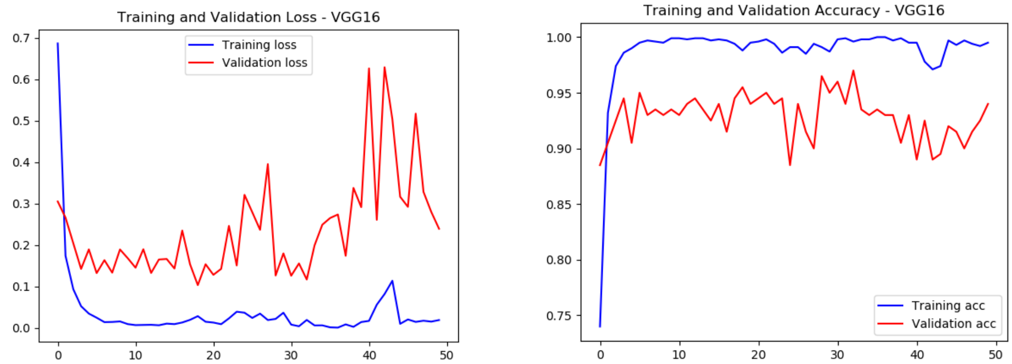

Figure 5: Accuracy and loss of VGG16

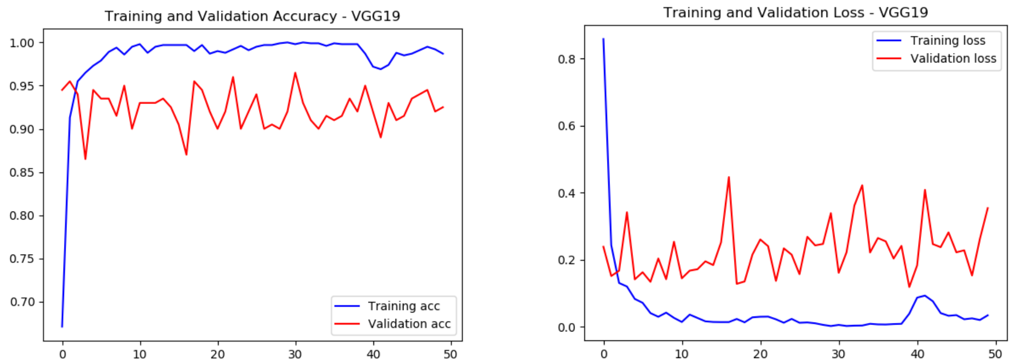

Figure 6: Accuracy and loss of VGG19

The actual composition of MOOC videos has been calculated manually and compared with the output of the classifier.

Table 3: Video Style composition testing

Platform – Coursera		
Course – Front End JavaScript Frameworks: Angular		
Video Name - Welcome to Angular		
Video Style	***Actual Composition (%)***	***Using Algorithm (%)***
Talking Head	10.7	11.3
Code	0	0
Slide	89.3	88.7
Animation	0	0
Writing	0	0
Conversation	0	0

Table 4:Video Style composition testing

Platform – Khan Academy		
Course – Introduction to logarithms		
Video Name - Introduction to logarithms		
Video Style	***Actual Composition (%)***	***Using Algorithm (%)***
Talking Head	0	0
Code	0	0
Slide	4	3.33
Animation	0	0
Writing	96	96.67
Conversation	0	0

3.2. Discussion

Among InceptionV3, VGG16 and VGG19 deep convolutional neural networks, VGG16 model was selected because it has shown the highest accuracy of 94% and lowest loss of 0.24. The video classification was carried out for 5 classes.

The testing that was done to determine the performance gain that can be achieved by parallelizing an image classification workload across a distributed computing platform can be said to have yielded positive results. While this method of classifying videos and images clearly does not yield performance gains that are comparable to GPU based parallel processing, it does however yield more than 75% performance improvements for a lesser cost as well.

4. CONCLUSION

MOOCREC V1 has classified only coding, slides and talking head styles. In MOOCREC v2 we were able to classify 5 styles with more accuracy. MOOC videos are now popular among most of the learners and this will be more helpful for those learners who have different learning styles.

Massive Open Online Courses have become a popular means of learning in a wide variety of subjects. MoocRec V1 introduced novel features to solve the problem of helping users find MOOCs. This paper proposes an improved version of MoocRec V1 by overcoming its issues and adding new features which would help users find the MOOCs that suit them the most. The results of the tests conducted show that the new features proposed are sufficiently effective at improving the system's usefulness.

5. REFERENCES

[1] X. a. L. C. a. L. S.-W. a. Z. V. Zhang, "Automated segmentation of MOOC lectures towards customized learning," IEEE, 2016.

[2] "Open CV Library," [Online]. Available: https://opencv.org/. [Accessed 8 2019].

[3] "How to use transferlearning when developing convolutional neural networkmodels," 08 2019. [Online]. Available: https://machinelearningmastery.com/how-to-use-

transfer-learning-when-developing-convolutional-neural-network-models/. [Accessed 08 2019].

[4] "Keras Documentation," [Online]. Available: https://keras.io/. [Accessed 08 2019].

[5] A. a. H. L. a. M. K. a. N. C. a. S. T. a. S. J. P. Hansch, "Video and online learning: Critical reflections and findings from the field," HIIG discussion paper series, 2015.

[6] S. a. L. F. a. A. W. a. S. N. a. S. R. Zha, "Exploiting image-trained CNN architectures for unconstrained video classification," *arXiv preprint arXiv:1503.04144,* 2015.

[7] "VGG16 – Convolutional Network for Classification and Detection," [Online]. Available: https://neurohive.io/en/popular-networks/vgg16/. [Accessed 08 2019].

[8] "VGG 19," [Online]. Available: https://www.kaggle.com/keras/vgg19/home. [Accessed 08 2019].

6. APPENDICES

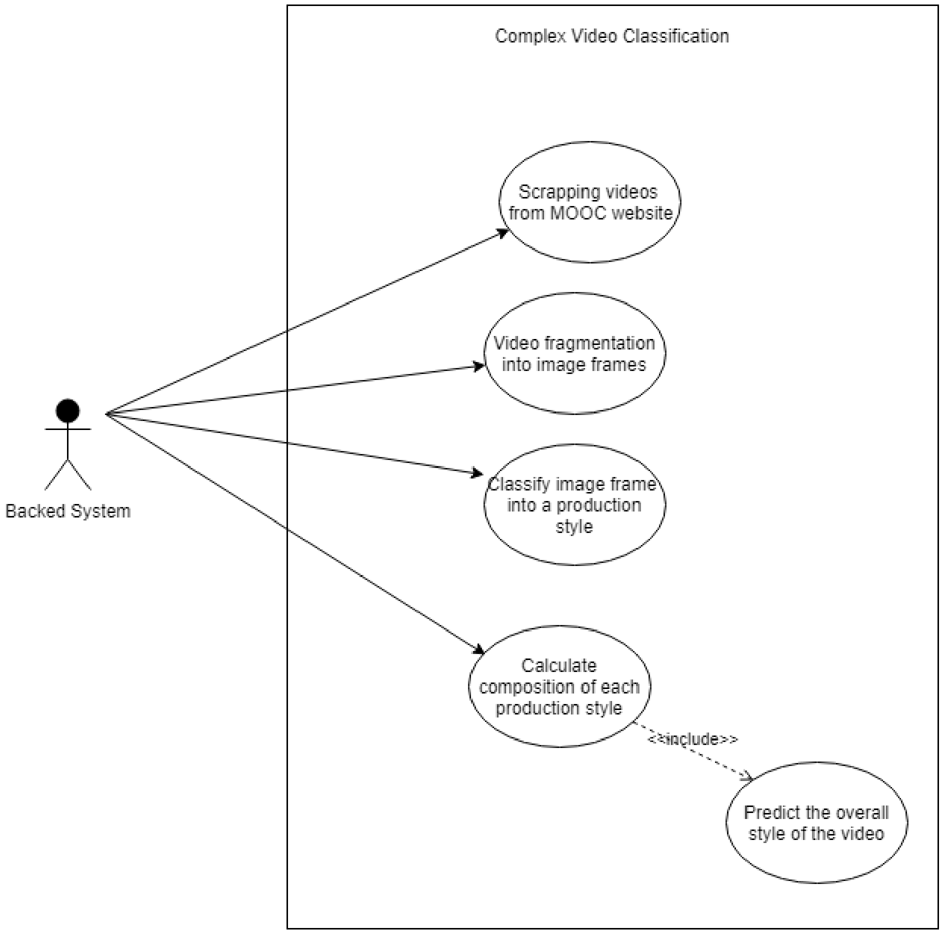

Figure 7: Use case diagram

Table 5: Use case scenario 1

Use Case Name	Splitting video files in to Image frames
Pre-Condition	Videos should be available in the database
Post-Condition	Image frames should be available after splitting
Actor	Video splitter – Backend System,
Main Success Scenarios	1. Select the desired video. 2. Split the video into consecutive image frames from the beginning to the end of the video.
Extension	1a. Processing an invalid video

Table 6:Use case scenario 2

Use Case Name	Classification of Image fragment into a style
Pre-Condition	Image frames should be available after fragmentation
Post-Condition	Image frames should be categorized into the correct video production style
Actor	Image classifier – Backend System,
Main Success Scenarios	1. Select the image frame. 2. Pass the image through CNN 3. CNN automatically classifies the image into a video production style
Extension	1a. Processing an invalid image frame 1b. Processing a corrupted file