Faculty of Science and Technology

Computer Engineering & Informatics

DISSERTATION REPORT

CST4090

BEHAVIORAL RECOMMENDATION ENGINE

Supervisor

Dr. Mohammad Osman

Date: 01-Oct-2021

Submitted By

Firas Hamid

MSIS:M00740777

Abstract

Video content is evolving as a critical source of information and is increasing in magnitude on social media and video sharing platforms. Recommendation Engines contribute to providing relevant content and hence affect the consumption rate massively. Techniques to analyze videos need to become more efficient giving more control to users to leverage value for their businesses. Behavioral Recommendation system articulates the idea of improvising recommender systems on image and video-based behavioral inputs. This system helps learners gain insights from behavioral images and classify videos based on these behaviors present in the videos. This research explores architectures and metrics to map behaviors with recommendation engines that facilitate users to view relevant content. This study uses transfer learning models for object detection that builds insights into behavioral recommendation workflow. It gives businesses and entrepreneurs an intuition to develop tools that analyze video databases and create valuable analytical insight impacting decisions for businesses and researchers. In this study, we propose and consolidate the idea of improving Recommendation Engines by classifies inputs as behaviors taken from images and videos. The study develops insights and experiments with multiple methods with varying degrees of success. From the perspective of data, the research mainly uses prebuilt trained models, and additionally, NFL sports data is used for improvisation. This research also focuses to train behavioral models and develop insights in the sports analytics domain.

Table of Contents

Abstract 1
Acknowledgement 4
1.Introduction 5
1.1 Background 6
1.2 Terminology 7
1.3 Problem Statement-Behavior of Interest: 8
1.4 Dissertation Organizations 9
2.Literature Review 10
2.1 Abstract 10
2.2 Keywords 10
2.3 Introduction 11
2.4 Review Methodology 12
A. Planning the review 12
B. Conducting The Review 13
2.5 Literature Review 17
Level 1: 17
Level 2: 18
Level 3: 18
2. 6 Outcome of the Review 20
3.Research Methodology 21
3.1 Introduction: 21
3.2 Data: 21
3.3 Experiment 1: 21
3.4 Experiment 2: 23
3.5 Experiment 3: 25
4. Results 29
5. Discussion 29

5.1 Data: 30
6.1 Blended-Hybrid Approach 31
6.2 Heterogenous Use Case 31
7. Bibliography 32
8.List of Figures 36
9.Apendices 37
9.1 Code 37
9.2 Ethical Form: 39

Acknowledgement

I would like to extend my gratitude to people who contributed and supported me directly or indirectly throughout my journey at Middlesex University. I would also like to extend my sincere thanks and gratitude to my Mentor, Professor Mohammad Osman who guided me in completing my dissertation. I would also like to extend my gratitude to my parents, peers, and friends who always help and motivate me to achieve my objectives in life. I am grateful to the faculty and professors who are always ready to help and share their knowledge and experience. I have learned a lot from my professors and coworkers whose advice resulted in the successful completion of this project.

1.Introduction

Recommender Systems have a great influence on the quality of information that users consume in today's time. How the recommender system, leverages information and presents relevant information and how it evolved is a point of curiosity that leads to this research. Diversification of this input information and quality of results has been an instrumental factor for the growth of RE's. Simultaneously videos are one of the primary classes of information that users consume, and this consumption is solely dependent on the videos that RE's recommend to the users. That infers the idea of video analysis can potentially act as fuel to improve the use of recommender systems. Besides rating and keywords RE'S need to be capable of learning from videos themselves. This research quantifies these metrics as behaviors that help customization of RE'S with video analysis based on behaviors. Exploring such features that enhance the RE'S is the primary objective of this research.
RE'S always are dependent on other domains and incorporate text analysis, ratings, and other domains for the functioning and delivery of results. Similarly, this research is thought, discussed, and conducted from a multi-dimensional perspective. Adding image processing, object detection, and using them in RE'S are the domains that are explored and the methodized for enhancing RE's adhered in this research. The focus of the research is to come out with metrics and architecture that can be used to build RE's that extract present content relevant to those extracted metrics.
As "briefcam" one of the video analysis startups leverages video analysis for security and building diverse products. The results from this research can be reused and further enhanced to build such products. The infrastructural costs of building these products are high but simultaneously have the capability to give solutions that have a high impact in multiple domains of business.

This research uses images split from a 30 second YouTube NFL video for training for different experiments conducted during the research. Thus, the research aims to analyze the behavior happening in NFL games and evaluate results in this domain. Helmet hit behavior is used to classify the behavior where one player hits the helmet of another while in the game.

1.1 Background

The purpose of this research is to develop an architecture that conditions the learning in RE's through video analysis. The research is intended to help researchers and entrepreneurs think of building products and experiment with the building of RE'S by utilizing video analysis. Online video content is increasing exponentially, and efficient techniques need to be thought of to gain control of utilizing information out of these data mines. Businesses can utilize these techniques by enhancing the use of recommending relevant content in their specific use cases. The domain of this research is diverse and complex simultaneously and is dependent on the developmental infrastructure. This is one of the highest challenges for entrepreneurs to sustain such costs. One of the challenges of this research is to find the dataset or transfer learning model relevant to the behavior for which you want to extract information.

There needs to be a way out for small business owners and entrepreneurs to make use of this large footage and a huge chunk of videos for efficient decision-making and investing. Behavioral analysis from videos leads to efficient insights for decision-makers and investors. The sports industry is one of the main industries which relies heavily on analyzing footage and video content for performance improvement of players, investing in teams, technique analysis, and gain deeper insights into sports they play. Coaches have a limited amount of time and multiple engagements. The video analysis team can define areas of interest from coaches and bring out behaviors from video footage that is relevant and valuable for decision. By building these insights can lead to the development of key performance indicators that are undiscovered through manual analysis. Injury prevention is one of the major factors that can cost teams money and video analysis can act as a mobilizing factor for analyzing behaviors that lead to such consequences. So, the applications scale to healthcare and medicine for disease predictions and can help doctors compare symptoms, analyze surgery footage, and prevent surgery failures. Utilizing large video datasets for certain behaviors is the domain that this research is addressing.

Moreover, efficient products are also needed to develop datasets and applications that run these software and artificial systems without fail. Clients don't' like to work with systems that bring uncertainties. High precision is the deciding factor of such products. While developing critical thinking and managing and maintaining quality products acts as the main factor for success. These organizations are high management and quality-driven and expect uncertainly and failures dealt with in advance.

1.2 Terminology

Recommendation Engine	RE	Referred to Artificial Intelligence-based software system which provides similar content to the user.
Mean Square Error	MSE	Measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value.
Structural Similarity	SSIM	SSIM is used for measuring the similarity between two images.
Behavior		It is referred to any instance that occurs in a video that could be of interest to analyze and infer further information.
Transfer Learning		Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.
Image Processing		Analyze digital images using different algorithms and statistical metrics.
Object Detection		Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects.
Video Analysis		The capability of automatically analyzing video to detect and determine temporal and spatial events.
Deep Learning		Artificial Intelligence technique used to analyze and learn non-linear patterns.
Thumbnail		It is the first image of the video or the image used as a hyperlink to access a video.
Machine Learning	ML	The subdomain of artificial Intelligence uses data to learn generalizations.
Artificial Intelligence	AI	It refers to the area of subject and methodologies that machines use to develop intelligence and generalizations of their own.

1.3 Problem Statement-Behavior of Interest:

This research analyzes Helmet hit detection behavior from NFL video footage and classifies it as a target variable to analyze in the video. The behavior is any point of interest of the user that it intends to develop and gain insight from the videos. To limit the scope of this research we analyze this specific behavior and develop a focus on the architecture of Behavioral RE.

Aims and objectives:

The main aim of the research is to explore and come out with an architecture for RE'S to analyze and recommend relevant videos. This research aims to explore and experiment with different methods from multiple domains that can contribute to diversifying the learning in REs. This research aims to be a contribution by enhancing the idea of using video analysis-based RE's.

Aim 1:

Explore and experiment with image processing and its implementations in RE'S.

The objective is to come out with metrics that RE'S can use for learning similarities between the images and comparing them to the video thumbnails. The aim is to build a foundation that helps RE's learn directly from images and further help in enhancing the procedure.

Aim 2:

Develop an Object detection for helmet hit detection and reuse detection metrics in RE's.

The aim is to also experiment with either object detection or build insight into how it can contribute to exploring digital video databases. The helmet detection behavior will be explored in a video and further provide a way for it to be scalable to all the databases. The aim is to articulate the behavioral analysis insight for videos and quantify its scalability to a video database.

Aim 3:

Experiment with learning techniques such as transfer learning for faster building of the behavioral model.

The aims are also to explore methodologies that help in simplifying the methodology and build a developmental process of such RE's. Transfer learning and bringing in models that are trained on sports analysis data is the prioritized methodology.

1.4 Dissertation Organizations

This research requires experimentation with multiple domains which express their dependency on image or video data. This research consists of exploring and developing metrics for Behavioral RE'S. The structure of the dissertation is as follows.

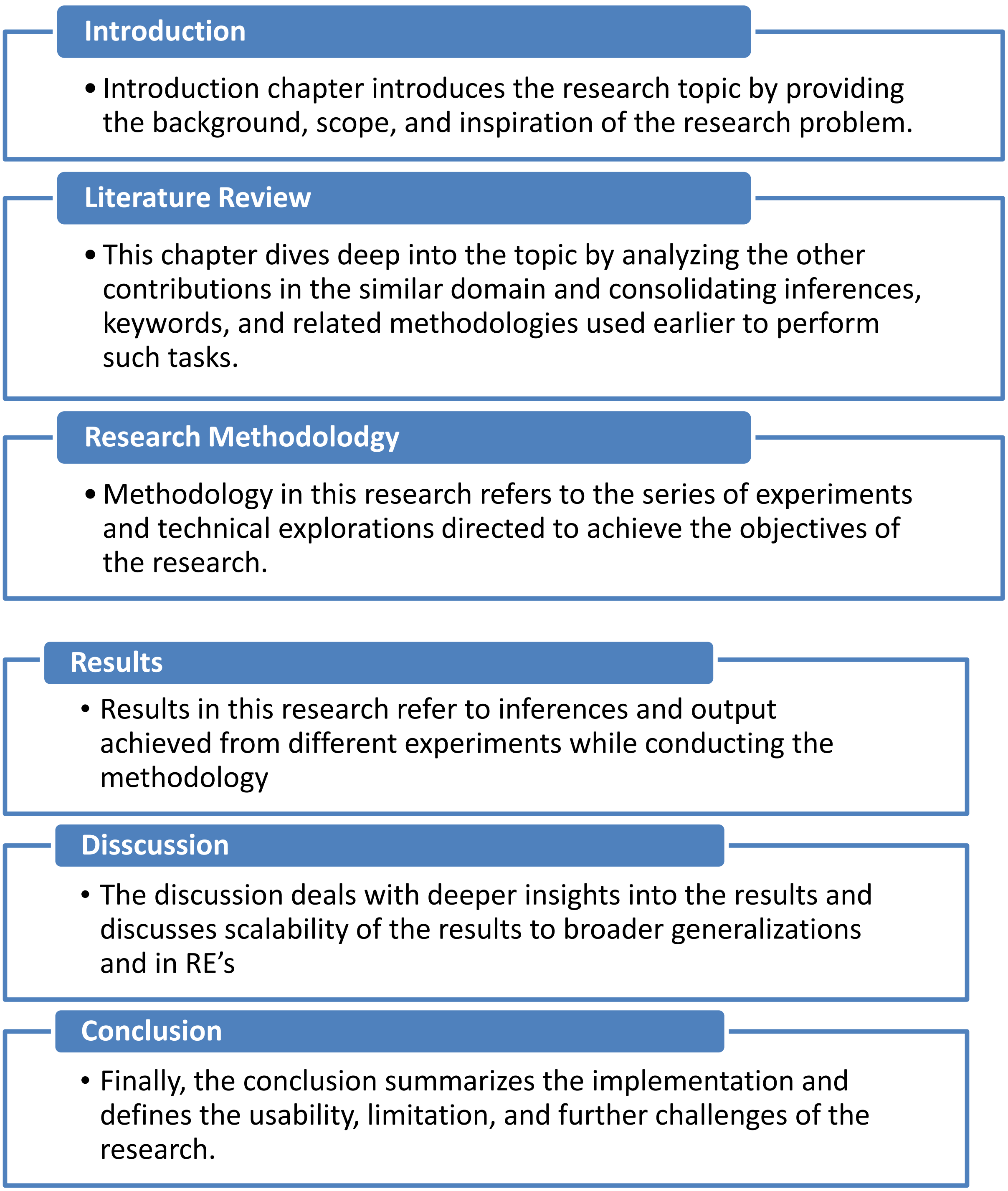

Figure 1 Thesis Structure

2.Literature Review

2.1 Abstract

Recommendation engines have been the primary facilitators of video consumption and mobilizers of good video production by recommending relevant and user-engaging content. Over time, this artificial intelligence contribution and its efficient improvisation by innovation and diversification have made video platforms the most benefited business in the technology industry. Good RE's have scaled up the screen time spent by the user, amount of video content, and no users for platforms like YouTube and Tik Tok. RE's leveraged from audio search, text, picture, and video rating over time and simultaneously produced audio, text, video as an output. This review highlights contributions and evolution of recommendation engines with video and images referred to as behaviors as input. The literature review articulates the metrics that define good quality literature for hybrid machine recommender and video analytics systems. This review acts as a contribution for developers working to explore methodologies and architectures in behavioral recommendation system domain.

2.2 Keywords

Video Analytics	**Behavioral Recommendation Artificial intelligence**	**Deep Learning Recommender System**	**video content analysis deep learning keyword**
Candidate Generation Recommendation	**Intelligent Surveillance System**	**Video Intelligence**	**social video recommendation**
Deep learning video analysis in sport			

2.3 Introduction

The video recommendation system is one of the exciting topics discussed for several years by researchers. This review aims to study methodologies that have worked on improvising RE's with diverse input data, particularly from the video analytics domain. The objective is to equip RE'S that train and reveal concealed inferences using video analytics components. The industry experiences that these ideas snowball into prototypes and products, which then receive appreciation from researchers, entrepreneurs, and investors. RE's are focused on user personalization, and video content has taken center stage. The efficient recommendation has multiplied the user engagement. Good performance of RE's have helped business gain profit by engaging customers and hence is a factor for immense competition between companies. This competition has led to the improvisation in evaluation and performance metrics for RE'S. This review also highlights the evolution of performance metrics for REs. Having the right suggestion can incite a purchase and increase conversion rate and footfall for e-commerce and retail businesses. Traditionally RE's have provided recommendations based on rating and collective opinions (Rahul and Bala, 2016).

The review aims to find those academic contributions that have discussed RE's with diverse inputs, including videos and images. Also, the academic content aligned to map video analysis with RE's is read and explored in detail. The assumption to begin the research is that no such method or architecture is discussed until now. The requirement and outcome of the review are to find elements from different domains that map a structure together to meet the objective of the research. This is the approach used for exploring content that builds a technical solution to the problem. The video recommendation based on analysis of frames has been discussed by (Jiana et al., 2019), and this provides us insights into the topic of deep learning for video analysis in sport. This review tries to center the scalability of frame analysis to video databases and reflects the architecture that could help build such RE's. (Kaklauskasa et al., 2018) has worked on property video recommendation videos and focus and created inferences from behavioral metrics arising from social, cultural, psychological, and personal factors considered in building RE's. Enhancing our training systems on specific behaviors further makes insights in mapping the build with RE's.

The term behavior is being taken broadly and is not limited to video or image snippets during the literature review. The research eventually desires to focus on these two categories of recommendation engines only. It also motivates us to explore the impact it can generate on the users of such systems. (Wang et al., 2012) leads in developing an RE by providing an understanding of the behaviors through content generations. (Cui et al., 2016) presents a perspective of behaviors from social media activities of users and leverages that for recommendation engines. As (Ismail, 2020) mentions, sports video analytics is an industry that started evolving from handcrafted measures, and it scaled up by deep learning methodology and learned features itself and created impact for sports players and coaches.

The keywords also evolved as the topic is very niche, and we tried a two-way approach for searching i: e from RE's and video analytics perspective. For example, "social video recommendation" and "deep learning video analysis in sport" are the closest keywords that helped to achieve relevant academic contributions.

2.4 Review Methodology

A. Planning the review

The main purpose of this review is to understand how behavior-based recommendation engines have evolved in academic and product sense. The review pursues to quantify the scale where these two systems have come close and produced any deliverable.

Some of the research questions which this review addresses are:

RQ1: How the whole idea of recommendation engines and video analytics evolved in each decade.

RQ2: How have the research topic titles diversified.

RQ3: What challenges has it experienced in each decade.

RQ4: Are there any architectures that describe behavior recommendation engines

The assumption which gets the research started is that this topic is not widely established, so the research of academic contributions will take a two-directional perspective one is video analytics and the second is recommendation engines. The research takes a chronological approach by decade and investigates every research question in each decade. The search will try to fill the gap by studying academic contributions which address the image-video analysis and recommendation together. The research starts with exploring google scholar, IEEE, and Middlesex University library resources. The keywords which come to use are Behavioral Video Analytics, Video Analysis, and Image-based recommendations.

As the research in the field has evolved so the criterion of the research is focused on keywords and preferences. This review investigates the journal which addresses improvisation in methods and tends to deliver products. All those academic contributions are prioritized which focus to present the architecture, solve challenges and simplify deployments. The research also explores the papers which provide infrastructural insight into these domains and their limitations. The research will also deep dive over the papers which have trained models on the nature of data that is relevant to the research. The research also explores the framework and data collection methodologies that help train these systems. The research also looks at the performance metrics to measure such systems. All papers that conduct some research hypotheses will be prioritized. Also, discuss papers that access privacy concerns with the behavioral data.

B. Conducting The Review

This review takes a multistage approach based on chronological order of how the idea of RE and video analysis evolved.

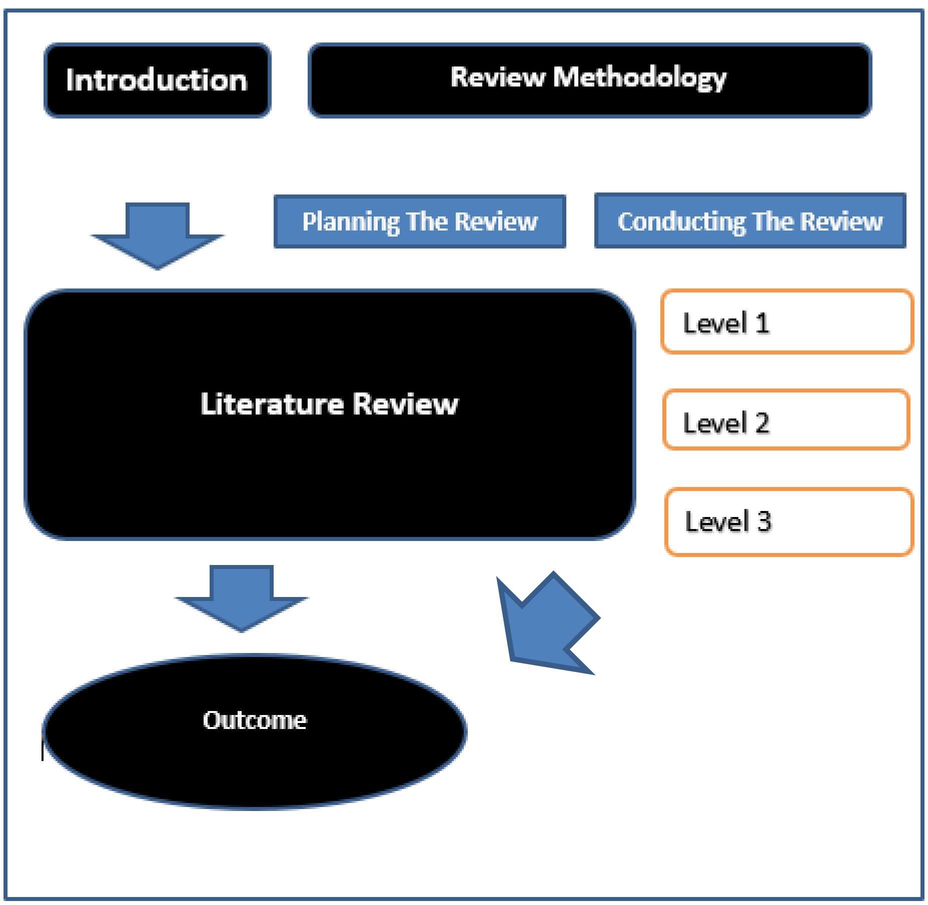

Figure 2 Literature Review Structure

Level 1:

At this stage, the papers are filtered from the 1990-2000 timeline. Considering the assumption that the resources and the contribution at this time would be less, so we set wider criteria for the search. This stage focuses to address RQ1 by filtering papers based on the abstract and conclusion. About 6 papers were filtered from MDX, IEEE, SPRINGER. The objective is to understand how the area of video analytics was perceived from 1990-2000.

Level 1	Type of Publication	Title	No of Articles
Journal	IEEE	Intelligent access to digital video: Informedia	1
	IEEE	Analytics Models And Characteristics Of Video Traffic In High Speed Networks	1
	IEEE	Fast And Globally Convergent Pose Estimation From Video Images	1
	Science Direct	Artificial Intelligence for intraoperative video analysis: Machine Learning's Role In Surgical Education	1
Article	ACM Digital Library	Artificial Intelligence Techniques In The interface to A Digital Video Library	1
Conference	ACM Digital Library	Expertise Recommender: A Flexible Recommendation System And Architecture	1
Total			6

Figure 3 Leve1

Level 2:

This stage filters about 3 papers based on the 2000-2010 timeline. This was the era in which the transition to video and the internet happened through smartphones and applications. We consider finding out answers to RQ2. The search will be conducted on the keywords which closely reflect the idea of the topic.

Level 2	Type of Publication	Title	No of Articles
Journal	Science Direct	"I know what you need to buy": context-aware multimedia-based recommendation system	1
	IEEE	JustClick: Personalized Image Recommendation via Exploratory Search From Large-Scale Flickr Images	1
Conference	Springer	A Collaborative Filtering Recommendation Methodology for Peer-to-Peer Systems	1
Total			3

Figure 4 Level 2

Level 3:

This stage filters 9 papers based on the 2010-2020 timeline. This stage approaches the literature review by diving deep into the architecture of recommendation systems and presenting the reflection of deep learning methods on video analytics.

Level 3	Type of Publication	Title	No of Articles
Journal	Wiley Online Library	A video recommendation algorithm based on the combination of video content and social network	1
Article	ELSEVIER	Comparison of performances of artificial intelligence versus expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video)	1
Conference Paper	Springer	Image-Based Fashion Product Recommendation with Deep Learning	1
	AAI Conference	Self-Supervised Video Representation Learning with Space-Time Cubic Puzzles	1
Journal	IEEE	Preference estimation for video recommendation using DCNN features and viewing behavior	1
	IEEE	JointRec: A Deep-Learning-Based Joint Cloud Video Recommendation Framework for Mobile IoT	1
	IEEE	Outfit Recommendation with Deep Sequence Learning	1
	IEEE	An Intelligent Video Tag Recommendation Method for Improving Video Popularity in Mobile Computing Environment	1
Conference	IEEE	Large-Scale Content-Only Video Recommendation	1
Total			9

Figure 5 Level 3

2.5 Literature Review

Level 1:

The idea of digital libraries saving content had evolved by 1990. The discussion had already reached a point where artificial intelligence techniques were used to retrieve information from digital libraries (Hauptmann, Witbrock and Christel, 1997). It considered techniques like Natural Language Processing, Image Segmentation, and other processes were thought where digital information processes become more efficient. This work defined the purpose of each AI technique in navigating information. It highlighted that Image processing can be used to break down the video into shots and help analyze these frames.
The topic was often referred to as Infomedia Digital Library as (Wactlar et al., 1996) discusses the techniques of retrieval. This research experimented focused on methodologies and considered information retrieval as a complex process by speculating rapid digital integration as the main reason. It consolidated different building blocks of an efficient search system.
The surveillance system has the capability of analyzing events that are providing behaviors to systems for decision making. This is presented by (Ivanov et al., 1999) as an event generation is a primary component of the system that accepts behavioral input. The instrumental thing about this system is how it generates events and parking maps together. This work scales up the degree of insights required for our research by bringing the idea of mapping between events and systems. This could potentially be used later for mapping events to content recommendations. The popular keywords used for research were "video browsing" and "Infomedia digital library".

Level 2:

(Byung Kwon, 2003) imperatively describes what is required for a web service to understand and interpret the information from an image. The outlines the intricacies of multimedia-based recommendations. The characteristic of this purchasing system is that it utilizes multimedia i:e image-based user preferences in a form that essentially has no criteria or context. The outcome facilitates the system by giving context called a context-aware multimedia agent. This hybrid approach can be used in behavioral Re's. The "Extended ARG" method discussed makes an image compact for use by a RE. This methodology builds the ground for how the context of images and behaviors benefits the REs.
RE's follow a developmental process that generalizes training by quantifying the behavioral activities. From a YouTube users' perspective, the systems reflect some interesting inferences which often make users curious. Such RE's follow the strong developmental process and simultaneously face technical challenges (Davidson et al., 2010a). It discusses the intricacies of ranking, user interface, and implementation for user content personalization. RE's account for about 60% of video clicks and impact business goals. To keep users engaged it is essential to capture new patterns and train systems regularly. CTR (Click-through rate) is the determinant metric to evaluate the clicks which resulted in users watching content.

Level 3:

Video recommendation based on a combination of social behavior and video content input is also discussed by (Cui et al., 2016). The video recommendation has taken a bidirectional approach which is highly relevant to the approach of our research and is based on trust friend's computation model and video quality evaluation model. It combines the social behavior of active users and video reputation. Trust-friend's models evaluate the similarity and influence between the users and followers and focus on the effectiveness of the video recommendation. The video model brings acceptance ratio as a metric to recommended videos to the users. The research finally presents the integrated work into the final model.
(Tuinhof, Pirker and Haltmeier, 2019) uses transfer learning for deploying a two-stage framework that recommends fashion images based on input from other images. The method focuses on making inferences from a single image to a return list of fashion products. There are two algorithmic network architectures taken into consideration Alex Net and Batch Normalization inception. Considering the dataset size transfer learning approach is mobilized to pre-trained models on the Deep Fashion Attribute prediction dataset. Primarily the CNN is used to extract features that are input into the secondary model. The impact of image-based fashion recommendations is evaluated through customer surveys. Thus for this kind of recommendation engine hybrid approach of developments Is highly preferred.

Learning the behaviors and deciding the criteria of what would be gained from images for learning is the critical point for this project. (Kim, Cho and Kweon, 2019) focuses to train 3D CNN on large datasets. The research uses kinetic datasets to use the self-supervised method of training using labels from videos. This approach caters to problems where the labeled data is scarce and there is an essential need for such data for building models. It auto-generates class labels by creating factors of images in either space or time where each unique pattern represents a label. When implementing 3D ReseNet is used an architecture on 400 human action classes where each video is trimmed to 10 seconds.

DCCN networks are used for the extraction of features from videos and can be transformed based on user's viewing behaviors (Ito, Ogawa and Haseyama, 2017). This work with user's behaviors for improving the accuracy and preference estimation of a video. They apply supervised Multiview canonical correlation analysis (svMVCCA) to features. The three features primary have targeted a video, user's viewing behavior, and user's evaluation scores and thus provide an approach to improve performance. It concluded the DCNN-based video features, as well as viewing behavior, features factors contribute to improving accuracy. They consolidated experimental results based on multiple dimensions of facial, head features, and body movements.

Federated training models where the deep learning algorithms learn by examining weights from local servers are updated into one final model (Duan et al., 2019). The training features are based on local user behaviors, descriptions, and user ratings. The research uses MovieLens composed of user ratings and information of the user. The data undergoes some statistical operations and defines data distribution on each cloud. The performance metrics are different for the centralized model and different for the distributed model. The experiment was performed using three interconnected cloud servers. Each has a local IOT dataset and one server acts as an aggregator which consolidates the different models. The data is partitioned over the systems on certain criteria and base on scenarios of the user.

(Jiang, XU and Cao, 2018) focuses on generating outfits by learning features from 160000 fashion items. The research is bidirectional wherein the model learns on relationships between categories and then do predictions based on the relationship between items themselves. Bidirectional LSTM and ResNet are the primary methodologies adhered to for achieving the goal. The dataset used in this research is providing outfits from a fashion website. The outfits consist of the order of the items in them. The output is an outfit recommendation method by sorting items in each category. The model also involves text queries from a user for recommendation where dimensionality projection methodologies are applied.

The collaborative filtering approach is instrumentally used for personalized video suggestions and encounters varied limitations in certain scenarios. Cold start problems where a new user signs up on the platform and there is not much data to infer to recommend videos. Collaborative Filtering methods are unable to handle such problems (Lee and Abu-El-Haija, 2017). In YouTube 300 hours of video content is uploaded every minute and simultaneous can have problems when the data is uploaded at such a high pace. This work discusses the methodologies of extracting the features from audio and visuals from the data uploaded to avoid cold start problems. Youtube8M dataset is being used for the research which consists of 8 million videos with 500k hours of video annotation.

Some of the contributions have also worked on leveraging recommender systems for creating the popularity of the content and focusing on the objective that business wants to obtain for their goal. Big data is essentially required to get results on that scale. (Zhou et al., 2019) takes two-directional approaches to attain social media attention particularly from the tags and original keywords of the video. It revolves around mobile and social media users and uses methods to scale the popularity of the content generated. The methodology addresses the problem of suggesting keywords on the first level. The second part works to suggest keywords based on the content of the video. Finally, third part TF-SIM algorithm which provides the top 15 keywords.

2. 6 Outcome of the Review

Building a behavioral Recommendation Engine is a challenging and complex process. The development of such a system requires clear use cases and data related to it. Scalability on diverse video data is going to be highly challenging and needs synchronization between developmental procedures and the architecture of the systems. The topic deals with the working of three essential components Image processing, Recommendation Engine, and Object Detection. The review additionally needs to explore the maintenance of RE’S which potentially contribute to the contributing efficient procedures for building RE’S. The data needs to be annotated for certain behaviors and the domain of implementation needs to be specified.

RE’S are hybrid systems that require data with sufficient features to determine good quality results. Behavioral RE’S needs good mapping techniques between the behaviors and videos.

3.Research Methodology

3.1 Introduction:

The proposed research intends to develop a pipeline for a behavioral recommendation system. The aim is to integrate methodologies that analyze behaviors from images and videos and maps them to RE. The research focuses to diversify the area of RE'S. The research analyzes multiple methodologies to read behaviors from images and further evolve to develop a criterion for recommending videos. Several experiments are further performed to improve and optimize different stages of the pipeline to increase performance and make the image processing faster. The research is essentially dependent on Python and is conducted in Jupyter Notebooks and scaled to Google Colab. The research is highly dependent and utilizes features of the OpenCV library.

3.2 Data:

The data used in different experiments of the research are taken from open-source resources like YouTube. This research aims to use the data for experimental use and infer results to build a pipeline. The data used at multiple stages of research is taken from multiple sources. The data part of the research was assumed to be complex and processing it on the local machine is heavy and takes time and thus was moved to Google Colab in later stage.

3.3 Experiment 1:

This experiment aims to frame the pipeline by building insights into the criteria of video recommendation. The result is in the form of a python dictionary which is obtained after processing some sample images. The processing starts with creating MSE and SSIM metrics to compare the two images. The comparison test of two images is intended to be able to be reused for videos. Images are loaded and formatted into the same dimensions. The comparison test function also uses and converts images to greyscale images. The three images are fetched and compared to a single original image. The image comparison happens between three images taken from American football videos. Taking insights from sports analytics this experiment adheres to a similar approach. Other images are also tested and evaluated for the change in MSE and SSIM values for comparison. The 0 value indicates less similarity and as it increases the similarity between the images becomes less. Structural Similarity Index (SSIM), developed by (Wang et al) is more inclusive in comparing two images. The values are between -1 and 1 and 1 indicates perfect similarity. For example, while comparing the images to themselves we got Mse:0 and ssim:1.

The matplotlib figure displays that the MSE and SSIM scores graphically and it is observed in this experiment that the similarity decreases as the video progresses.

```
In [101]: score

Out[101]: {'image_1': {'ssim': 0.46857773370108174},
           'image_2': {'ssim': 0.4714907570339792},
           'image_3': {'ssim': 0.49775101108084},
           'image_4': {'ssim': 0.5293293314239084},
           'image_5': {'ssim': 0.550441766265617},
           'image_6': {'ssim': 0.5479844349553629},
           'image_7': {'ssim': 0.5472383266621849},
           'image_8': {'ssim': 0.5466952581741058},
           'image_9': {'ssim': 0.5464154951523921},
           'image_10': {'ssim': 0.5460738735477677},
           'image_11': {'ssim': 0.5458607572595084},
           'image_12': {'ssim': 0.5457938364874673},
           'image_13': {'ssim': 0.545487869236237},
           'image_14': {'ssim': 0.5452919921038216}}
```

Figure 6.1 SSim Score between the images

The matplotlib figure displays the MSE and SSIM scores graphically and it is observed in this experiment that the similarity decreases as the video progresses.
The experiment also builds insight into how comparison tests can be mapped to use with videos and recommendation engines. For that reason, the experiments split the video into images and uses one original image for the comparison. We split the 'Thesis_V2_Trim30.mp4' video file into 944 images. To simplicity and speed of processing the experiment just uses a 30-second video for splitting into images. The split is random and evolves into logical splitting for better recommendations in the following experiments.

The result generates a dictionary that compares the original image to every split image of the video, and we observe the SSIM values for each comparison. It makes 944 comparisons and generates a score dictionary for the video.

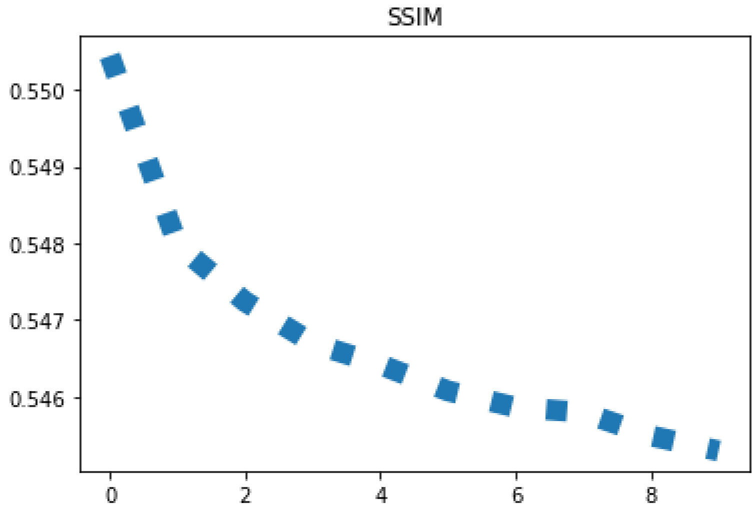

Figure 7 SSim Comparison for 30 second video

This experiment was done by using image libraries from python. The metrics MSE and SSIM were calculated by using functions from the same library. Opencv package is also used for processing images here

3.4 Experiment 2:

This experiment aims to understand features generated from images that help in recommending videos and utilize patterns from these images. The transfer learning approach used here is aimed to perform object detection on images split from videos in the last experiment. This approach uses python libraries for image segmentation and object detection. Following are the primary libraries used for object detection.

- scikit-image
- pillow
- pixelib
- tensorflow-gpu

The prebuilt model used in this experiment is an open-source contribution (He et al., 2017) and is used for diverse applications. This research uses this model by leveraging images split from NFL video and detects objects and segments objects inside the images. The detected objects and similarity metrics are potentially used for recommending similar videos. The model is not trained for the coco images dataset.

The model “mask_rcnn_coco.h5" is trained on open-source **MS COCO (Microsoft Common Objects in Context)** dataset released in 2014 containing 164K images. The data is annotated for object detection, captioning, key points detection, stuff image segmentation, panoptic, dense pose. The dataset has about 80 object categories. The dataset is not focused on sports objects or images or any specific use case.

At this point, this research builds insights and quantifies the magnitude of the problem related to recommender systems. The transfer learning approach is building insights to recommender systems and improvisation on specific datasets can yield more results. This experimental method takes us to explore training models which are trained on data related to images from the sports domain which is a video of our concern. Although the behavioral recommendation is a diverse problem this research keeps in focus a specific domain and builds scalable methodologies.

The “mask_rcnn_coco.h5” is a two-stage model which is based on RPN(Regional Proposal Network) and Faster R-CNN respectively. The input image is fed to a pre-trained network such as Resnet101 that uses classification. The second uses the ROI pooling layer mainly focused on the bounding box to the images of interest. The model has undergone improvisation in data inputs techniques and accuracy metrics. The model uses five accuracy metrics (Kumar, 2019). This model can take any input size and is based on a convolutional neural network.CNN acts as a backbone for segmentation and object detection for the model.

- rpn_class_loss, Lcls1Lcls1: RPN (bbox) anchor binary classifier loss

In the scope of this research, we care about the similarity each classification generates that could be reusable in recommender systems. This method combines object detection and segmentation. The bounding boxes potentially become a more focused metric to the problem of interest.

3.5 Experiment 3:

To enhance the development of the transfer learning model used in experiment 2, research further employs a customized object detection model. Keeping in view the model and the nature of data split from NFL video we train a model for sports behaviors. The purpose of this experiment is to enhance the model “"mask_rcnn_coco.h5" on a customized dataset.
Some of the images were annotated using labelimg tool.

Figure 8 Annotation tool

As experiment 2 was trained on limited behaviors which are different from the use case, we define for the sports video it does not yield stronger results.

Use case: This experiment uses NFL video and images split from the video to train and read sports behaviors such as helmets hits as the target of the training. The purpose of defining a specific use case for a RE's is to channelize the training and act as a method of quantification of behaviors for models.

Exploring the multiple transfer learning models primarily does not infer that they perform better on different data. Transfer learning models are customized to the labeled data of their concern. The primary issue with such a method is labeling the data and images in accordance with the behaviors we intend object detection system to extract from the video.
This method uses TensorFlow 2.6.0 and its built-in models as the primary resource for modeling used in our project. Additionally, it also uses the coco dataset for the models used by TensorFlow models imported. It also installs object detection API. This experiment uses packages

OS	**NUMPY**
pathlib	**TIME**
Cv2	**object_detection/Tensorflow**
argparse	**PANDAS**
googlecolab	**MATPLOTLIB**

The local machine could not sustain the computational complexity of the experiment, so this experiment is performed on google Colab as it provides improved computational power in the form of GPU'S. The google Colab packages also uses functions for optimizing hardware resources and mapping them to different components of the project. It also uses resent packages that get utilized in object detection models and for better object detection.

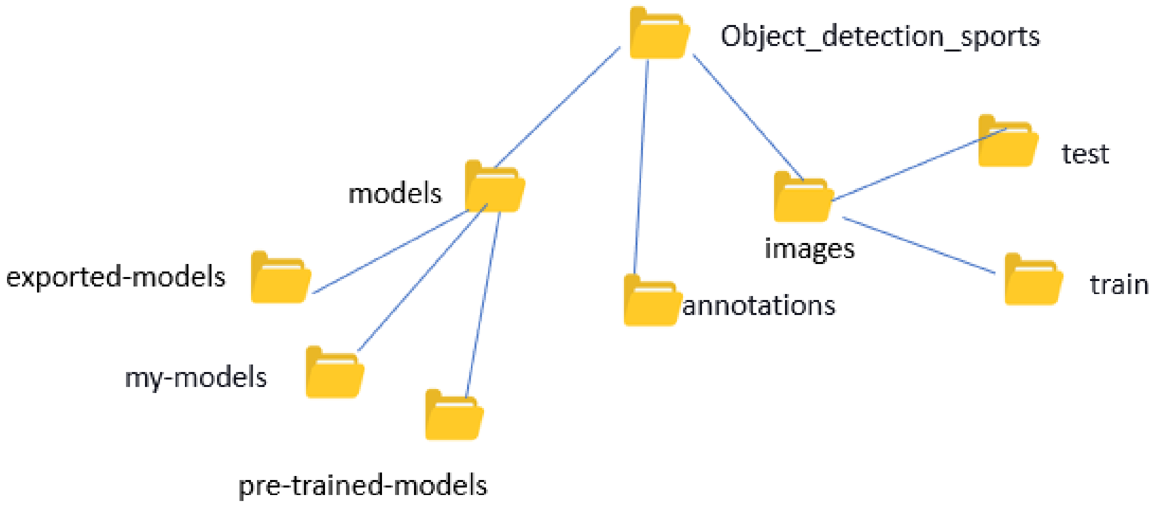

Figure 9 Folder Structure

This experiment is structure into multiple folders
Object_detection_sports is the main folder that consolidates images, models, and annotations folder.
Images folder further is divided into test and train folder for using images to train and test the model.
Models folder is divided into the pre-trained-models exported-models my-models folder where each contains transfer learning models exported models and final model respectively.
As specified by the official documentation following files are updated into the main project.
1. export_tflite_graph_tf2
2. exporter_main_v2
3. generate_tfrecord
4. model_main_tf2
Also following files are uploaded into the annotations folder

1. label_map_pbtxt
2. train.record
3. test.record

These folders all consolidate and generate a model that leverages files from other folders after learning and processing images into a my-model folder.

This experiment also needs to develop a dataset with annotations and classify behaviors. These classified behaviors are used as annotations and can further evolve as more exploration

happens within the domain. A 30-second video generates 944 images and for experimentation, this experiment annotated 20 images classifying helmet hit as behavior
We used tools for annotating images and added them to train and test folders of the project. Pip install xys and after annotating it stored the images into .xml files to be reused for the purpose of modeling and modes itself converted them into "test. record" and "train .record" files that get input into the model directly later.

After defining the folder structure and importing the required files we begin the process of utilizing these dependencies for our project. We set up the parameters for our models by importing the "pipleline.config" to my models.
Since we define the use case as classification as Helmet hit is detected o not detected in the image, we set the classes as two. From the infrastructure perspective, we want the project to run faster and we deploy batch normalization by setting the batch size as 8. Also, the default learning rate of model is 0.01333 and no of steps is 25000. Also, reduce the no of steps to 2000 to achieve faster results.
Here the paths to the required folders are also set which utilize annotated data and set checkpoints as detection. This configuration is all objected to solve this particular use case of detecting helmet hits from images. Finally, we run the model and extract it to the exported folder and by defining the folder path.

Finally, we start training the model after fixing some configuration mismatch errors. Taking some time for 20 images to train we move on to test and define the functions for displaying the tested and other images for detection. To make things run faster we test the model first on 1 epoch. The maximum number of images or objects to be detected in the image is set to 200 and the threshold as 0.5 to display an image with helmet hit detection.
The challenge faced in this experiment is the data itself.20 images is not enough data to train these systems for accurate results. The hardware needs to be equally scalable for better training of data and train the resnet on multiple epochs. To experiment with multiple methodologies of image processing and object detection transfer learning does not always suit the use case. The research could not find a transfer learning model which is scalable to such sports use cases. From all the above factors combined, we could get expected results and satisfying accuracy is a challenging part of the research. For image segmentation models the processing is computationally high and local machines couldn’t prove to be a good option.

4. Results

The results collected from different experiments in the research yield different results which RE’s potentially can use collectively or individually for improvisation in behavioral recommendation based on features in images and videos. Image to image comparison experiment provides a dictionary of MSE and SSIM scores which evaluates behaviors in the video and creates an inference for RE’S to recommend videos. Experiment 2 lays the foundation of modeling behaviors and creating a way out for healthy datasets. This experiment also contributes to the importance of customization in the transfer learning model in experiment 3. The models do not provide high accuracies due to different challenges at various stages of the research. The preprocessing part results in the annotation of images that leads to XML files where rectangular boxes in the images are positive behavioral detections and labeling refers to the text which defines the behavior. The transfer learning object detection model used these XML annotated image files for additional training. The features from the Object Detection transfer learning model do not lead to a convincing performance and accuracy that can be standalone to be reused by RE's.

5. Discussion

This research is constructing an idea by proposing architectural metrics to create RE’s which take inputs from graphical features of images and videos. The architecture is explored through several experiments performed between different domains. Image to image comparison delivers the similarity between the images by providing MSE and SSIM metrics done between the images of interest referred to as behaviors with the images split from videos. It outputs a Python dictionary data structure that stores these similarities in the form of keys and values. The second experiment obtains insights from Object detection modeling of Helmet hit detection by making use of the transfer learning model. In the second experiment, the realization of annotated data and customizing training for transfer models is developed. The third experiment starts with data preparation by splitting the video data into images annotating 20 images for helmet hit detection. The model is run and is highly dependent on the computational power for quicker results. Annotating dataset manually transformed into a challenge a 30-second video splits into 2K images. Exploration for transfer learning models trained for sports object detection is performed however no suitable results were found. The evaluation metrics were not adopted as the goal was to explore and make this model reusable in RE’S. The resultant images from the model were observed manually for experimental purposes. The result is not highly accurate as of the data at this point is less and exploration for more relevant annotated data is done. The optimization of the model for better results was done by adding more annotated data and tuning hyperparameters. The modeling uses “resnet” trained on the coco dataset and the second experiment added customized data for training. Image comparison experiment is performed on local machine using Jupyter and Customized object detection for performed on Google Colab.

5.1 Data:

The primary dependency of this research is on transfer learning models trained on the coco dataset. The model “"mask_rcnn_coco.h5" is trained on open-source **MS COCO (Microsoft Common Objects in Context)** dataset released in 2014 containing 164K images. The data generated from experiment 1 refers to the MSE(Mean Square Error) and SSIM (Structural Similarity Score). The influence of these two scores in RE’S can give a better understanding of behaviors by doing a logical comparison of the interesting behavior and videos or any sequence of instances in the video. More relevant the similarity with the video by similarity score more is the priority of video in the recommendation hierarchy.
The annotation of image data is done using "labelimg" tool which annotates and marks Helmet hit in the image. The annotated images are in the XML format given to the model for training. The challenge of annotating the dataset is that the images are taken from the wide-angle where stadium view is presented, and helmet hit detection behavior is a specific thing and is not good graphic quality images. The modeling struggles with noisy graphical data and impacts training accuracies.

Helmet hit instance occurs when any player hits his helmet with the opposite player while getting pushed or dragged during the play.

Figure 10 Helmet-Hit

6 .Conclusion

The research Behavioral Recommendation Engine is an exploration into the idea of diversifying the RE's by blending them with images processing and video analysis. This research articulates in detail the complexities and challenges in the construct of the idea and its implementation. The research performs several experiments from different domains like image processing, object detection, and transfer learning to conclude an architecture that forms the basis of executing and improvising the idea of Behavioral RE's. The research discusses the evolution and perception of hybrid RE'S. It focuses on the architectural design and input parameters that resulted in the diversification in the development of RE's. The research concludes that the idea is excessively heterogeneous in nature and needs high clarity with use cases and the availability of the right data. The developmental procedures need to be aligned with the construct of uses cases in Behavioral RE's.

6.1 Blended-Hybrid Approach

The research performs an image comparison test to evaluate similarity scores that RE's can use for recommending relevant videos. Additionally, object detection is also performed to analyze behaviors in videos. To obtain quicker results and discover the right methods transfer learning approach is the backbone of this research. The transfer learning approach is experimented with to reduce the magnitude of complexity in the project. The transfer learning approach gives better results as data quantity and quality is enhanced with proper annotations of behaviors. This approach has bought agility in inferring insights for methodologies to develop such systems. The research needs healthy infrastructural capabilities to experiment and progress swiftly to conclude engaging results. The Blended-Hybrid approach of developing RE's inherits deficiencies and strengths from the domains it incorporates. The impact of these deficiencies and strengths impacts highly on the quality of the results it generates. This research ignores the quality of the results and focuses to develop metrics and architecture to tackle such a problem. This approach provides a wide area for improvisation as it creates focus by secluding the multiple problems that exist in different with RE'S domains.

6.2 Heterogenous Use Case

For product development, the research needs to clearly define the behaviors of interest, the scale of the RE'S in the business domain, and the relevancy of the data needed. The unavailability of the right annotated data can potentially deaccelerate the research process. Lack of a clear use case for Behavioral RE's can affect the multiply developmental costs. Thus, the impact of RE'S in any domain needs to be quantified in advance. The planning and development of such products need to be done with standard developmental practices with identifiable resources and the degree of required skill. The stakeholders should develop an understanding of what the quality of results of RE'S can impact the business, client, or efficiency of any decisions which the RE' is empowering. The planning needs to be aligned and coordinated well between each stakeholder of the research and development of the behavioral RE. To avoid any detrimental consequences due to a shortcoming in the developmental procedure or use case risk mitigation should be done by reducing the scale of and deploying efficient testing procedures in R&D.

7. Bibliography

Byung Kwon, O. (2003). “I know what you need to buy”: context-aware multimedia-based recommendation system. *Expert Systems with Applications*, 25(3), pp.387–400.

Cui, L., Dong, L., Fu, X., Wen, Z., Lu, N. and Zhang, G. (2016). A video recommendation algorithm based on the combination of video content and social network. *Concurrency and Computation: Practice and Experience*, 29(14), p.e3900.

Davidson, J., Liebald, B., Liu, J., Livingston, B., Nandy, P., Van Vleet, T., Gargi, U., Lambert, M., Gupta, S. and He, Y. (2010a). The YouTube video recommendation system. In: *Fourth ACM conference on Recommender systems*.

Davidson, J., Livingston, B., Sampath, D., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y. and Lambert, M. (2010b). The YouTube video recommendation system. *Proceedings of the fourth ACM conference on Recommender systems - RecSys ’10*. [online] Available at: https://dl.acm.org/citation.cfm?id=1864770.

Del Carpio, A.F. and Angarita, L.B. (2020). *Trends in Software Engineering Processes using Deep Learning: A Systematic Literature Review*. [online] IEEE Xplore. Available at: https://ieeexplore.ieee.org/document/9226311 [Accessed 10 Aug. 2021].

Duan, S., Zhang, D., Wang, Y., Li, L. and Zhang, Y. (2019). JointRec: A Deep Learning-based Joint Cloud Video Recommendation Framework for Mobile IoTs. *IEEE Internet of Things Journal*, 7(3), pp.1–1.

Fan, J., Keim, D.A., Gao, Y., Luo, H. and Li, Z. (2009). JustClick: Personalized Image Recommendation via Exploratory Search From Large-Scale Flickr Images. *IEEE Transactions on Circuits and Systems for Video Technology*, [online] 19(2), pp.273–288. Available at: https://ieeexplore.ieee.org/abstract/document/4703261 [Accessed 1 Sep. 2021].

Fukuda, H., Tada, T., Kinoshita, K., Horie, M., Ishihara, R., Kato, Y., Matsunaga, T., Nishida, T., Yamada, T. and Ogiyama, H. (2020). Comparison of performances of artificial intelligence versus expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video). *Gastrointestinal Endoscopy*, [online] 92(4), pp.848–855. Available at: https://www.sciencedirect.com/science/article/pii/S0016510720343789 [Accessed 3 Sep. 2021].

Hauptmann, A.G., Witbrock, M.J. and Christel, M.G. (1997). Artificial intelligence techniques in the interface to a Digital Video Library. *CHI '97 extended abstracts on Human factors in computing systems looking to the future - CHI '97*.

He, K., Gkioxari, G., Dollár, P. and Girshick, R. (2017). *Mask R-CNN*. [online] arXiv.org. Available at: https://arxiv.org/abs/1703.06870.

Ismail, S. (2020). Deep learning in sport video analysis: a review. *core.ac.uk*, [online] 18(1693-6930). Available at: https://core.ac.uk/reader/295344720.

Ito, Y., Ogawa, T. and Haseyama, M. (2017). *Preference estimation for video recommendation using DCNN features and viewing behavior*. [online] IEEE Xplore. Available at: https://ieeexplore.ieee.org/abstract/document/8229258 [Accessed 3 Sep. 2021].

Ivanov, Y., Stauffer, C., Bobick, A. and Grimson, W.E.L. (1999). Video surveillance of interactions. In: *Proceedings Second IEEE Workshop on Visual Surveillance (VS'99) (Cat. No.98-89223)*.

Jiana, M., Zhanga, S., Wua, L., Zhanga, S., Wangb, X. and Hea, Y. (2019). Deep key frame extraction for sport training. *Neurocomputing*, [online] 328, pp.147–156. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0925231218309482 [Accessed 30 Aug. 2021].

Jiang, Y., XU, Q. and Cao, X. (2018). *Outfit Recommendation with Deep Sequence Learning*. [online] IEEE Xplore. Available at: https://ieeexplore.ieee.org/abstract/document/8499079 [Accessed 3 Sep. 2021].

José R., Villar, Suárez, R., María, Sedano, J. and Mateos, F. (2009). Unsupervised Feature Selection in High Dimensional Spaces and Uncertainty. In: *International Conference on Hybrid*

Artificial Intelligence Systems.

Kaklauskasa, A., Zavadskasa, E.K., BanaitisaI, A., Kavaliauskienea, M., Libermanb, A., DzitaccI, S., Kuzminske, A., Naumcika, A., Cerkauskasa, J. and Ubartea, I. (2018). *Middlesex University Login*. [online] ethos.mdx.ac.uk. Available at: https://www-sciencedirect-com.ezproxy.mdx.ac.uk/science/article/pii/S0040162517309332?via%3Dihub.

Kim, D., Cho, D. and Kweon, I.S. (2019). Self-Supervised Video Representation Learning with Space-Time Cubic Puzzles. *Proceedings of the AAAI Conference on Artificial Intelligence,* 33(AAAI-19), pp.8545–8552.

Kitchenham, B. (2004). *Procedures for Performing Systematic Reviews*. [online] Available at: http://www.elizabete.com.br/rs/Tutorial_IHC_2012_files/Conceitos_RevisaoSistematica_kitchenham_2004.pdf.

Kumar, H. (2019). *Quick intro to Instance segmentation: Mask R-CNN*. [online] kharshit.github.io. Available at: https://kharshit.github.io/blog/2019/08/23/quick-intro-to-instance-segmentation.

Kyeong Kim, H., Kyeong Kim, J. and Ho Cho, Y. (2005). A Collaborative Filtering Recommendation Methodology for Peer-to-Peer Systems. In: *International Conference on Electronic Commerce and Web Technologies*.

Lee, J. and Abu-El-Haija, S. (2017). *Large-Scale Content-Only Video Recommendation*. [online] openaccess.thecvf.com. Available at: https://openaccess.thecvf.com/content_ICCV_2017_workshops/w18/html/Lee_Large-Scale_Content-Only_Video_ICCV_2017_paper.html [Accessed 3 Sep. 2021].

Rahul, K. and Bala, P. (2016). *Recommendation engine based on derived wisdom for more similar item neighbors.* [online] ethos.mdx.ac.uk. Available at: http://web.a.ebscohost.com.ezproxy.mdx.ac.uk/ehost/detail/detail?vid=0&sid=b08021f2-170c-41a4-bb91-98b3fe123398%40sdc-v-sessmgr03&bdata=JkF1dGhUeXBlPXNzbyZzaXRlPWVob3N0LWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=125223803&db=bth [Accessed 18 Aug. 2021].

Ricci, F., Rokach, L., Shapira, B. and B. Kantor, P. (2011). *Middlesex University Login*. [online] ethos.mdx.ac.uk. Available at: https://link-springer-com.ezproxy.mdx.ac.uk/content/pdf/10.1007%2F978-0-387-85820-3.pdf [Accessed 18 Aug. 2021].

Tuinhof, H., Pirker, C. and Haltmeier, M. (2019). Image-Based Fashion Product Recommendation with Deep Learning. *Machine Learning, Optimization, and Data Science*, 11331(0302-9743), pp.472–481.

Wactlar, H.D., Kanade, T., Smith, M.A. and Stevens, S.M. (1996). Intelligent access to digital video: Informedia project. *Computer*, [online] 29(5), pp.46–52. Available at: https://ieeexplore.ieee.org/abstract/document/493456 [Accessed 1 Sep. 2021].

Wang, Z., Sun, L., Zhu, W., Yang, S., Li, H. and Wu, D. (2012). *Middlesex University Login*. [online] ethos.mdx.ac.uk. Available at: https://ieeexplore-ieee-org.ezproxy.mdx.ac.uk/document/6397621 [Accessed 30 Aug. 2021].

Zhou, R., Xia, D., Wan, J. and Zhang, S. (2019). *Middlesex University Login*. [online] ethos.mdx.ac.uk. Available at: https://ieeexplore-ieee-org.ezproxy.mdx.ac.uk/document/8938802.

Briefcam (2021). *Search & Review Hours of Video in Minutes*. [online] BriefCam. Available at: https://www.briefcam.com/solutions/review-search/ [Accessed 25 Jun. 2021].

Kwon, O.B. (2017). Real-time event detection for online behavioral analysis of big social data. *Future Generation Computer Systems*, [online] 66(3), pp.137–145. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0167739X16300899.

Ramezani, M. and Yaghmaee, F. (2016). A novel video recommendation system based on efficient retrieval of human actions. *Physica A: Statistical Mechanics and its Applications*, [online] 457, pp.607–623. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0378437116300991 [Accessed 25 Jun. 2021].

Sulthana, A.R., Gupta, M., Subramanian, S. and Mirza, S. (2020). Improvising the performance of image-based recommendation system using convolution neural networks and deep learning. *Soft Computing*, 24(19), pp.14531–14544.

Tuinhof, H., Pirker, C. and Haltmeier, M. (2019). Image-Based Fashion Product Recommendation with Deep Learning. *Machine Learning, Optimization, and Data Science*, 11331(14 February 2019), pp.472–481.

8.List of Figures

Figure 1 Thesis Structure 9
Figure 2 Literature Review Structure 13
Figure 3 Leve1 14
Figure 4 Level 2 15
Figure 5 Level 3 16
Figure 6.1 SSim Score between the images 22
Figure 7 SSim Comparison for 30 second video 23
Figure 8 Annotation tool 25
Figure 9 Folder Structure 27
Figure 10 Helmet-Hit 30
Figure 11 Splitting Videos into images 37
Figure 12 Comparing images on SSIM,MSE score 38
Figure 13 Final Model 38

9.Apendices

9.1 Code

Splitting the video

```
vidcap = cv2.VideoCapture('Thesis_V2_Trim30.mp4')
success, image = vidcap.read()
count = 1
while success:
  cv2.imwrite("V4/image_%d.jpg" % count, image)
  success, image = vidcap.read()
  print('Saved image ', count)
  count += 1
```

```
Saved image  1
Saved image  2
Saved image  3
Saved image  4
Saved image  5
Saved image  6
Saved image  7
Saved image  8
Saved image  9
Saved image  10
Saved image  11
Saved image  12
Saved image  13
```

Figure 11 Splitting Videos into images

Comparison Testing

```
def mse(imageA, imageB):
    # the 'Mean Squared Error' between the two images is the
    # sum of the squared difference between the two images;
    # NOTE: the two images must have the same dimension
    err = np.sum((imageA.astype("float") - imageB.astype("float")) ** 2)
    err /= float(imageA.shape[0] * imageA.shape[1])

    # return the MSE, the lower the error, the more "similar"
    # the two images are
    return err

def compare_images(imageA, imageB, title):
    # compute the mean squared error and structural similarity
    # index for the images
    global m
    m= mse(imageA, imageB)
    global s
    s = ssim(imageA, imageB)
    # setup the figure
    fig = plt.figure(title,figsize=(30, 30))
    plt.suptitle("MSE: %.2f, SSIM: %.2f" % (m, s))
    # show first image
    ax = fig.add_subplot(1, 2, 1)
    plt.imshow(imageA, cmap = plt.cm.gray)
```

Figure 12 Comparing images on SSIM,MSE score

Final Model

```
--trained_checkpoint_dir /content/training_demo/models/my_sad  --output_directory /content/training_demo/exported_models/my_model
```

```
2021-09-20 01:46:26.710710: E tensorflow/stream_executor/cuda/cuda_driver.cc:271] failed call to cuInit: CUDA_ERROR_NO_DEVICE: no CUDA-capable device is detected
2021-09-20 01:46:26.710772: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:156] kernel driver does not appear to be running on this host (bb067219cfd7): /proc/driver/nvidia/version does not exist
WARNING:tensorflow:From /usr/local/lib/python3.7/dist-packages/tensorflow/python/autograph/impl/api.py:463: calling map_fn_v2 (from tensorflow.python.ops.map_fn) with back_prop=False is deprecated and will be removed in a future version.
Instructions for updating:
back_prop=False is deprecated. Consider using tf.stop_gradient instead.
Instead of:
results = tf.map_fn(fn, elems, back_prop=False)
```

Activate Windows
Go to Settings to activate

Figure 13 Final Model

9.2 Ethical Form:

Application for Ethical Approval for Research Projects

This is an application form for ethical approval for research undertaken by any Middlesex University Dubai staff and students. The person who completes this form should be the principal (or sole) Middlesex University Dubai researcher on the proposed study. After completion, this form (along with accompanying documents) should be submitted to the Research Ethics Committee (REC) for review. Student Researchers should submit to their Supervisor.

Section 1 – Applicant details

1.1 Details of Applicant (Principal Investigator or Student Researcher)		
Name: Firas Hamid	Department/Position: Computer Engineering & Informatics/Student	
Qualifications:	Email: ff272@live.mdx.ac.uk	Tel:
1.2 Details of Supervisor for student applicants (if applicable)		
Name: Muhammad Osman	Programme of study/module: MSc Data Science/CST4090	
Qualifications:	Email:	Tel:
1.3 Details of any co-investigators (if applicable)		
Name:	Organisation:	Email:
Name:	Organisation:	Email:
Name:	Organisation:	Email:
1.4 Details of External Funding (if applicable)		

Section 2 – Details of the proposed study

2.1 Research project title	Behavioural Recommendation Engine		
2.2 Proposed start date	05/05/2021	**2.3 Proposed end date**	30/09/2021
2.4 Describe the aim and rationale of this study?			
A study of video analysis and recommendation engines in artificial intelligence domain.			
2.5. Discuss the research questions and/or hypotheses of this study? This research analyzes Helmet hit detection behavior from NFL video footage and classifies it as a target variable to analyze in the video. The behavior is any point of interest of the user			

that it intends to develop and gain insight from the videos. To limit the scope of this research we analyze this specific behavior and develop a focus on the architecture of Behavioral RE.
2.6 Details of study design, data collection methods to achieve the research aims (e.g., interviews, questionnaire, observation etc.) and/or secondary data sources (e.g., National Statistics) to be used in the research, proposed hypotheses, data analysis, with references and citations (where applicable). Include details of any online data collection (ie online survey, via social media).
None

Section 3 – Initial Checklist to be completed by the applicant

3.1 Does this research involve human participants	☐Yes	☒No
If yes, please provide the following details:		
Who are your participants? Please specify any specific groups of human participants: (e.g., students, general public, specific groups etc.)		
How many participants will you have? (Under each category)		
How will participants be recruited and approached?		
Do you need access to groups of participants (e.g., through gatekeepers, e.g., organisations, managers, parents, schools etc.)	☐Yes	☒No
If yes, please provide details including no objection certificate(s):		
3.2 Does this research involve secondary data collection?	☐Yes	☒No

If yes, please Indicate your response below:		
3.2.1 Do you have the necessary approval to access the data*? (*If yes, please provide evidence of approval) (If no, please provide details and plan of action)	☐Yes	☐No
3.3 The outputs from research (e.g., products, reports, publications, etc.) are not likely to cause harm to others and are in-line with the local legislation	☐Yes	☒No
If no, please explain how this can be avoided or managed:		
3.4 Will the study require data collection by proxy (someone else doing part of or all of your data collection)	☐Yes	☐No
If yes, please provide details including its rationale: *Note: When collecting data by proxy, you need to ensure that the highest ethical standards and procedures are adopted by all research partners/fieldworkers*		

Section 4 – Anonymity, confidentiality, and consent for primary and secondary research

4.1 Will the research involve collecting or analysing personal data or sensitive personal data? or involve sharing of confidential information beyond the initial consent given (i.e., personal data refers to information that may identify individuals e.g., name, address, date of birth, opinion, specific event, set of characteristics that would clearly identify individuals or very small groups. Sensitive personal data refers to racial or ethnic origin, political opinion, religious beliefs, trade union membership, sexual life, physical or mental health, criminal matters.)	☐ Yes	☒ No	☐ NA
If yes, please provide details: (e.g., Justification for use personal data or sensitive personal data? How you plan to anonymise the data? Where the data will be kept and care/storage facilities etc.) *Alternatively, if personal or sensitive personal data is required for the research, you must comply with the GDPR act and understand your responsibilities under the GDPR and have received data protection training. Please complete the Data Protection Checklist for Researchers*			
4.2 Will lists of identity numbers/codes or pseudonyms for individuals and/or organisations (i.e., linking keys to personal identifiers) be stored securely and separately from the research data and destroyed after the study to avoid any risk of confidentiality being compromised?	☐ Yes	☐ No	☒ NA
If no, please provide details on how this can be avoided or managed:			

4.3 Will you tell participants that their data will be treated confidentially and the limits of anonymity will be made clear in your <u>Participant Information Sheet</u>? (e.g., their identities as participants will be concealed unless prior consent is given to include the name of the participant in any documents resulting from the research. Consider how participants' narratives, quotes or involvement in specific events may make anonymity difficult to maintain.) Attach: Participant information sheet	☐ Yes	☐ No	☒ NA
If yes, provide details on how you will ensure this:			
4.4 Will you obtain <u>Written Informed Consent</u> directly from research participants (if applicable)? Attach: Informed Consent sheet	☐ Yes	☐ No	☐ NA
If no, please explain why? If yes, please specify how and when this will be achieved?			
4.5 Will you obtain <u>Written Informed Consent</u> directly from gatekeepers (if applicable)? Attach: Informed Consent sheet	☐ Yes	☐ No	☒ NA
If no, please explain why? If yes, please specify how and when this will be achieved?			
4.6 Will you inform participants that their participation is <u>voluntary</u> and that they have a <u>right to withdraw</u> from the research at any time?	☐ Yes	☐ No	☒ NA
If no, please explain why?			
4.7 Will you have a process for managing <u>withdrawal of consent</u>? Please provide details:	☐ Yes	☐ No	☒ NA
If no, please explain why? If yes, please provide details on how this will be managed?			
4.8 Will it be necessary for <u>participants to take part in the study without their knowledge and consent at the time</u>, or <u>by deception</u> e.g., covert observation?	☐ Yes	☐ No	☒ NA
If yes, please provide justification and details of how this will be managed to respect the participants/third parties involved to respect their privacy, values, and to minimise any risk of harmful consequences:			
4.9 Will you provide a <u>Written Debriefing Sheet</u>? (if applicable, also attach)	☐ Yes	☒ No	☐ NA
If no, please explain why?			

4.10 Will you need <u>consent from people</u> who appear in <u>visual data</u> (e.g., photos or films or social media)?	☐ Yes	☐ No	☒ NA
If yes, please provide details on how this will be managed: If no, please explain why?			
4.11 Will you <u>audio or video record</u> interviews and/or observations?	☐ Yes	☐ No	☒ NA
If yes, please provide details on how participants' anonymity will be maintained:			
4.12 Will your research involve **<u>participants responding to internet surveys, emails, chatroom discussions, blogs, interactive games, social media and networking sites etc</u>,**	☐ Yes	☐ No	☒ NA
If 'yes', please explain how will you obtain permission from the website authors, or informed consent from participants, and ensure anonymity and protect confidentiality in an environment that generates significant amounts of background information e.g., data logs, IP addresses, cookies and caches and/or with low levels of system security?			
4.13 Do you have a Data Management Plan? *(E.g.: Where the data will be stored, who will have access to data, how will the data be shared, how long the data will be stored, how it will be deleted/destroyed after your research completion etc.)*	☐ Yes	☐ No	☒ NA
If no, please explain why?			

Section 5 – Avoiding harm: risk assessment and management, safety and legal issues

5.1 Will you use an **<u>experimental research design</u>** (ie., implement a specific plan for assigning participants to conditions and noting consequent changes?)	☐ Yes	☒ No	☐ NA
If yes, please provide details of treatment/intervention (and specify is these are intrusive interventions e.g., hypnosis or physical exercise, or include the use of drugs, placebos or other substances e.g., vitamins, food substances etc.) and provide details of required resources for this study:			
5.2 Will the research involve **discussion of sensitive topics**? (e.g., sexual activity, drug use, national security etc.)	☐ Yes	☒ No	☐ NA

If yes, please provide details of how possible adverse reactions will be avoided and what support will be in place to manage any adverse consequences:			
5.3 Is pain or more than mild discomfort likely to result from the study?	☐ Yes	☒ No	☐ NA
If yes, please provide details on how this can be avoided or managed:			
5.4 Could the study induce **psychological stress or anxiety** or **cause harm or negative consequences** beyond the risks encountered in normal life?	☐ Yes	☒ No	☐ NA
If yes, please provide details and state how participants will be supported:			
5.5 Will the study involve prolonged and repetitive testing?	☐ Yes	☒ No	☐ NA
If yes, please provide details, justification and state how participants will be supported and length of each data collection session, number of sessions and location of data collection:			
5.6 Will this research be conducted **off-site** (i.e., not on Middlesex University Dubai premises)?	☐ Yes	☒ No	☐ NA
If yes, please provide details of other locations and explain how you will minimise any risks to your own health while off-site.			
5.7 Will you **being alone** with individual participants or group of participants place you at risk?	☐ Yes	☒ No	☐ NA
If yes, please state how this can be avoided or managed?			
5.8 Are there any **adverse risks** or **safety issues** (e.g., from **potential hazards)** that your methodology raises for you and/or for your participants or others?	☐ Yes	☒ No	☐ NA
If yes, please specify and provide details of mitigating actions that will be taken (e.g., travelling alone, working in hazardous conditions, discussing illegal activities on-line etc.) and how you, and your participants/third parties will be supported?			
5.9 Is the research or outputs from the research **likely to cause harm** to others (e.g., to their physical well-being, mental health, dignity or personal values) to an extent greater	☐ Yes	☒ No	☐ NA

than that encountered in ordinary life?			
If yes, please state how this can be avoided or managed?			

Section 6 – Research Sponsorship and/or Collaboration (if applicable)

6.1 Does the research have a sponsor *(i.e., any person or organisation who provides support for the research in the form of income, use of data, facilities, materials, assistance with data collection etc.) that may have ethical implications for the research?*	☐ Yes	☒No	☐ NA
If 'yes' please provide details of the role of the funder and issues:			
6.2 Does the research involve an international collaborator or research conducted overseas?	☐ Yes	☒No	☐ NA
If 'yes', what ethical review procedures must this research comply with for that country, and what steps have been taken to comply with these: (e.g., Do you need local permission/approval? Are there any country specific cultural social or legal considerations that need to be taken into account? Who will be collecting the data overseas? Have you considered intellectual property issues?)			
6.3 Does this research already have or require Approval from an External Research Ethics Committee?	☐ **Yes**	☒**No**	☐ **NA**
If 'yes' please provide details:			
6.4 Will this research or part of it be conducted in a language other than English?	☐ Yes	☒No	☐ NA
If 'yes', full translations of all non-English materials will need to be submitted.			

Section 7 – Other Issues

7.1 Does the research involve any ethical and/or legal issues not already covered that should be taken into consideration?	☐ Yes	☒ No	☐ NA
If yes, please give details:			
7.2 Do you require training on the requirements of GDPR for researchers?	☐	☒	☐

	Yes	No	NA
If yes, please give details:			
7.3 Does the research raise any other risks to safety for you or others that would be greater than in normal life?	☐ Yes	☒ No	☐ NA
If yes, please provide details and state how this can be avoided or managed? If appropriate, complete a separate state **Risk Assessment Form** along with this application			
7.4 Will participants receive any reimbursements or payments for participating?	☐ Yes	☒ No	☒ NA
If yes, please provide details and justification:			
7.5 Are there any conflict of interests to be declared in relation to this research?	☐ Yes	☒ No	☐ NA
If yes, please complete and attach the "Disclosure of Potential Conflict of Interest Form" along with this application			

Section 8 – Pre-Submission Checklist

Please mention the documents (where applicable) you will be attaching with this application:

Please check and attach the following documents where applicable:			
1. Participant Information Sheet	☐ Yes	☐ No	☐ NA
2. Informed Consent Sheet	☐ Yes	☐ No	☐ NA
3. Debriefing Sheet	☐ Yes	☐ No	☐ NA
4. Copy of questionnaire/interview guide/details of materials for data collection (including translations, visual images etc.)	☐ Yes	☐ No	☐ NA
5. Letter of permission (if required from organisation where research is to be conducted)	☐ Yes	☐ No	☐ NA
6. Evidence of external approval – for access to secondary data	☐ Yes	☐ No	☐ NA
7. Completed Risk Assessment Form	☐ Yes	☐ No	☐ NA

8. Data Protection Checklist for Researchers			
9. Disclosure of Conflict of Interests	☐ Yes	☐ No	☐ NA
10. Evidence of external approval – from external ethics body	☐ Yes	☐ No	☐ NA
11. Evidence of relevant licence for research with animals/animal by-products	☐ Yes	☐ No	☐ NA
12. If you are attaching any other documents, please provide details below:	☒NA		

Section 9: Declaration – to be completed by student, supervisor and reviewers

As <u>principal investigator</u> or <u>student researcher</u> I confirm that:

1. I have read and agree to abide by the relevant Code(s) of Ethics appropriate to my research field and topic.
2. I have reviewed the information provided in this form and believe it accurately represents the proposed research.
3. I have read and agree to abide by the University's Code of Practice for Research: Principles and Procedures.
4. I agree to inform my Supervisor of any adverse effects or changes to the research procedures.
5. I understand that research/data may be subject to inspection for audit purposes and I agree to participate in any audit procedures required by the Research Ethics Committee (REC) if requested.
6. I have completed and signed a risk assessment for this research study (if applicable).
7. I understand that it is my own responsibility and not that of Middlesex University Dubai to assess the personal risks involved with undertaking this research and to do my best to limit them.
8. I understand that Middlesex University Dubai is not accountable or liable for any adverse personal circumstances I may encounter as a result of the risk factors involved with undertaking this research.
9. No data collection will be undertaken before receiving approval for this application. If there is any alteration in the research methodology after approval, then submission of a Change in Ethics Approval form is required.
10. I understand that the owner of the data from this research will be the supervisor for undergraduate and master's level students' projects.

Principal Investigator or Student Researcher

Name: Firas Hamid Signature: Date:30/09/2021

As <u>Supervisor,</u> I confirm that (Student Applicant only):

1. I have reviewed all the information submitted with this research ethics application and believe it accurately represents the proposed research.
2. I accept responsibility for guiding the applicant so as to ensure compliance with the terms of the protocol and with any applicable Code(s) of Ethics.
3. I understand that research/data may be subject to inspection for audit purposes and I agree to participate in any audit procedures required by the Research Ethics Committee (REC) if requested.
4. I confirm that it is my responsibility to ensure that students under my supervision undertake a risk assessment to ensure that health and safety of themselves, participants and others is not jeopardised during the course of this study.
5. I understand that personal data about me contained in this form will be managed in accordance with the GDPR Act.
6. I have seen and signed a risk assessment for this research study (if applicable).

Supervisor's recommendation to the REC		
This is a low risk project and all ethical, legal and safety issues have been sufficiently addressed	☒ Yes	☐ No

Supervisor Name :……………………..…….. Signature:………….……… Date:………………

As <u>peer-reviewer</u> I confirm that (Student Research Applications only):

1. I have carefully reviewed the ethics application
2. I have relevant knowledge of the research topic
3. I have no involvement in the study
4. Declare any conflicts of interest which may influence the peer review process
5. Act in confidence and not disclose the content or outcome of the process to anyone other than to REC and those responsible in research supervision)

Peer Reviewer Assessment	
This is a low risk project	☐
This is a high risk project and therefore recommends full review by the University REC	☐

Peer Reviewer Decision (For Low Risk Projects) **(Please select one)**

1. Approved	☐
2. Approved with minor amendments (Please provide details):	☐
3. Revisions and further information required (Please provide details):	☐
4. Not Approved for the following reasons:	☐

Peer Reviewer Name:……………………..Signature:………….……… Date:………………

FOR RESEARCH ETHICS COMMITTEE (REC) USE ONLY

Research Committee Decision (For Staff Application or High Risk Student Projects) (Please select one)	
1. Approved	☐
2. Approved with minor amendments (Please provide details):	☐
3. Revisions and further information required (Please provide details):	☐
4. Not Approved for the following reasons:	☐

Name of the Chair of the Research Ethics Committee or nominee (If applicable):……………………..……..

Signature:…………….……… Date:………………

ADDITIONAL NOTES FOR COMPLETING THIS FORM

1) Refer to Middlesex University Research Ethics section on the University intranet
2) Please read Middlesex University's Code of Practice for Research: Principles and Procedures available on the University intranet
3) Please read and ensure compliance with Data Protection under the General Data Protection Regulation (GDPR)
4) Please note that a student (UG, PG taught or research) cannot be the Principal Investigator for ethics purposes
5) External ethics approval is required from some organisations, agencies and local authorities that have their own ethics processes and require completion of additional ethical approval forms and processing in addition to the MU process. It is your responsibility to check whether additional permissions apply to you.
6) Accompanying forms and checklists are available on the University Intranet. This include but not limited to:
   - The Middlesex University Risk Assessment Form is available on the University intranet

- Disclosure of Potential Conflict of Interest form is available on the University intranet
- Data Protection Act Checklist for Researchers is available on the University intranet
- Child Parent Consent Form
- Gate Keeper Letter

7) Templates for Participant information sheet, Informed consent sheet, Debriefing guide and other related materials are available on the University intranet