Research Proposal-2.docx

by Ashish Dhyani

Submission date: 03-Aug-2021 04:20PM (UTC+0100)
Submission ID: 158397510
File name: Research_Proposal.docx (370.42K)
Word count: 5407
Character count: 28476

ACTION RECOGNITION IN DARK USING DEEP LEARNING

ASHISH DHYANI
Student ID - 975129

Under the supervision of

KARTHICK KALIANNAN NEELAMOHAN

Research Proposal

Master of Science in Data Science

LIVERPOOL JOHN MOORES UNIVERSITY, UNITED KINGDOM

AUGUST 2021

Table of Content

Abstract

Action recognition in the dark is the need of the hour where world is trying hard to build different algorithms which can augment our capabilities in this area. But videos captured in the low light conditions or in dark suffer from poor visibility and due to that extracting features from these videos becomes very hard. Most of the feature extraction methods are designed to work on high quality input videos. There is some work done on the low-light enhancement methods for the images which we can be utilized in this paper. ARID Dataset, specially created for dark videos, is used for this research and the research explores different low light enhancement methods which are efficient in illumination of the each frame of videos and eventually help in feature extraction from videos. Furthermore, using the features extracted from video, this research proposes to build deep neural network on the enhanced videos which will eventually help in recognizing the human actions in the videos. Additionally, this research benchmarks some of the state-of-the-art methods for low light illumination and human action recognitions and proposes the method which is best in accuracy and low in computational time. Additionally, here plan is to find out whether the best state-of-the-art method for low light enhancement is really required for human action recognition because here focus is primarily on the human action recognition and not on the visual quality of the videos.

1. Background

In today's world, machines are helping us in various ways by identifying various situations and problems around us. Human action recognition is one of the areas where these machine learning algorithms can help us and augment our capabilities around it. We need the capability to identify human actions not only in good lighting conditions but also in dark or low lighting conditions.

Action recognition on normal videos with good visibility has done a lot of progress in recent times and many machine learning algorithms are able to do it efficiently but action recognition in the low light situations is still pretty challenging and a hard nut to crack because the pictures or videos captured during the night time or low light situations suffer from poor visibility and contains a lot of noise and sometimes these videos are even difficult for human eyes to recognize accurately.

Unfortunately, there is not enough work done on this in machine learning space but we see lot of work going on at hardware side where cameras are getting equipped with night vision capabilities but as we all know and experience that night vision sensors degrade the quality of the images or videos by introducing lot of noise and it gets more difficult for present day feature extraction algorithms to deal with these noises. Another drawback with these sensors is that they make the surveillance cameras more expensive and it's difficult for everyone to afford and adopt these technologies. So there is a need of software solutions which are capable of recognizing the actions in videos which are shot in dark without the help of any night vision capabilities.

There are lot of research going on for

- Night surveillance at very sensitive areas around the cities for crime investigations
- Surveillance at the country borders to stop illegal infiltration
- Self driving cars at night time
- Outdoors home surveillance during night hours or in low light situations
- Studies for wild animals behaviour during night time without disturbing them.

These all issues are very important which world would like to address.

According to an article (TheSleepJudge Editorial Team, 2020) about crimes in United States, more than 50% of the major crimes happen during night time. And most of the crimes are never reported which happen in the dark because there are either no proofs or if it's get reported,

video is not clear to identify the criminal. So it's a major issue for which we need the model which can enhance the dark videos and identify the human action.

One of the articles (Guglielmi, 2018) tells about that how wild mammal animals turning nocturnal because of the human activities in the forests during day time. Also humans use cameras with flash lights to study the wild animals' behaviours at night which actually disturbs then a lot during night time as well. We need to learn to see in the dark without the use of flash lights.

2. Problem Statement / Related Works

In last decade, many low light enhancement methods were proposed with different capabilities and different methods to enhance the dark images. Here, only few of the most popular methods are highlighted which work the best with ARID dataset and eventually help in identifying the human action from the dataset.

Some methods are based on histogram equalization, some are based on Gamma Intensity Correction where they try to amplify the low light in the images. Some methods are deep learning based where we need to have images in pairs of dark and bright images. Some methods are based on Retinex Theory where they suggest splitting the image in two parts, Reflectance and illumination where sometimes reflectance is considered as the enhanced image.

Recently, (Xu et al., 2021) proposed a new dataset (**ARID**) which contains 3784 video clips with 11 categories where each video is of around 1.2 seconds with 36 frames and all these videos are low contrast and low brightness videos. Together with dataset, a comparative study is done on the five different illumination methods which are based on Histogram Equalization, Gamma Intensity correction and Retinex Theory followed by human action classification on the videos.

According to this study, authors advocate Gamma Intensity correction (GIC) which gives the highest accuracy (78.03%) among all the methods. Authors have also done comparison using 3D-ResNext-101(Hara et al., 2018) classification model with original dark videos and enhanced video with various methods and concludes that GIC method gives the best improvements (3.30%) in the accuracy of the model and KinD(Zhang et al., 2019) enhancement method actually lowers the accuracy of classification model by (5.11%) in ARID Dataset.

One interesting thing to note here is, GIC doesn't necessarily do the best job in terms of image enhancement when we look at the image from the human lenses but its giving the highest accuracy (according to this paper) in action recognition from ML classification model point of view.

KinD (Zhang et al., 2019) (Kindling the Darkness), is one of the major work which was done recently for image enhancement. It's a deep learning based method which is inspired by Retinex Theory. This paper mostly focuses on the poor visibility and different types of degradation like noise and colour distortion. This research is done on LOL Dataset which

contains 500 low/normal light image pairs. 450 out of 500 images are used for training and 50 images are used for test.

KinD proposes layer decomposition and reflectance restoration neural networks, where restoration network adopts 5-layer UNet. It also proposes the Illumination Adjustment Net which can help in generating the ground-truth light level for images. In this research, batch size as 10 for layer decomposition net and batch size of 4 for reflectance and illumination adjustment net and stochastic gradient descent as optimization technique are used while building the models.

A comparative study is done on the 10 different image enhancement methods with the method proposed in the paper. There are some visual comparison and some metrics (i.e. PSNR, SSIM) comparison between images (enhanced with different methods) are present in the paper and based on that, it concludes that the proposed method gives clear advantage over all the state-of-the-art methods selected in the paper and it outperforms all the competitors using LIME (Guo et al., 2017) and NPE (Wang et al., 2013) datasets as well. One thing to note here is, by looking at the visual comparison, it can be clearly seen that the images, enhanced with KinD**,** do not suffer from any noise or any sort of colour distortion.

LIME (Guo et al., 2017) (Low-Light Image Enhancement via Illumination Map Estimation) is another very successful illumination technique and beats most of the state of art techniques of its time. This research warns that just recalling the visibility of the dark regions can create another problem of light saturation at the relatively bright regions so there will be a loss of details in the image.

Lime method is also based on Retinex theory but it mostly focuses on the illumination part. This method builds an illumination map and refines it by finding the maximum intensity of each pixel in R G B channels. For the refinement of illumination map, Augmented Lagrangian Multiplier (ALM) based algorithm is used which is also quite efficient algorithm and reduces the computational cost. Illumination map estimation and refinement actually considers the neighbouring pixels which help in local consistency of illumination which helps in getting uniform illumination across the image. Gamma correction and denoising and recomposition are done to get the best results for illumination map.

HDR dataset is used for this study and some low light images were chosen from the dataset and their LOE numbers are compared with all the other competitors. A comparative study is done on the 9 different image enhancement techniques with the proposed method and based on

the visual comparison and LOE numbers, it can be clearly seen that the proposed method gives the clear advantage over all those state-of-the-art methods selected in the paper.

Naturalness Preserved Enhancement Algorithm (Wang et al., 2013), is also based on Retinex Theory but it essentially focuses on preservation of naturalness of the image which is difficult to achieve in non-uniformly illuminated images.

This paper proposes a new algorithm for non-uniform illumination images where it also proposes lightness-order-error measure for naturalness and bright-pass filter which helps in decomposing the image into reflectance and illumination and bi-log transformation is done to consider the balance between naturalness and details. This paper tries to improve the local variation in the image and also doesn't harm the global trend of the intensity at the same time. So the paper focuses on the Reflectance extraction and relative order illumination compression.

This paper introduces its own dataset which contains more than 150 images. A comparative study has been done on these images with 6 state-of-the-art methods together with the proposed algorithm and based on the visual comparison, quantitative measurement results of discrete entropy and LOE numbers, it can be clearly seen that the proposed method gives the clear advantage over all those state of the art methods selected in the paper. According to the paper, their solution doesn't scale well with video files because it introduces slight flickering in the video files and authors plan to do further research in this area.

EnlightenGAN: Deep Light Enhancement Without Paired Supervision (Jiang et al., 2021), is another successful method which proposes a unsupervised dubbed EnlightenGAN (one path GAN), that can be trained and evaluated without normal/low-light image pairs unlike most of the deep learning based methods. Instead of ground truth data, information is extracted from input image and used for unpaired training.

In this paper, proposed method uses an attention U-Net neural network as the generator and uses global and local discriminator to maintain the texture and details of the image. This paper advocates for self-regularized attention map instead of supervised learning. Also in the experiment, it uses data from multiple dataset where around 900 low light and 1K+ normal light images are collected from various datasets.

EnlightenGAN is trained by 100 epochs and with learning rate of 1e-4 with Adam Optimizer and then another 100 epochs where learning rate was linearly decayed to zero. A comparative study is also done with 6 state-of-the-art methods together with EnlightenGAN and

based on the visual comparison and based on NIQE scores, it can be clearly seen that the proposed method gives the clear advantage over all the methods selected in the paper.

Human action recognition based on recognition of linear patterns in action bank features using convolutional neural networks (Ijjina and Mohan, 2014), proposes a deep CNN architecture for recognizing human actions in videos using action bank features. Action bank is nothing but a predefined set of videos converted to linear representation and saved as action bank features. This research proposes a convolutional network with linear mask which can capture the localized patterns for each action.

Generally, in case of action recognition in a video requires a 3D CNN to learn the spatiotemporal features from videos but this paper advocates using 2D CNN with action bank which is based on the concept of speech recognition using the spectrogram of audio data. They have designed a CNN which exploits this similarity in action bank features and this will drastically reduce the computational time for action recognition.

UCF50 dataset containing 205 videos is used in this research, these input videos first get processed by the feature extraction module to extract the action bank, and then these action bank features are given as input to pattern recognition module for training. The CNN utilizes the similarity patterns to assign an action label to the videos. By looking at the results with this approach looks quite promising where the proposed method is able to achieve 93-94% accuracy in action recognition.

Human Action Invarianceness for Human Action Recognition (Sjarif and Shamsuddin, 2016), proposes to use human action shape or silhouette uniqueness to recognizing the human actions. Human action features can be extracted by using integration moment invariant. Action features are actually based on how silhouette moves in video frames. In human action invariance, the paper proposes three processes like extracting global features, similarity measurement between features and intra and interclass analysis.

Authors have used IXMAS dataset which contains 13 different actions performed by 10 people. The experiment is done with different video frames i.e. 30, 120 and 300. This research uses various techniques to achieve the higher accuracy like wavelet, PCA, Normalization, pre and post discretization. A comparative analysis on other methods is also done which other methods are not able to give the good accuracy with this dataset. But the method proposed by this paper is able to perform very well and is able to predict the human action with high accuracy up to 98-

99%. This paper has given promising results and could be one of methods in my research in identifying the human action recognition in dark.

Learning Spatiotemporal Features with 3D Convolutional Networks (Tran et al., 2015), proposes deep 3D ConvNets which are trained on large dataset and suggest that these 3D ConvNets are better than 2D ConvNets for spatiotemporal feature extraction. This paper evaluates both the networks and concludes that 3D ConvNets gives much better accuracy (98%) compare to 2D ConvNets.

Learning to See Moving Objects in the Dark (Jiang and Zheng, 2019), method proposes a new optical system which can capture both bright and dark videos of the same scene, to have both training and ground truth dataset.

They discouraged the use of infrared sensors in the cameras because in forests it can disturb the animals and might trigger uncontrollable animal reactions. So they advocate enhancing the images which are captured by ordinary cameras with ND filter so the proposed camera can click both dark and bright images simultaneously. This paper introduces a new dataset which contains 179 pairs of videos consisting of 35800 extremely low-light images and their corresponding properly lighted images.

This paper proposes a 3D U-net based network for low-light enhancement. A comparative study is done between state-of-the-art methods and the proposed method which concludes that the proposed method is the best among all the methods and also according to this paper, proposed method is able to tackle the flickering issues in the enhanced videos.

Action Recognition in the Dark via Deep Representation Learning (Ulhaq, 2018), proposes action recognition from multiple video streams using deep multi-view representation learning. This paper talks about video captured in low lighting condition and multi-sensor scenario which is a novel method to fuse spatio-temporal deep correlations from multiple streams.

This paper proposes an approach to deal with very low quality dark night time videos. In the experiment, Night-Vision Video dataset (NV) is used. This dataset contains videos recorded by two different cameras Raytheon Thermal IR-2000B and Panasonic WV-CP470.

They were able to achieve average precision of 80% which is pretty impressive. But this method requires multiple videos captured via different sensors in cameras.

3. Research Questions

The below research questions are formulated based on Literature Review done in the field of human action recognition in dark:

- Are there any conclusions to use a particular low light enhancement method which actually helps in human action recognition?
- Is the best low light enhancement method really required to get high accuracy in human recognition or a simple low light enhancement method is good enough for the model to perform well in identifying the human actions?

  Here we are going to focus on the human recognition in dark, not on the visual quality of the videos.
- What will be the Human action recognition model's accuracy for different low light enhancement methods?
- Can we use different algorithms for video enhancement based on business needs? User might be interested to see the actual enhanced video after human action recognition is done by the model.

4. Aim and Objectives

The aim of this research is to do the comparative study on the present capabilities of low light enhancement methods and build a model and recognize the human actions in ARID dataset. The goal of this study is to identify the best low light enhancement method for videos which are shot in dark and which works the best and gives the highest accuracy with an action recognition deep learning model.

The research objectives are formulated based on the aim of this study, which are as follows:

- To analyze the most popular low light enhancement methods and identify the best method among all of them which works the best for human action recognition.
- To propose a deep learning model which works the best in identifying the human actions with low light enhanced videos.
- To perform comparative study on 2D CNN and 3D CNN methods for human action recognition.

5. Significance of the Study

The significance of this study is to explore various low light enhancement methods and find the best method which is efficient in recognizing the human action in the dark videos and will augment human capabilities during night time.

There are many areas where this research will help like night surveillance. To make our city and country safe, we need the capabilities to do surveillance at very crime sensitive areas. This model will help in recognizing the human actions during the night hours; we don't need to deploy a human at every place for surveillance. Most of the times it is very difficult for the humans as well to see in the dark without night vision capabilities. This action recognition method will help our police to protect our neighbourhood and people.

The second use case is Self Driving Cars; there are lot of researches going on for autonomous cars. As we know, self driving cars is still a big challenge and auto maker companies still facing a lot of challenges to make a perfect autonomous car where driving in the dark will be another hard nut to crack. I hope this study will be able to help this area to some extent because we need a model which can recognize things in dark.

This model, with the comparatively analysis, will be able to help data scientists to find different methods to identify human actions in various scenarios. This study will be extended to give suggestions about different low-light enhancement methods which are best suited for different dark environment scenarios.

6. Scope of the Study

Due to lack of time frame, the scope of the study will be limited as below:

- The dataset is taken from ARID (Xu et al., 2021) dataset which is publicly available and data validation is not the part of this study. The scope of study is to analyze only 11 human actions mentioned in the data but this can be extended to any number of human actions in the future.
- This study contributes to the comparative study on different low light enhancement techniques, comparative study on human action recognition techniques and suggests the best combination of both which help in recognizing the human action in the videos which are shot in the dark.
- In case of low-light illumination, this study will limit to use some of the state-of-the-art techniques like KinD (Zhang et al., 2019), LIME (Guo et al., 2017), EnlightenGAN (Jiang et al., 2021) and Gamma Intensity Correction and plans to do the comparative analysis by comparing the different metrics.
- In case of human action recognition, this study will limit to use some of methods like action bank feature extraction (Ijjina and Mohan, 2014), utilize a rolling prediction average and utilize the 2D, 3D CNN, RNN methods for model building.

7. Research Methodology

Today, as our population is increasing, it is getting difficult to do the night surveillance manually by humans, we need some kind of solutions to solve the problems we face during night time; some scenarios are like night surveillance, illegal infiltrations at country borders, self driving cars during night time and analyze animal behaviours in the wild during night. Humans need to augment their capabilities during the night. A lot of work in done on Human action recognition with the properly lighted videos and images but not enough work is done on action recognition in dark videos. However some of the state-of-the-art methods have done awesome job in enhancing the low lighted images and we are in the right direction.

The motive of this research is to analyze human actions in videos which are shot in the dark and with normal cameras. Human action recognition directly on the dark videos will not be efficient because it will be difficult to extract features out of the videos. So this paper suggests the research in two parts. Firstly, find the best method to pre-process and do the low light enhancement in the dark videos. Secondly, do the human action recognition on enhanced videos. There are some of the methods which are available for low light enhancement for videos and images. A comparative study needs to done on various methods and propose the best method which gives the highest accuracy and lowest computation time. Some methods enhance the images visually very well but a question to ask is "do we really need these kind of refinements for action recognition in the dark?"

7.1 Dataset

This research uses ARID dataset; this dataset contains more than 3700 video clips with 11 different kinds of human actions where each class contains around 110 clips For example: Drinking, Jumping, picking, pouring, running, sitting, standing, turning, walking and waiving. Each video clip has fixed frame rate of 30 frames/sec and 1.2 seconds long. (Xu et al., 2021) have created this dataset specifically for analyzing dark videos. These dark videos are shot from the normal cameras without using any kind of night vision sensor. Some of videos are so dark that it is very difficult for even human eye to identify the actions in the videos.

The purpose of this dataset is to provide the dark videos which were really shot in the dark because we see most of the researches are done on synthetic dark videos or the dark videos

which were created using some camera filters. Models built on synthetic dataset don't perform very well with the actual dark videos.

7.2 Data Pre-Processing

Before starting to build the model for action recognition, we need to pre-process the data. Here this research plans to do the low light enhancement in the videos as part of the pre-processing. For low light enhancement, some of methods will be used which make sense to use with ARID [1] dataset. Some of methods which will work well this dataset are: KinD (Zhang et al., 2019), LIME (Guo et al., 2017), EnlightenGAN (Jiang et al., 2021), Histogram Equalization (HE), Gamma Intensity Correction (GIC). EnlightenGAN seems to be the most impressive and promising methods as of now among all the low-light image enhancement methods.

These methods are chosen because they directly enhance the low lighted images without the need of paired bright light images for training the model. These all the enhancement methods were invented for images, so these methods need to be applied on each of the video frames. During enhancement of the videos, a comparative study is planned where their computational time will be on focus.

7.3 Model Building

After the dark videos enhancement, it is required to use multiple CNN based neural networks for feature extraction and find the best neural network which gives the highest accuracy and the lowest computational time on ARID dataset.

For model building, Keras framework is required for designing the 2D or 3D CNN models where Tensor Flow will act as a backend to Keras. There are multiple ways and methods to do the human action recognition, but this research plans to use three different methods for video classification

- Using feature extraction module to extract the action bank features (Ijjina and Mohan, 2014). We will create the action bank for all the 11 action categories in the dataset and train the model on the action bank and then build the 2D CNN model to do the video classification.

This method is chosen because it will reduce the computational time drastically as we are using 2D CNN for video analysis instead of 3D CNN and this method is giving more than 90% accuracy in action recognition.

- Another way is to use 2D CNN and utilize a rolling prediction average (Rosebrock, 2019) where author suggested using a 2D CNN model and classifying each frame individually and keeping the list of predictions and taking the average of last K predictions and predicting the label with largest probability.

  Here this method is chosen because it's a 2D CNN model which will help in reducing the computational time and according to the results shown by the author its working nicely on the action recognition on the videos.

- Using neural network architectures like Long short-term memory (LSTMs) and Recurrent Neural Networks (RNN), because videos are nothing but sequence of images (frames). RNNs are well suited for sequential data classification.

  This method is chosen because a comparative study is required between 3D CNN model and 2D CNN model approaches to find the best method for ARID dataset.

7.4 Model Training and Evaluation

Dataset will be split into two parts (70% and 30%), 70% of the videos will be used for training the dataset and rest 30% will be used for validation of model.

While model training, below hyper parameters will be considered

Parameter	**Value**
No. of Epochs	100
Batch Size	30
Optimization Technique	Adam
Learning Rate	Default for Adam
ReduceLROnPlateau	Patience=5, factor=0.1

Table 1 Hyper Parameters

As shown in Table 1 and Figure 1, model will be trained with 100 Epochs with batch size of 30. These values are chosen based on the resource availability during this research. Other than that,

ADAM will be used as the optimization technique while finding the global minima and reduce the learning rate at the plateau with the factor of 0.1 and patience level of 5.

Below metrics will be considered which will tell us how the model is performing during model training and evaluation:

Metrics	**Description**
categorical_accuracy	Accuracy of the model on the training data
val_categorical_accuracy	Accuracy of the model on the validation data
Precision	Proportion of positive identifications was actually correct
Recall	Proportion of actual positives was identified correctly
F1-Score	This a harmonic mean of Precision and Recall

Table 2 Model Evaluation Metrics

We have to watch out for categorical accuracy and validation categorical accuracy especially so that we have balanced performance on train and test set. Both the accuracies should be in the same range and more than 80% for a good model. If categorical accuracy is too high and validation accuracy is too low then it will considered as over fitting which we have to avoid.
Only accuracy may not determine the performance of the model but other metrics like Precision and Recall are required to be analyzed which will help in determining the performance of the model.

Methodology described in Figure 1, will be repeated for every low light enhancement technique and for every action recognition algorithm. All the models will be saved to Model Repository. All the models will be evaluated and compare against various metrics and propose the best low light and human action recognition algorithm for ARID dataset.

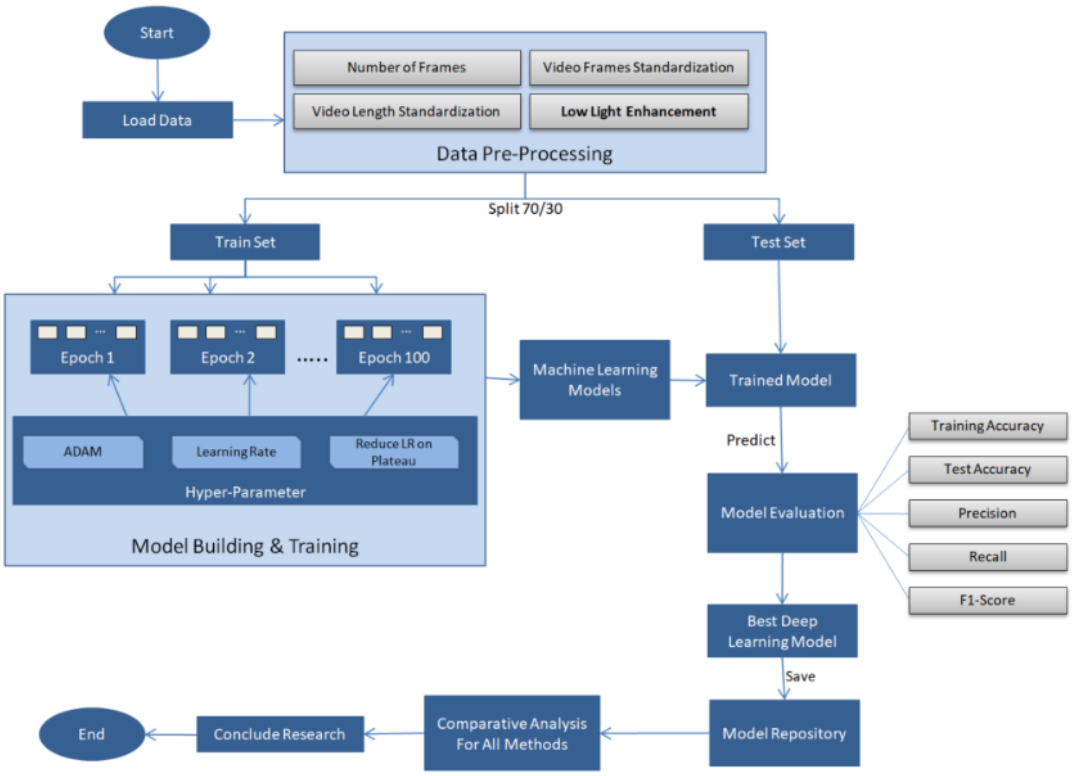


Figure 1 Research Methodology

8. Required Resources

The research will need below hardware and software resources throughout the implementation

8.1 Software Requirements

The Predictive modelling will be implemented using python.

- Package Manager: Anaconda Navigator 1.9.12
- Presentation Layer: Jupyter lap 0.35.4
- Kaggle Environment to have access to GPUs and TPUs
- Language: Python 3.6.X
- Python Libraries:
    - Pandas and NumPy for data processing
    - Matplotlib and Seaborn for data visualization
    - Keras framework and Tensor Flow as its backend for CNN and RNN model building and for model evaluation.

8.2 Hardware Requirement

A laptop with below configuration will be used:

- Operating System. Windows 10: 64-bit
- Processor: Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz 2.90 GHz
- Memory: 16 GB

At Kaggle Environment below configuration will be used:

- Processor & Accelerator: NVIDIA TESLA P100 GPUs or TPU v3-8
- Duration: At least 20-30 Hours / Week
- Memory: 8GB per kernel

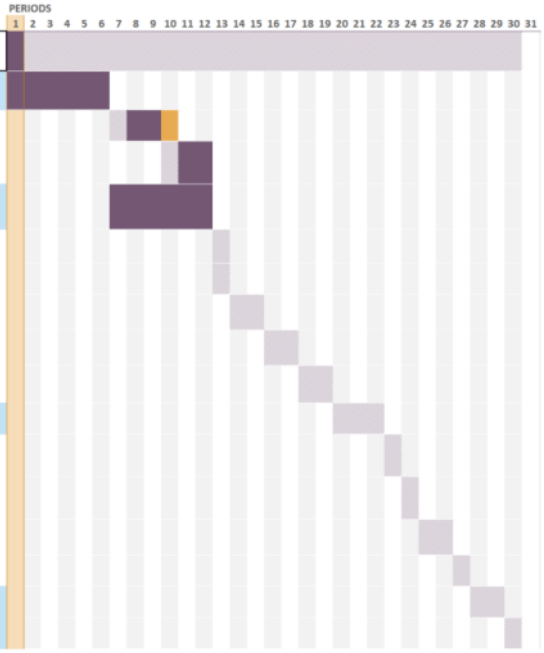
9. Research Plan

Project Planer - Action Recognition in Dark

Select a period to highlight at right. A legend describing the charting follows.

Period Highlight: 1 | Plan Duration | Actual Start | % Complete | Actual (beyond plan) | % Complete (

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Research Project: Action Recognition in Dark	1	30	1		30%
Milestone 1: Topic Submission with dataset	1	6	1	6	100%
Literature Review	7	3	8	3	100%
Document the Content for Research Proposal	10	3	11	2	100%
Milestone 2: Research Proposal Submission	7	6	7	6	100%
Get the data and prepare with hardware and software	13	0.5			0%
EDA and Data Preprocessing	13	1			0%
Low-Light Enhancement to All The Videos	14	2			0%
Build Deep Learning Models for Action Recognition	16	2.5			0%
Evaluate the prediction by deep learning models	18	2			0%
Milestone 3: Interim Report	20	3			0%
Comparitive Study between diffirent low light enhancement techniques	23	1			0%
Explaination for comparitive study between state-of-the-art methods	24	1			0%
Explaination for Best Deep Learning model for Action Recognition with	25	2			0%
Final Conclusion	27	1			0%
Milestone 4: Final Report	28	2			0%
Milestone 5: Video Presentation	30	1			0%



References

Guglielmi, G., (2018) *Mammals turn to night life to avoid people*. [online] Nature (Nature). Available at: https://www.nature.com/articles/d41586-018-05430-4 [Accessed 3 Aug. 2021].

Guo, X., Li, Y. and Ling, H., (2017) LIME : Low-Light Image Enhancement via. 262, pp.982–993.

Hara, K., Kataoka, H. and Satoh, Y., (2018) Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet? *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.6546–6555.

Ijjina, E.P. and Mohan, C.K., (2014) Human action recognition based on recognition of linear patterns in action bank features using convolutional neural networks. *Proceedings - 2014 13th International Conference on Machine Learning and Applications, ICMLA 2014*, pp.178–182.

Jiang, H. and Zheng, Y., (2019) Learning to see moving objects in the dark. *Proceedings of the IEEE International Conference on Computer Vision*, 2019-OctobIccv, pp.7323–7332.

Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P. and Wang, Z., (2021) EnlightenGAN : Deep Light Enhancement. 30, pp.2340–2349.

Rosebrock, A., (2019) *Video classification with Keras and Deep Learning*. [online] pyimagesearch. Available at: https://www.pyimagesearch.com/2019/07/15/video-classification-with-keras-and-deep-learning/ [Accessed 3 Aug. 2021].

Sjarif, N.N.A. and Shamsuddin, S.M., (2016) Human action invarianceness for human action recognition. *SKIMA 2015 - 9th International Conference on Software, Knowledge, Information Management and Applications*.

TheSleepJudge Editorial Team, (2020) *Crimes that Happen While You Sleep*. [online] Available at: https://www.thesleepjudge.com/crimes-that-happen-while-you-sleep [Accessed 3 Aug. 2021].

Tran, D., Bourdev, L., Fergus, R., Torresani, L. and Paluri, M., (2015) Learning spatiotemporal features with 3D convolutional networks. *Proceedings of the IEEE International Conference on Computer Vision*, 2015 Inter, pp.4489–4497.

Ulhaq, A., (2018) Action Recognition in the Dark via Deep Representation Learning. *IEEE 3rd International Conference on Image Processing, Applications and Systems, IPAS 2018*, pp.131–136.

Wang, S., Zheng, J., Hu, H.M. and Li, B., (2013) Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 229, pp.3538–

3548.

Xu, Y., Yang, J., Cao, H., Mao, K., Yin, J. and See, S., (2021) ARID: A New Dataset for Recognizing Action in the Dark. *Communications in Computer and Information Science*, 1370, pp.70–84.

Zhang, Y., Zhang, J. and Guo, X., (2019) Kindling the darkness: A practical low-light image enhancer. *MM 2019 - Proceedings of the 27th ACM International Conference on Multimedia*, pp.1632–1640.

Research Proposal-2.docx

ORIGINALITY REPORT

9%	3%	5%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Liverpool John Moores University Student Paper	4%
2	Anwaar Ulhaq. "Action Recognition in the Dark via Deep Representation Learning", 2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS), 2018 Publication	1%
3	Earnest Paul Ijjina, C. Krishna Mohan. "Human Action Recognition Based on Recognition of Linear Patterns in Action Bank Features Using Convolutional Neural Networks", 2014 13th International Conference on Machine Learning and Applications, 2014 Publication	1%
4	Haiyang Jiang, Yinqiang Zheng. "Learning to See Moving Objects in the Dark", 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019 Publication	1%

5	academic.odysci.com Internet Source	1 %
6	"Advances in Informatics and Computing in Civil and Construction Engineering", Springer Science and Business Media LLC, 2019 Publication	1 %
7	"Computer Vision Systems", Springer Science and Business Media LLC, 2019 Publication	1 %
8	Xiaojie Guo, Yu Li, Haibin Ling. "LIME: Low-Light Image Enhancement via Illumination Map Estimation", IEEE Transactions on Image Processing, 2017 Publication	1 %
9	link.springer.com Internet Source	1 %

Exclude quotes	Off	Exclude matches	< 1%
Exclude bibliography	On		