iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** III **Month of publication:** March 2019

DOI: http://doi.org/10.22214/ijraset.2019.3398

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue III, Mar 2019- Available at www.ijraset.com

Deep Learning Model for Motion Video Processing

Yugesh. E[1], Santhosh. M. M[2], Gopikrishnan. M[3]
[1, 3]*Department of Computer Science and Engineering, Prathyusha Engineering College, Thiruvallur – 602 025*

***Abstract:** In the film industry for motion picture processing, it requires different rate (fps) of motion capture. Utilizing single sequence compromises on quality. Different frames are required to be inputted to video editing tools as per motioning requirements. The goal of the paper is to construct a model to utilise the captured frame sequence and generate slow motioning using clustering and Convolutional Neural Networks.*
***Keywords:** Deep Learning, Convolutional Neural Networks, Video Processing, Segmentation, Motioning*

I. INTRODUCTION

In the film industry for motion picture processing, it requires different rate (fps) of motion capture. Utilizing single sequence compromises on quality. Different frames are required to be inputted to video editing tools as per motioning requirements. The goal of the paper is to construct a model to utilise the captured frame sequence and generate slow motioning using clustering and Convolutional Neural Networks.

II. PROBLEM IDENTIFICATION

The major problem in the filming industry is to generate slow motioning with existing sequence of frames. With slow motion capture it is possible to generate fast motion by removing the intermediate frames. It is possible to generate slow motioning from fast motion capture by inserting intermediate frames, but it experiences jitters and jerks during the final rendering. This problem is seen in the video editing tools like Adobe Premier Pro, Avid Media Composer.

III.PROBLEM SOLUTION

Despite there is not a complete solution for the problem, the model can be trained so that the jitters and jerks can be minimized to certain extend. The deep learning model uses clustering mechanism to group the similar set of frames. The duplicate frames are identified using Deep Neural Networks for unstructured data. They are further grouped together and then they are ordered .The sequence is then fed into the tool to render the final output. The final output is the slow motion video. The final output will have less distortion when compared with regular video editing tools. The accuracy of the model depends upon the training of the model.

IV.DISADVANTAGES IN THE EXISTING SYSTEM

The existing video editing tools will experience jitters and jerks during the final rendering of the output. The distortion experienced in the final rendering of the output is a major problem experienced in video motioning

V. CLUSTERING

The similar set of frames that are recognized are grouped together .Clustering is done on the similar set of frames or identical frames that was detected using the Deep Neural Networks from the pixel data. Clustering is basically grouping the images based upon the similarity in the patterns. This helps in discovering the patterns in the underlying data. It is based upon the distance metrics and similarity. It calculates the data points at each position and measures the similarity. The primary aim is to partition, segregate various different data and group the similar ones. Few partition algorithms are very dynamic such that they easily transit from one cluster to another. They gain knowledge through the prototypes and through the size and shape of the clusters.

VI.ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks are the model adapted from the biological Neural System. It consists of a large number of interconnections and nodes. Each of the nodes is addressed as neurons capable of computation. The computation is performed from the inputs and the nodes are trained by collective way of learning in order to optimize the final output .The input is taken in the form of a multi dimensional vector and the data is travelled across the hidden layers.The hidden layers improvise the decisions based upon the previous layers. These hidden layers are stacked upon each other.This is how the Deep Learning process takes place.The Artificial Neural Network has a high computational complexity, so it finds hard to compute the image data and also to store them.

VII. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks is used for image processing and recognition. Each neuron in the network takes the pixel data as input and stores it. The particular neuron is compared with the neuron of another similar image for recognizing the similarity based upon the color level. This process is done to identify the similar images. The images can be cropped and rescaled, the video is a collection of pictures where each pictures are being processed. It is based on the connectivity pattern i.e., it may be Early Fusion or a Late Fusion .The early fusion takes the entire information and combines them at pixel level. This allows the pixel data to be detected easily. The late fusion model processes two separate frames and shares the parameters. Then the streams are merged into a single stream. The Convolutional layer determines the output based upon the connectivity towards the inner regions. The calculation is based upon the weights and connectivity. The scalar product is done between the weights and the connectivity. Optimization of a model that was built might take some time and it happens gradually. The model is optimized by repeated training. The individual layers are explored and the weights are calculated. With the deep Convolutional Networks (ConvNets) architectural model [6] image recognition is done [4, 5].

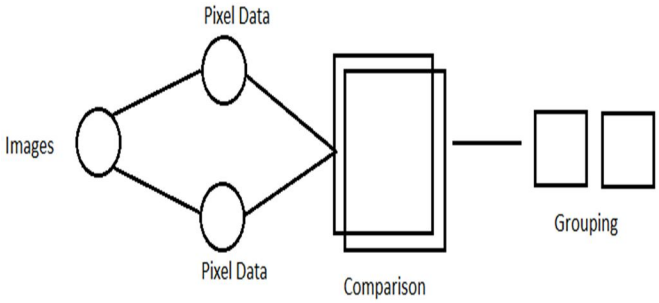


Fig. 1 Image Processing

VIII. SUPERVISED LEARNING

The supervised learning is based upon the pre labelled inputs. In a training example, there exist some input values and some approximately predicted output values are mapped to the input values. This model reduces the classification error. This basically aims at finding the correct value output for the given input based upon the training set. The accurate results are provided as input based upon which the learning is carried out. This is most commonly used approach for Machine Learning. The output is in the form if vector scores and based upon which the further processing is done. The gradient vectors are calculated so as to level the weight vectors. The machine learns and adjusts the adjustable inner parameters by itself.

IX. UNSUPERVISED LEARNING

The unsupervised learning does not require any labelling. The unsupervised learning uses unlabelled data as per mentioned in [2]. The success rate is usually being determined by the minimization of the cost function. The most important in this learning is to note that the image recognition, pattern recognition highly rely on unsupervised algorithms. The initial usage of unsupervised criterion to pre-train each layer leads to a better solution as mentioned by the journal paper [1]. The accurate results are not provided at the initial stages. It computes and update at each stages [3].

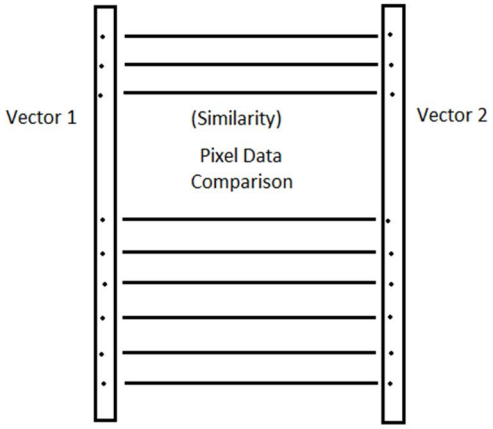


Fig. 2 Pixel Data Comparison

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue III, Mar 2019- Available at www.ijraset.com

X. DEEP LEARNING

Deep Learning has multiple processing layers. With the help of those layers the patterns are learned and recognized by processing multiple levels or layers of abstraction. Deep Learning is essential to handle the data in raw form and especially if the data has no structure to it .If the data is unstructured then deep learning algorithms are the only algorithms that can process such data. New algorithms and architectures for Deep Learning are under progress. This will enrich and enhance the existing system by integrating with new features. The Deep Learning methodologies are based upon representational learning, which consists of multiple layers of abstractions. It is used to handle various non linear modules which transform the representations at each layer.

XI. SEQUENCE DIAGRAM

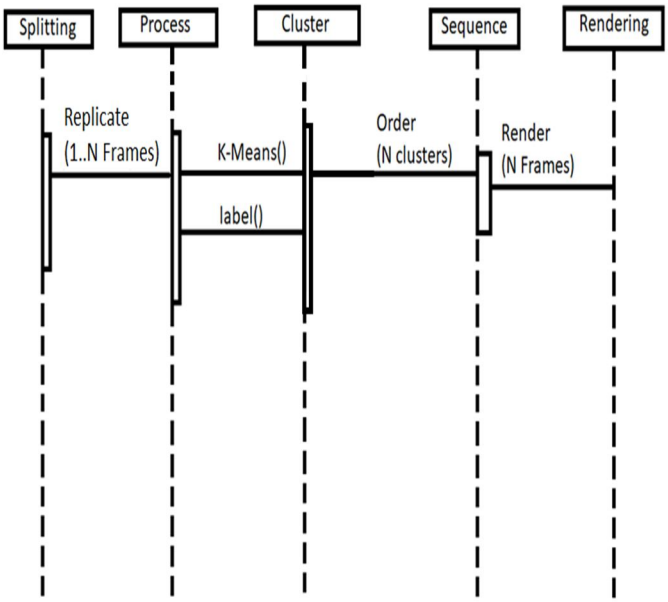


Fig. 3 Process Flow

XII. EXPERIMENTAL IMPLEMENTATION

The following stages are carried out.

A. Splitting

The extraction of frames from the fast motion picture is splitting. Splitting is the process of dividing the video into multiple parts.

B. Replication

The frames which were extracted are then made multiple copies.This process is a duplication process where the exact copy of a frame is made.

C. Processing

The copies which are identical are grouped together. The identical copies are recognized in this process.

Fig. 4 Similar Image Recognition

D. Sequencing

The frames are arranged in a sequence for final rendering. This is ordered the frames in a particular order so that it is ready to be rendered.

E. Rendering

The derived sequence, the slow motion video can be rendered by combining the images. The final output is the slow motion video.

Fig. 5 Original Video

Fig. 6 Slow Motion Video

XIII. PROGRAM OUTCOME

Deep Learning model for generating a slow motion video by extracting frames from a fast motion video and training them using Deep Neural Networks, so that the final outcome is distortion less. The intermediate frames will be inserted as per the motioning requirements. And the final slow motion video will be generated.

XIV. CONCLUSIONS

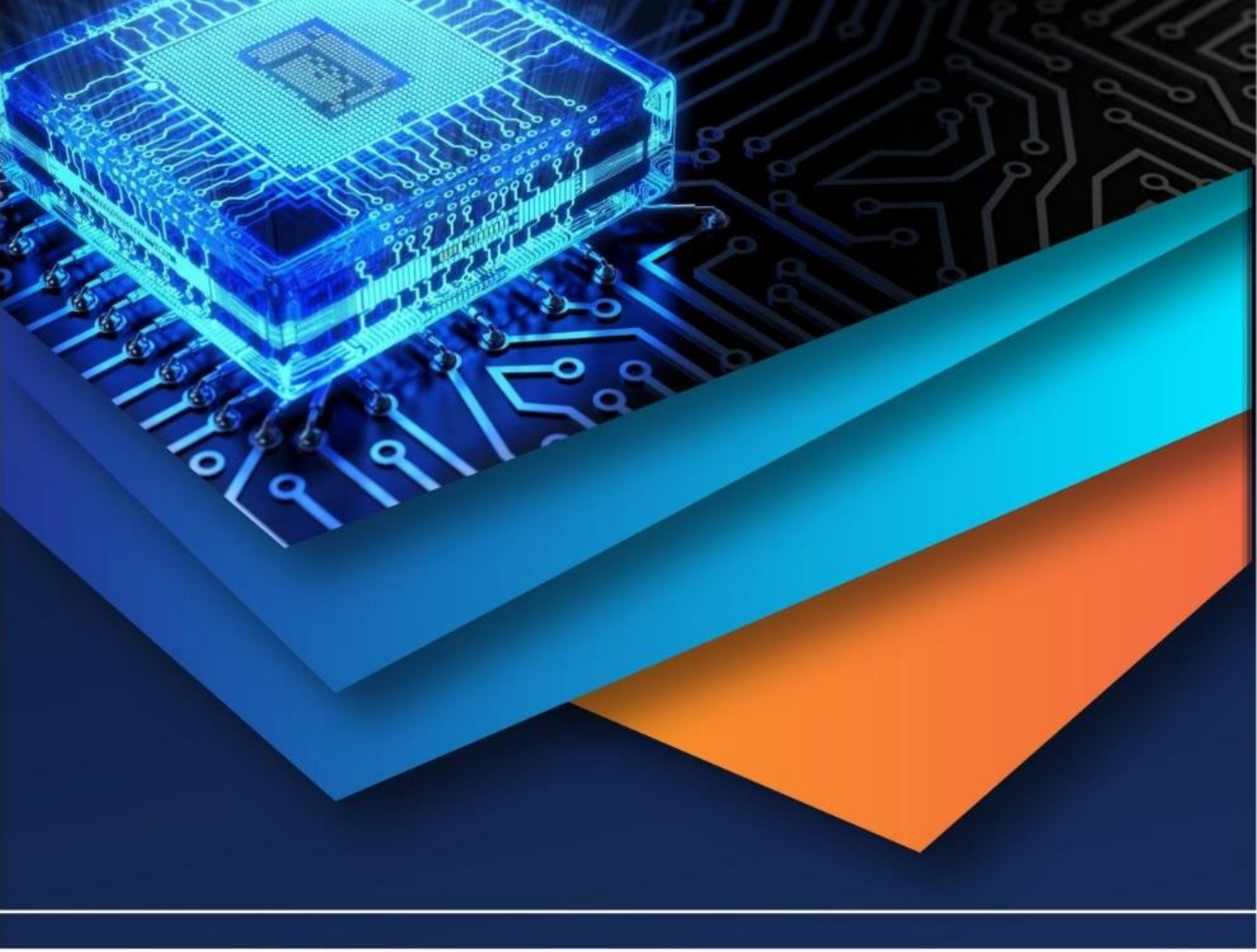
Deep Learning model for generating a slow motion video by extracting frames from a fast motion video and training them using Deep Neural Networks, so that the final outcome is distortion less. The intermediate frames will be inserted as per the motioning requirements. And the final slow motion video will be generated.

XV. ACKNOWLEDGMENT

We would like to show our gratitude to the Institution for sharing their pearls of wisdom with us during the course of this paper. We thank M.GopiKrishnan for his comments on an earlier version of the manuscript, although any errors are our own and should not tarnish the reputations of these esteemed persons.

REFERENCES

[1] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, “Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion,” The Journal of Machine Learning Research, vol. 11, pp. 3371–3408, 2010.
[2] Farabet, C. et al. Large-scale FPGA-based convolutional networks. In Scaling up Machine Learning: Parallel and Distributed Approaches (eds Bekkerman, R., Bilenko, M. & Langford, J.) 399–419 (Cambridge Univ. Press, 2011).
[3] Raina, R., Madhavan, A. & Ng, A. Y. Large-scale deep unsupervised learning using graphics processors. In Proc. 26th Annual International Conference on Machine Learning 873–880 (2009).
[4] D. C. Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. In Proc. CVPR, pages 3642–3649, 2012.
[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. In NIPS, pages 1106–1114, 2012.
[6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY