

PROJECT SYNOPSIS

on

Audience Measurement Using Digital Video Watermarking

SUBMITTED BY

Name	SAP ID
1) Akshay Bhogan	60002160007
2) Yogesh Deshpande	60002160021
3) Viren Baria	60002160005

Under the guidance of

Dr. Sunil H. Karamchandani
Assistant Professor

DEPARTMENT OF
ELECTRONICS AND TELECOMMUNICATION ENGINEERING
Academic Year : 2019-2020

Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

Plot no. U-15, JVPD Scheme, Bhaktivedanta Swami Marg,
Vile Parle (W), Mumbai – 400 056

Shri Vile Parle Kelavani Mandal's
Dwarkadas J Sanghvi College of Engineering
Plot No. U – 15, JVPD Scheme, Bhaktivedanta Swami Marg,
Vile Parle (W), Mumbai – 400 056

Department of Electronics and Telecommunication Engineering

This is to certify that the Project Report Stage – I

“Audience Measurement Using Digital Video Watermarking”

Submitted by:

1. **Akshay Bhogan**
2. **Yogesh Deshpande**
3. **Viren Baria**

Students of **Electronics and Telecommunication Engineering** have successfully completed their **Project Stage – I** required for the fulfillment of **SEM VII** as per the norms prescribed by the **University of Mumbai** during the second half of the year 2019. The project synopsis has been assessed and found to be satisfactory.

Internal Guide

External Guide

Head of Department

Principal

Internal Examiner

External Examiner

Table of Contents

Sr No.	Contents	Page No.
1.	Introduction	4
2.	Theory	6
2-a	Working Principle	6
2-b	Block Diagram	13
3.	Extensive Literature Review	14
4.	Work to be carried out in semester 8	18
5.	Project Budget Sheet	19
6.	Applications	20
7.	References	21

1. INTRODUCTION

ABSTRACT: TV audience measurement involves estimating the number of viewers tuned into a TV show at any given time as well as their demographics. First introduced shortly after commercial television broadcasting began in the late 1940s, audience measurement allowed the business of television to flourish by offering networks a way to quantify the monetary value of TV audiences for advertisers, who pay for the estimated number of eyeballs watching during commercials. The first measurement techniques suffered from multiple limitations because reliable, large-scale data were costly to acquire. Yet despite these limitations, measurement standards remained largely unchanged for decades until devices such as cable boxes, video-on-demand boxes, and cell phones, as well as web apps, Internet browser clicks, web queries, and social media activity, resulted in an explosion of digitally available data. TV viewers now leave digital traces that can be used to track almost every aspect of their daily lives, allowing the potential for large-scale aggregation across data sources for individual users and groups and enabling the tracking of more people on more dimensions for more shows. Data are now more comprehensive, available in real time, and cheaper to acquire, enabling accurate and fine-grained TV audience measurement. In this project, we discuss the evolution of audience measurement and what the recent data explosion means for the TV industry and academic research.

BACKGROUND: Audience Measurement has been in use from 1940's. The ideal role of such measurement was to provide the television viewership data. Earlier the broadcasters and advertisers determined how many people are watching their channel, but now due to the evolution in the entertainment sector, channel providers need to determine the **3W** questions. **Who** are watching? **When** they are watching and **Why** they are watching? These data are broken down according to the target audience and regions.

Earlier in the past era, audience measurement began with the use of Diaries. It was used as one of the first method to record information. Data was written and collected to the level of customer opinion, cross referenced with their age, race and economic status. In 1942, IBOPE (Instituto Brasileiro de Opinião Pública e Estatística or Brazilian Institute of Public Opinion and Statistics) was established in Sao Paulo, Brazil. IBOPE provided research on media ratings, public opinion and other polls required by their clients. During 1960s and 1970s, as stated by Amanda D Lotz in "The Television will be Revolutionized", Nielson introduced Storage Instantaneous Audimeter, a device that sent daily viewing information to company computers using phone lines. But, due to ineffective performance and

inaccurate ratings, Nielsen introduced a more advanced version of his people's meter. The Local People Meter (LPM) provided more accurate ratings of particular local markets. The LPM system allows year-around period ratings rather than, quarterly 'sweeps' period. The introduction of digital terrestrial television (DTT) introduced digital and analog measurement in mixed television broadcast. Due to large number of newly emerging content, it resulted in technological convergence. Thus, new methodologies like Audio or Video matching and watermarking were required, for the measurement of television audiences.

In India, BARC (Broadcast Audience Research Council) which is a joint industry company founded by the stakeholder bodies that represent Broadcasters, Advertisers and Advertising & Media Agencies. BARC India sets up design, commission, supervise and owns an accurate, timely and reliable television audience measurement system. It provides information relating to the viewing patterns of customers. It allows advertisers to understand viewing patterns allowing informed decisions to be made on where to place advertisements to reach the correct target audience at the lowest cost. Our aim is to build a device that can measure the number of viewers for a particular channel and also the duration for which the channel is being viewed.

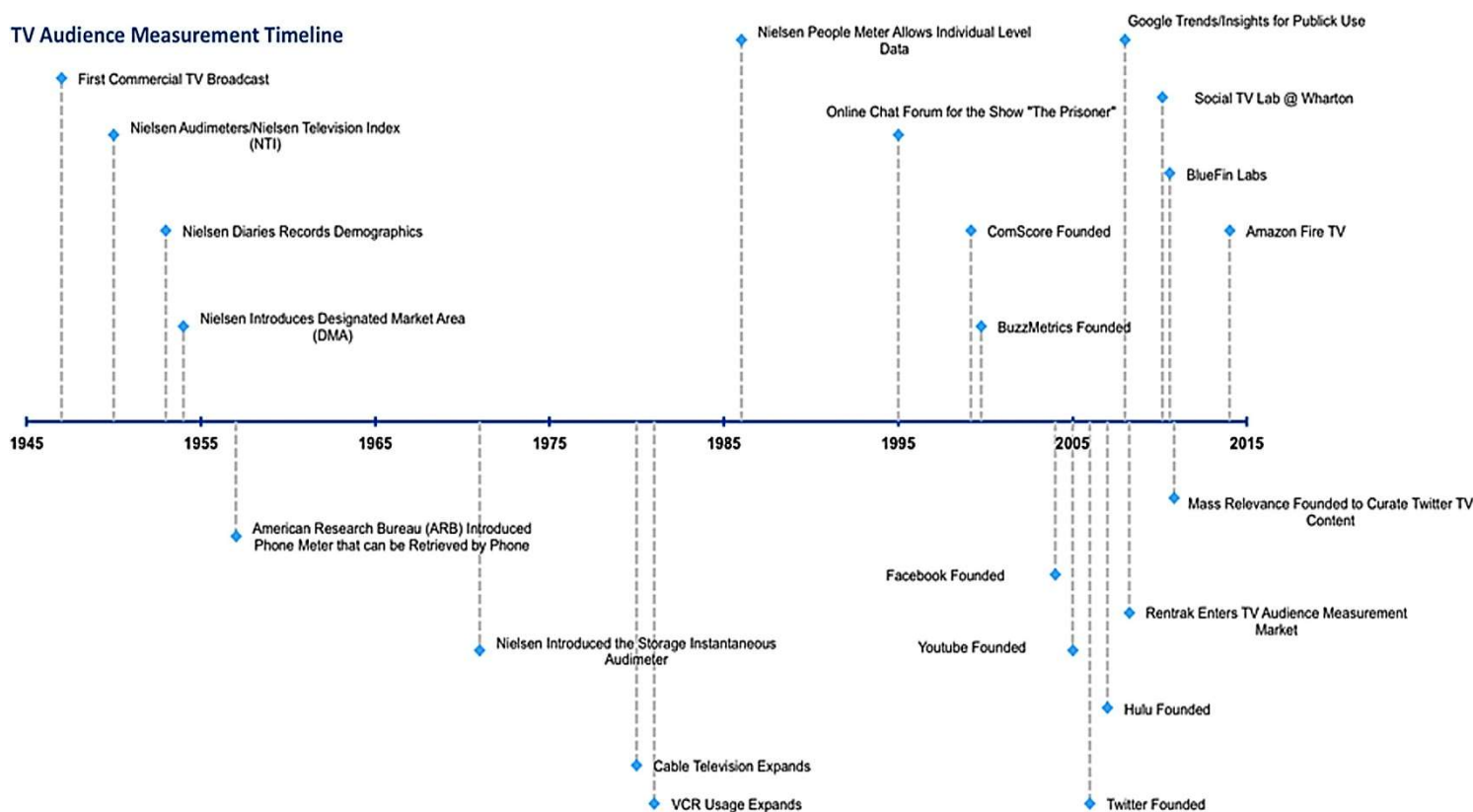


Figure 1. The Evolution of Audience Measurement

2. THEORY

2-a. Working Principle:

The project is intended to come up with an audience measurement system which works as follows. The system will include a basic IR based remote which will trigger the receiver system which will be using TSOP1738 IR receiver to obtain the IR burst from the remote. The intended flow will be working as follows, once the signal is being passed on from the transmitter to the receiving system the received signal will be considered as the code for the specific watermark or channel. This signal will trigger the specific video channel assigned to the respective watermark based on this received signal from various users using the system. Thus, based on the data tracked i.e. the count of the specific watermark being triggered the audience count is considered which will be useful for the overall audience measurement. Other than this the moods and the exact age of the viewers will also be tracked which will help identify the age group watching a specific channel at the same time knowing the instances or moods based on which either they tend to change the channel or react to a specific content of the channel.

1. Video Watermarking

We will be using this system to imply specific watermarks to the respective channels. Based on which the channels will be differentiated and the audience measurement will be heavily considered. We will be using the DWT algorithm for processing the implementing the watermarking process. The DWT Embedding Algorithm works as follows:

- Read the cover image and the watermark logo to be embedded.
- Disintegrate the cover image into four sub-bands namely LL, HL, LH and HH using Haar wavelet.
- Disintegrate the LL band further to the 4" level.
- Apply SVD to the HH (high frequency) band $[U_h \ Sh \ V_h] = \text{svd}(HH, 'econ')$, where econ is the economy decomposition of the matrix.
- Decompose the watermark using SVD.

$$[U_w \ Sw \ V_w] = \text{svd}(\text{watermark logo}, 'econ')$$

- Swap the singular values of the HH band with the singular values of the watermark.
- Generate signature for watermark using algorithm 2.

- The above generated signature is embedded in the LL band by using algorithm 3 and the watermark is embedded in the HU band of the cover image by using the condition

if (length(watermark logo) = 256)

Sh_diag(1:length(Sh), :) = Sw_diag(1:length(Sh),:)

Where, Sh and Sw are the singular values of host image and watermark logo respectively.

- Apply inverse SVD to get the modified HH band which now holds the SVs of the watermark logo.
- Apply inverse DWT with modified HH and LL band to obtain the watermarked image.

Then comes the DWT Extracting Algorithm which works as follows:

- First level DWT is performed on the host image to decompose it into four sub bands LL1, HL1, LH1 and HH1.
- The second level DWT is performed on the LL1 sub band to get four smaller sub bands LL2, HL2, LH2 and HH2.
- The third level DWT is performed on the LL2 sub band to get four smaller sub bands LL3, HL3, LH3 and HH3.
- First level DWT is performed on the watermarked image to decompose it into four sub bands nLL1, nHL1, nLH1 and nHH1.
- The second level DWT is performed on the LL1 sub band to get four smaller sub bands nLL2, nHL2, nLH2 and nHH2.
- The third level DWT is performed on the LL2 sub band to get four smaller sub bands nLL3, nHL3, nLH3 and nHH3.
- Then following extracting is performed to get wLL3 with the extraction formulae with same value of 'a' as in embedding $wLL3 = \text{new LL3} - LL3 / a$
- Apply inverse DWT on wLL3 with all other sub bands (LH ,HL, HH) equal to zero to get wLL2
- Repeat the last step two times each level to get the extracted watermarks.

2. Emotion recognition

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlaps to cover the entire visual area.

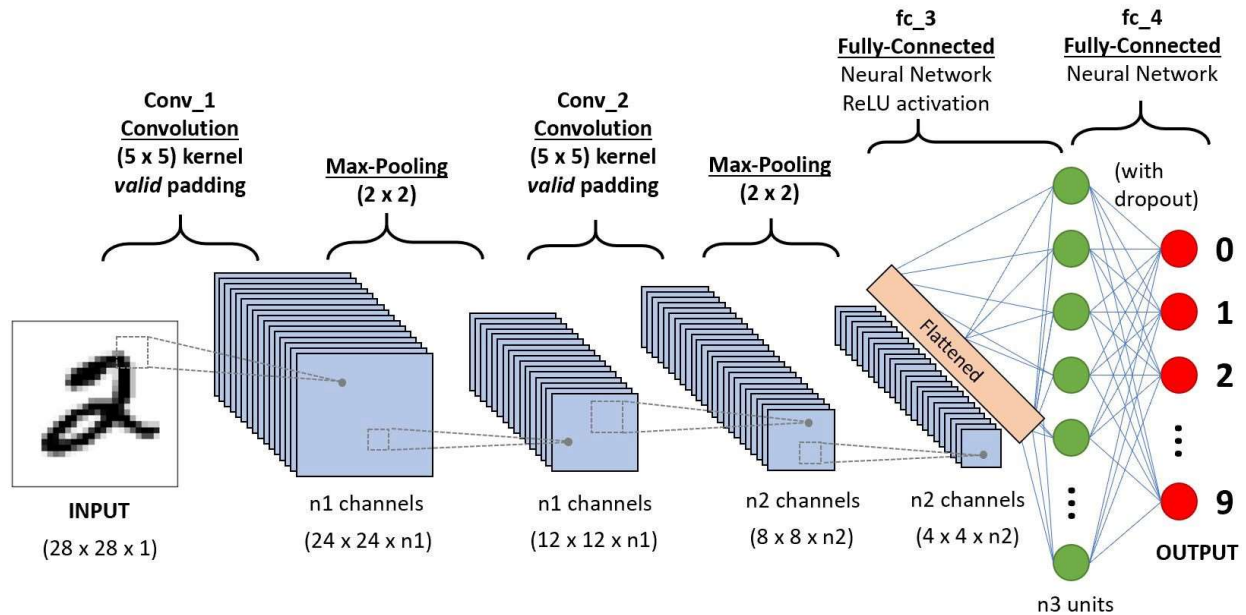


Figure 2. Neural network with many convolutional layers

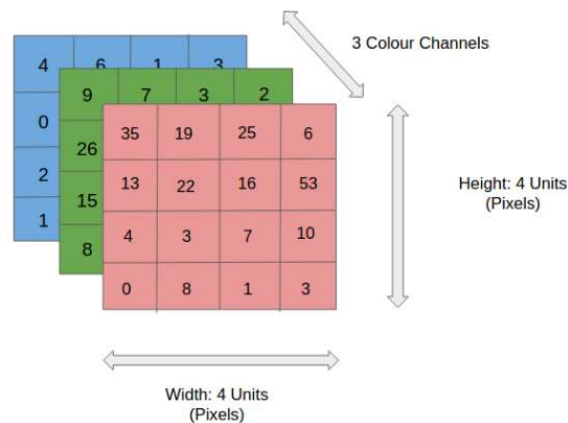


Figure 3. 4x4x3 RGB Input Image

In the figure, we have an RGB image which has been separated by its three-color planes — Red, Green, and Blue. There are a number of such color spaces in which images exist — Grayscale, RGB, HSV, CMYK, etc.

You can imagine how computationally intensive things would get once the images reach dimensions, say 8K (7680×4320). The role of the ConvNet is to reduce the images into a form which is easier to process, without losing features which are critical for getting a good prediction. This is important when

we are to design an architecture which is not only good at learning features but also is scalable to massive datasets.

Convolution Layer — The Kernel

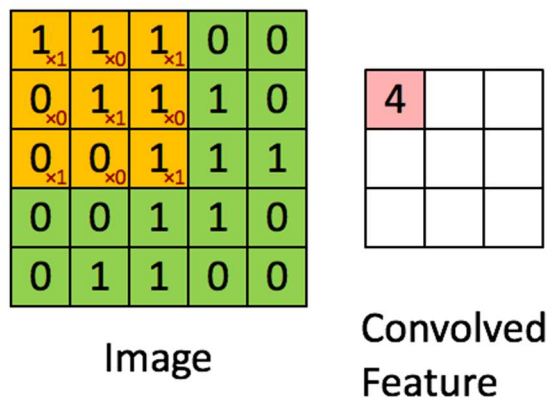


Figure 4. Convoluting a 5x5x1 image with a 3x3x1 kernel to get a 3x3x1 convolved feature

Image Dimensions = 5 (Height) x 5 (Breadth) x 1 (Number of channels, eg. RGB)

In the above demonstration, the green section resembles our 5x5x1 input image, I . The element involved in carrying out the convolution operation in the first part of a Convolutional Layer is called the Kernel/Filter, K , represented in the color yellow. We have selected K as a 3x3x1 matrix. The Kernel shifts 9 times because of Stride Length = 1 (Non-Strided), every time performing a matrix multiplication operation between K and the portion P of the image over which the kernel is hovering.

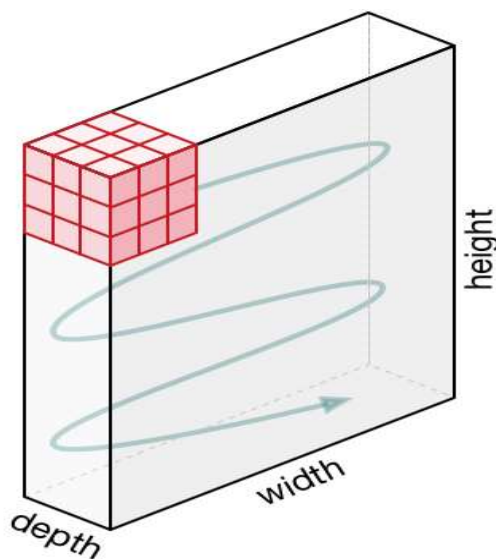


Figure 5. Visualizing convolution in an image

The filter moves to the right with a certain Stride Value till it parses the complete width. Moving on, it hops down to the beginning (left) of the image with the same Stride Value and repeats the process until the entire image is traversed.

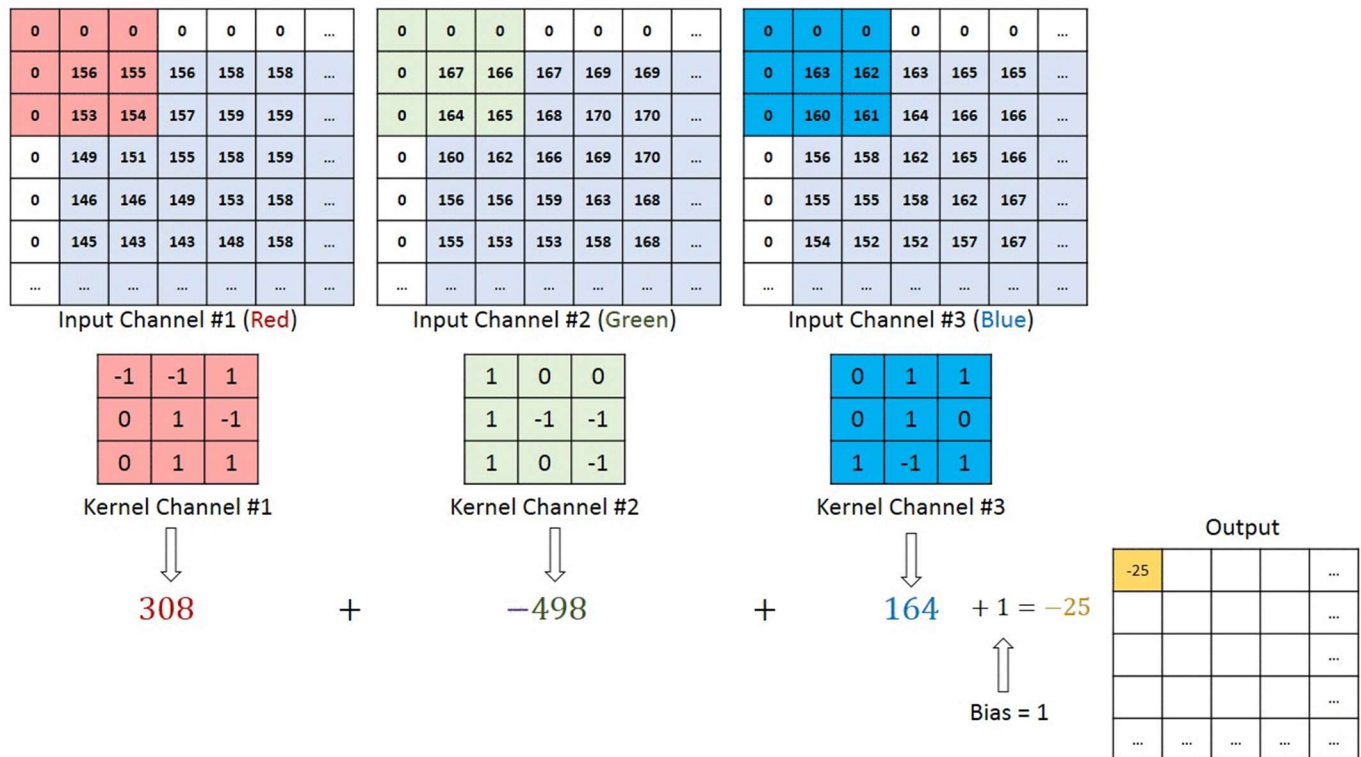


Figure 6. Computation of Output Matrix

In the case of images with multiple channels (e.g. RGB), the Kernel has the same depth as that of the input image. Matrix Multiplication is performed between K_n and I_n stack ($[K1, I1]; [K2, I2]; [K3, I3]$) and all the results are summed with the bias to give us a squashed one-depth channel Convolved Feature Output.

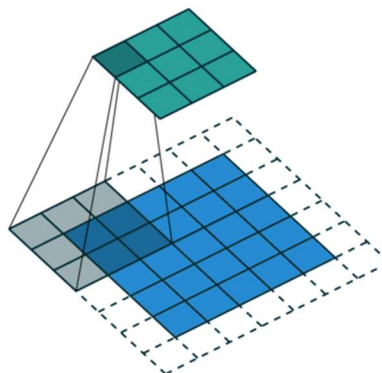


Figure 7. Convolution Operation with Stride Length = 2

The objective of the Convolution Operation is to extract the high-level features such as edges, from the input image. ConvNets need not be limited to only one Convolutional Layer. Conventionally, the first ConvLayer is responsible for capturing the Low-Level features such as edges, color, gradient orientation, etc. With added layers, the architecture adapts to the High-Level features as well, giving us a network, which has the wholesome understanding of images in the dataset, similar to how we would. There are two types of results to the operation — one in which the convolved feature is reduced in dimensionality as compared to the input, and the other in which the dimensionality is either increased or remains the same. This is done by applying Valid Padding in case of the former, or Same Padding in the case of the latter.

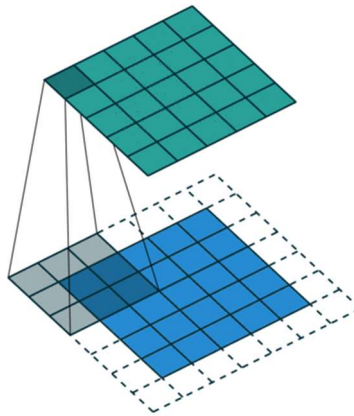


Figure 7. Convolution after Padding

When we augment the $5 \times 5 \times 1$ image into a $6 \times 6 \times 1$ image and then apply the $3 \times 3 \times 1$ kernel over it, we find that the convolved matrix turns out to be of dimensions $5 \times 5 \times 1$. Hence the name — Same Padding. On the other hand, if we perform the same operation without padding, we are presented with a matrix which has dimensions of the Kernel ($3 \times 3 \times 1$) itself — Valid Padding.

3. Dataset Description

We conduct experiments on the FEREC-2013 dataset, which is provided on the Kaggle facial expression competition. The dataset consists of 35,887 gray images of 48×48 resolution. Kaggle has divided into 28,709 training images, 3589 public test images and 3589 private test images. Each image contains a human face that is not posed (in the wild). Each image is labeled by one of seven emotions: angry, disgust, fear, happy, sad, surprise and neutral. Some images of the FEREC-2013 dataset is shown below.



Figure 8. Sample images of the Dataset

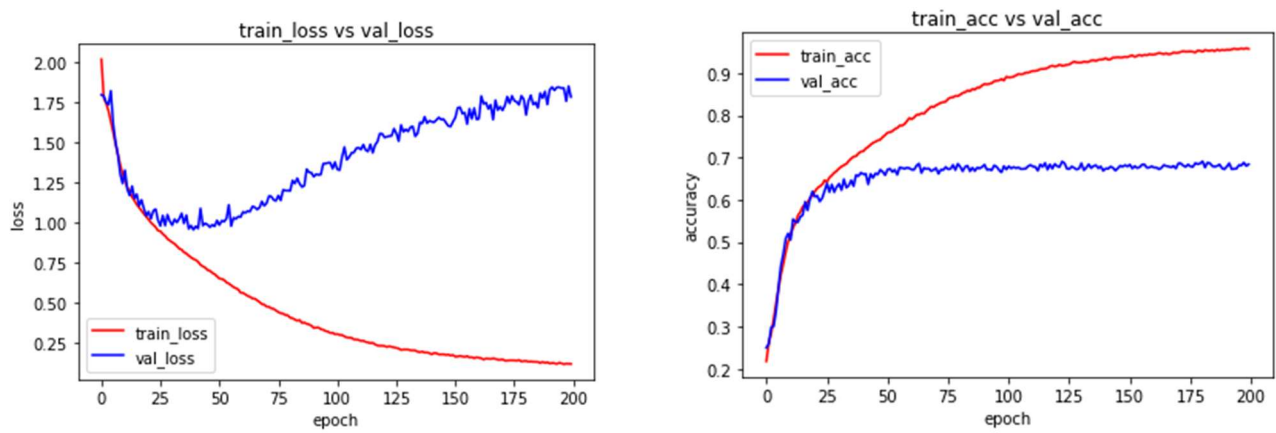


Figure 9. Training Loss and accuracy of the model

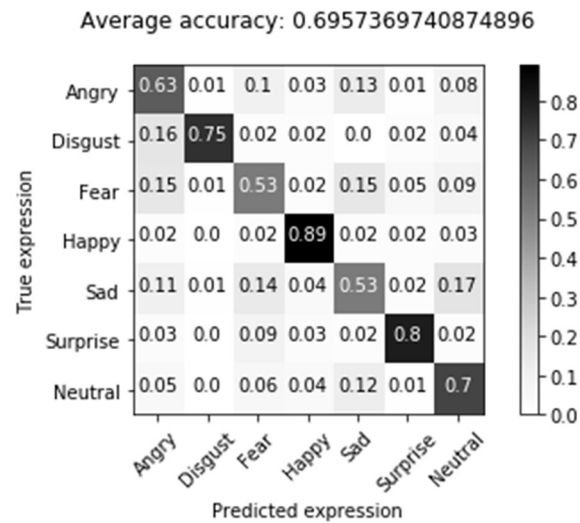


Figure 9. Confusion Matrix with Accuracy

2-b. Block diagram of complete project

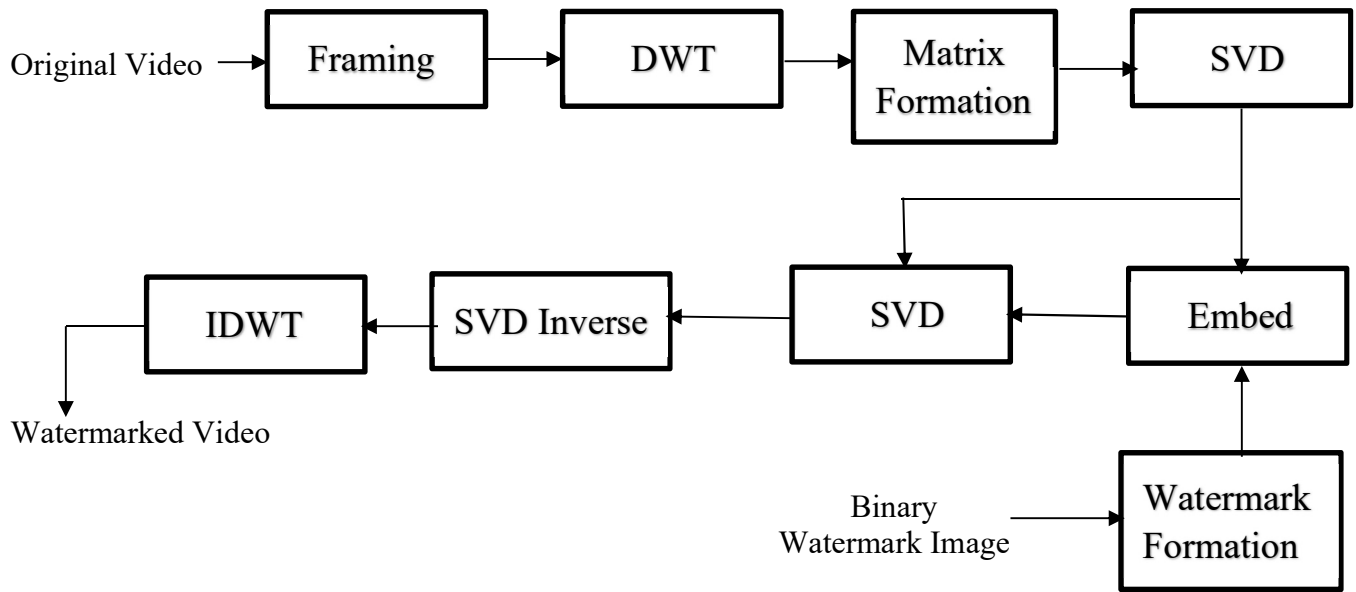


Figure 10. Video Watermarking Process

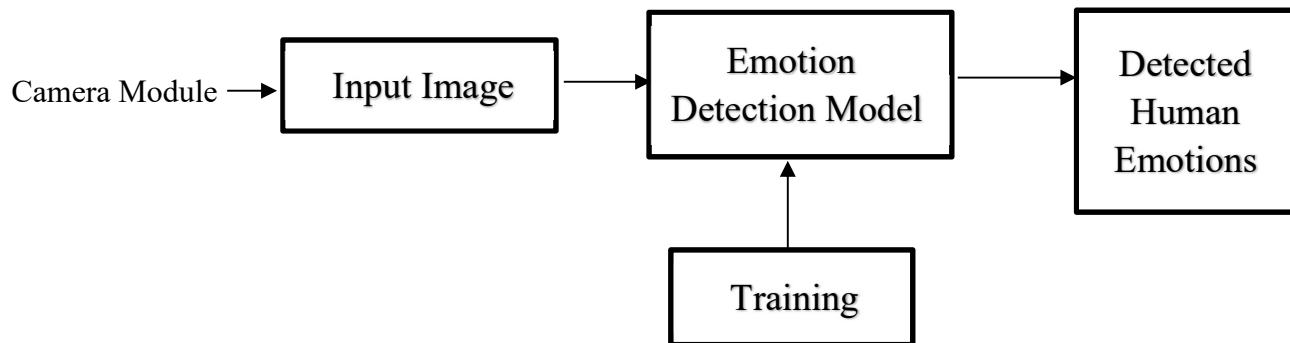


Figure 11. Emotion Detection Process

3. EXTENSIVE LITERATURE REVIEW

➤ **History of Audience Measurement.**

The television industry started a long time back in 1927, but its audience measurement started in 1940s. An overview of television audience measurement through the lens of the various methodologies that have been employed throughout its history and of its evolution as a result of the amount and types of available data. Better TV audience measurement leads to better targeted advertisement opportunities, which should, in turn, lead to increased TV viewer engagement and greater profits for broadcasters and advertisers [3].

Birth (1950): Phone surveys

Since the advent of commercial television in 1947, systems have been in place to measure what audiences are watching. These ratings systems used methodologies developed to measure radio audiences, relying primarily on phone calls to the audience. Clark Hooper eliminated bias and problems with earlier random survey methods that had respondents recall what they had watched/listened to by developing a survey methodology that asked only what listeners were tuned into at the moment they received the call, plus demographic information about who was listening. Known as telephone coincidental, these calls created many of the standard measurements used by television ratings companies such as audience shares. The technique had been considerably refined by 1950 when Nielsen acquired Hooper's business and began its audience monitoring of national television with the Nielsen Television Index.

Infancy (1953): Meters and Diaries

Nielsen primarily used metering devices to measure audiences, another technique developed for tracking radio listeners. These audimeters measured what was being viewed on a television set and when. This eliminated dependence on often unreliable and costly phone surveys, but the system only collected information as to what was on the TV, not who was watching. To remedy this gap, Nielsen began to record more detailed information thanks to a subset of the sample population who kept viewing habits in Nielsen Diaries. The demographic information thus supplemented audiometer collected data. In 1971, Nielsen introduced their version of an American Research Bureau meter whose data could be retrieved by phone line, lowering data-preparation time for the market. This storage instantaneous audimeter stored data during the day and transmitted them overnight. This newfound speed made

Nielsen ratings and reports increasingly important for advertisers, as demographic information on viewers dominated advertising decisions. This system remained largely unchanged until 1986, when the people meter was introduced. This new viewership measurement method allowed for individual data collection from multiple members of a household; individual users recorded their viewing patterns along with demographic information on this device. This major change in data collection created a huge, readily available database that included much more detailed information about exactly who was watching what than had ever been available before. Businesses could now tailor their advertising messages more specifically. This major change in data collection created a huge, readily available database that included much more detailed information about exactly who was watching what than had ever been available before.

Childhood (1986): Cable TV and VCRs

The spread of cable television further altered the type of data collected for audience measurement and how it was used. Cable television was first established in the 1940s to provide TV signals to remote communities. Operators took signals from areas with good reception and then distributed them by coaxial cable to subscribers. Cable systems were able to handle more stations and, beginning in the 1970s, networks designed to be distributed specifically by cable were created, and increasingly diverse programming was introduced. By 2011, more than 5300 systems were in operation, with around 60 million subscribers in the United States. The people meter system's capacity to measure small, demographically targeted audiences made it possible to tailor programming content and develop shows for specific audiences. At the same time, cable's advertiser-supported networks could collect the detailed data needed to lure advertisers of niche products, who could place ads specifically targeted at particular demographic groups. This proliferation of cable networks enhanced the importance of TV ratings and made user data even more valuable for advertisers. They could now dispense with appealing to the lowest common denominator and target their products to those groups most likely to be interested. The widespread adoption of VCRs in the 1980s signaled a further shift in television viewing habits. People could now record television programs and watch them at a later date, a phenomenon known as time shifting. Expanding with the later introduction of digital TV recorders, time shifting dramatically changed how data are collected and used in programming and advertising decisions. The need to keep up with the ever-increasing numbers of shows, as well as take time shifting into account, led to the development of new solutions to automatically detect which programs were being viewed using digital signatures. The System for Measuring and Reporting Television was developed by SRI in 1994, and

Nielsen competed with Active/Passive shortly after. These digital tracking solutions allow the detection of both TV show and TV advertisement signatures. The new possibilities led to an expansion of the measurement market; in parallel with these developments, newcomer Rentrak outpaced its competition by devising solutions for measuring VOD. Fully entering the market in 2008, Rentrak swiftly became a major player on the basis not only of offering viewership and demographic data, but also of being the only company to provide VOD measurements obtained from set-top box data as well as measurement of box office ticket sales, all of which can be linked by time and location.

Adolescence (2000–2010): The Internet and social media

The development of the Internet in the early 2000s altered how people watched TV and integrated it into other aspects of their lives. As people spent more time online, companies recognized that clicks, searches, locations, tweets, purchase, and demographics could all be measured relatively easily at large scale in real time. As the public entered the new world of the Internet, they gave rise to a new world of data. This could be used to measure the impact of TV and advertising on viewers in terms of attention while watching and what they “thought” while watching, as well as what they purchased. Various methods exist for doing this. Starting in 1999, comScore measured demographics, clicks, and purchases, and has since evolved to measure attention across many key platforms (e.g., both home and mobile Internet). Google Trends allow tracking of keyword searches over time and by geographic location. Yet while users are likely to search for a brand online after seeing it advertised on TV, the first online measurement solutions were unable to take TV viewing into account because of data inaccessibility. This has since changed. Ever since a chat forum for discussing the show *The Prisoner* was launched in 1995, a plethora of social TV platforms have popped up, and their usage rates have skyrocketed. Users now engage more with TV shows because shows are also prompting two-way communication.

The coming of age (2013–): VOD and TV everywhere

The proliferation of the Internet and mobile devices such as smartphones and tablets has resulted in major disruption for the TV industry by fundamentally changing the way people watch TV and therefore calls for new models of advertising. The appearance on the scene of services such as Netflix (established in 1997), YouTube (2005), Hulu (2007), and Amazon Fire TV (2014) has led to a world of infinite content—anytime, anywhere, for anyone. In terms of global viewership as a result of online VOD, YouTube is catching up with TV. The vastly different metrics of these media create an image problem for the TV industry among both consumers and investors. During this year's UpFronts, a number of tech

writers highlighted a presentation where Google chairman Eric Schmidt seemed to say that YouTube's success was because the site now has some one billion unique users who consume some six billion videos a month. There is now more content online at any time than could possibly be presented on television. These infinite viewing choices also represent infinite advertising possibilities, meaning greater opportunities to identify users.

Networks have responded with TV Everywhere services. Premiered in 2009 by Time Warner Cable and followed by many other providers, this model allows users to subscribe to specific channels via a user account and view them across multiple devices. The existence of specific accounts means that for the first time, firms can collect data and target users at the individual level, linking such details as personal preferences, locations, demographics, purchases, and social media activity across devices.

Social TV spaces represent repositories of preference information, consumption patterns, social trends, market segmentation information, and opportunities for customization of very fine granularity at the individual and community levels, all of which can be creatively extracted and leveraged to provide enhanced value to customers and advertisers. This makes today an exciting time for TV measurement. This article offers only a brief overview of TV measurement, highlighting only key sources of data and not comprehensively covering all firms and efforts, but Figure 1 illustrates the extent to which its evolution has speeded up in recent years, mainly because of the explosion of Internet and social media data. Just as it is clear that limited data was at the root of the mostly static nature of TV measurement, the removal of such limitations has left the field dynamically evolving. Viewers are no longer guaranteed to be sitting in their living rooms while watching TV, but it is now possible to track where and when they watch shows. These new data sources provide fertile ground for current and future research, as my recent work detailed below demonstrates.

➤ **Digital watermarking using DWT (Discrete Wavelet Transform)**

The popularity of digital video-based applications is accompanied by the need for copyright protection to prevent illicit copying and distribution of digital video. Copyright protection inserts authentication data such as ownership information and logo in the digital media without affecting its perceptual quality. In case of any dispute, authentication data is extracted from the media and can be used as an authoritative proof to prove the ownership. As a method of copyright protection, digital video watermarking has recently emerged as a significant field of interest and a very active area of research. Watermarking is the process that embeds data called a watermark or digital signature into a multimedia object such that

watermark can be detected or extracted later to make an assertion about the object. The object may be an image or audio or video. For the purpose of copyright protection digital watermarking techniques must meet the criteria of imperceptibility as well as robustness against all attacks for removal of the watermark. Many digital watermarking schemes have been proposed for still images and videos. Most of them operate on uncompressed videos, while others embed watermarks directly into compressed videos. The work on video specific watermarking can be further found. Video watermarking introduces a number of issues not present in image watermarking. Due to inherent redundancy between video frames, video signals are highly susceptible to attacks such as frame averaging, frame dropping, frame swapping and statistical analysis [4].

➤ **Facial Feature detection using Haar Classifiers.**

The human face poses even more problems than other objects since the human face is a dynamic object that comes in many forms and colors. However, facial detection and tracking provides many benefits. Facial recognition is not possible if the face is not isolated from the background. Human Computer Interaction (HCI) could greatly be improved by using emotion, pose, and gesture recognition, all of which require face and facial feature detection and tracking. Although many different algorithms exist to perform face detection, each has its own weaknesses and strengths. Some use flesh tones, some use contours, and other are even more complex involving templates, neural networks, or filters. These algorithms suffer from the same problem; they are computationally expensive. An image is only a collection of color and/or light intensity values. Analyzing these pixels for face detection is time consuming and difficult to accomplish because of the wide variations of shape and pigmentation within a human face. Pixels often require reanalysis for scaling and precision. Viola and Jones devised an algorithm, called Haar Classifiers, to rapidly detect any object, including human faces, using AdaBoost classifier cascades that are based on Haar-like features and not pixels [5].

4. WORK TO BE CARRIED OUT IN SEMESTER VIII

- **Audience Age and Gender Detection:** Currently our project is able to detect Emotions of the audience. We plan on expanding the project to further detect the Age and Gender of Audience
- **Audience Mood Detection:** Along with the Age and emotion detection we also plan to detect the Mood of Audience using Speech Processing

- Providing the user, a detailed report about their usage of each channels, hence they can choose an effective plan for themselves by dropping the least viewed or unattended channels.
- Work with the hardware to integrate all the systems, and expect full flow of the system.

5. PROJECT BUDGET SHEET

Component	Quantity	Price (in rupees per piece)
Rpi – 3	1	2200
Ethernet Cable	1	300
SanDisk 128GB Class 10 microSDXC Memory Card with Adapter (SDSQUAR-128G-GN6MA)	1	1300
He Retail Supplies Ir Infrared Receiver V1838b T11838 Infrared Sensor 1838 38khz Electronic Components Electronic Hobby Kit	6	1200 (200*6)
Rpi Camera Module v2	1	2000
Infrared Remote Control	1	60
		Total = Rs. 7600

6. APPLICATION OF THE PROJECT

- The system can compute various important parameters like TRP (Television Rating point), GRP (Gross Rating Point) etc. that effectively represents viewership of a particular channel. This collected data is of great value to various advertisement companies for optimizing their marketing framework. These companies will have access to the real time data that can be utilized to decide on how much resources are to be invested for advertising. Details like age distribution, gender distribution, and regional distribution etc. of the audience can boost the revenue of the companies as now they can access this information.
- Taking advantage of the efficient, robust algorithm and transparent infrastructure any kind of manipulation in data can be avoided by collecting more samples and considering a wider demographic region.
- For implementation customer is provided monthly subscription and other value-added offers. Also, he/she has the added privilege to control the content being displayed on the television as him watching a particular show will motivate the content providers to come up with more of that kind of shows.

- Based on the viewership any vital information regarding various Government policies like Pradhan Mantri Jan Dhan Yojana (PMJDY), Pradhan Mantri Sukanya Samriddhi Yojana (PMSSY) or any measure to be implemented like demonetization can be spread more effectively over the required demographic areas.
- Information regarding free medical camps, polling booths etc. can be advertised based this viewership data.

7. REFERENCES

- [1] Gonzales and Woods, —Digital Image Processing, Pearson Education, India, Third Edition
- [2] Singh, Prabhishek, and R. S. Chadha. "A survey of digital watermarking techniques, applications and attacks." International Journal of Engineering and Innovative Technology (IJEIT), 2013
- [3] S. Hall, "TV audience measurement with big data," Jun. 2014.
- [4] Sinha, Sanjana, Prajnat Bardhan, Swarnali Pramanick, Ankul Jagatramka, Dipak K. Kole, and Aruna Chakraborty. "Digital video watermarking using discrete wavelet transform and principal component analysis." International Journal of Wisdom Based Computing, 2011.
- [5] Wilson, Phillip Ian, and John Fernandez. "Facial feature detection using Haar classifiers." Journal of Computing Sciences in Colleges, 2006.
- [6] Chen, Lawrence S., Thomas S. Huang, Tsutomu Miyasato, and Ryohei Nakatsu. "Multimodal human emotion/expression recognition." In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 366-371. IEEE, 1998.