ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Life-Logging through User Centric Data Analysis

Aadithya G Bharadwaj, Kedaar R Rao, Mahesh Kumar, Vijay Vardhan T

Junior Networks Engineer in Avaya Inc., Bengaluru, India

Junior Software Engineer, Analytics Quotient Services India Pvt Ltd, Bengaluru, India

Software Engineer, Cerner Healthcare, Bengaluru, India

Trainee Engineer, IMS Health, Bengaluru, India

ABSTRACT: Life-logging represents a phenomenon whereby people can digitally record their own daily lives in varying amounts of detail, for a variety of purposes. In a sense it represents a comprehensive black box of a human's life activities and may offer the potential to mine or infer knowledge about how we live our lives. We have recently observed a convergence of technologies to foster the emergence of Life-logging as a mainstream activity. Computer storage has become significantly cheaper, and advancements in sensing technology allows for the efficient sensing of personal activities, locations and the environment. This is best seen in the growing popularity of the quantified self-movement, in which life activities are tracked using wearable sensors in the hope of better understanding human performance in a variety of tasks. This paper aims to provide a comprehensive summary of Life-logging, to cover its research history, current technologies, and applications. Thus far, most of the Life-logging research has focused predominantly on life-logging in order to capture life details of life activities, hence we maintain this focus in this review. However, we also reflect on the challenges Life-logging poses to an information retrieval scientist. This paper is a suitable reference for those seeking information retrieval scientist's perspective on Life-logging and the quantified self.

KEYWORDS: Life-logging**,** User-Centric, Image Segmentation, Data Acquisition, Wearable Cameras, Object Detection,Image Captioning, Text to Speech Conversion, Geo-tagging

I. INTRODUCTION

Life-logging [1] is referred to as "a form of pervasive computing, consisting of a unified digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive". The unified digital record uses multi-modal captured data which has been gathered, stored, and processed into semantically meaningful and retrievable information and has been made accessible through an interface, which can potentially support a wide variety of use-cases. It is a trend that is rapidly increasing thanks to advances in wearable technologies over recent years. Nowadays, wearable cameras are very small devices that can be worn all day long and automatically record the everyday activities of the wearer in a passive fashion, from a first-person point of view. As an example, Fig. 1 shows pictures taken by a person walking down a street while wearing such a camera. Most wearable cameras on the market such as GoPro, MeCam, Looxcie or Google Glass (see Fig. 2 (a) and (c)) are video cameras, which have relatively high temporal resolution (HTR) (25-30 frames per second) and have become popular among sportsmen over recent years. A limited number of wearable cameras, such as Narrative Clip and SenseCam (see Fig. 2 (b) and (d)) are photographic cameras, which have low temporal resolution (LTR) (2-3 frames per minute), and hence are more suitable for acquiring data over long periods of time.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Figure 1: Example of a sequence acquired by the Narrative Clip wearable

The following section gives the reason behind implementing life-logging.
The current method is to capture images and store them as raw data. The issue with this is the number of images is continuously increasing and the amount of data is increasing exponentially. Even though we have cloud storage with the option for users to box images into a clusters and separate them on the basis of the day or place where they take the image, the issue still remains the same which is the amount of data is very vast and will be very tedious for the user to go through them again and the information received from the images is very less. To prevent this we make use of life-log. Life-logging is becoming more accessible and feasible to users because of the following reasons. Computer storage has become incredibly cheap, both on the cloud and as personal storage. In fact we have seen exponential growth in disk storage capacity over the lifetime of digital storage. We are seeing advances in sensors for sensing the person as well as sensing the person’s environment which are making such sensors cheap, robust and unobtrusive. There is growing social interest in the phenomenon of sensing and Motivation of recording oneself, the so-called quantified-self movement. Sometimes this is driven by applications like sports and health/wellness, other times it is sensing just because we can. We can observe an increased openness to storing and sharing information about ourselves as can be seen in social networks.

II. RELATED WORKS

The following section reviews the current state of art methodologies used for life-logging.
In machine learning, a Convolutional Neural Network [2] (CNN) is a type of feed-forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex, whose individual neurons are arranged in such a way that they respond to overlapping regions tiling the visual field. Convolutional networks were inspired by biological processes and are variations of perceptron. When used for image recognition, convolutional neural networks (CNNs) consist of multiple layers of small neuron collections which process portions of the input image, called receptive fields. The outputs of these collections are then tiled so that their input regions overlap, to obtain a better representation of the original image; this is repeated for every such layer. Tiling allows CNNs to tolerate translation of the input image. Convolutional networks may include local or global pooling layers, which combine the outputs of neuron clusters.

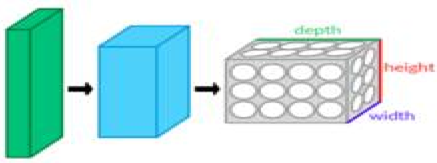

Figure 2: CNN layers arranged in 3 dimensions

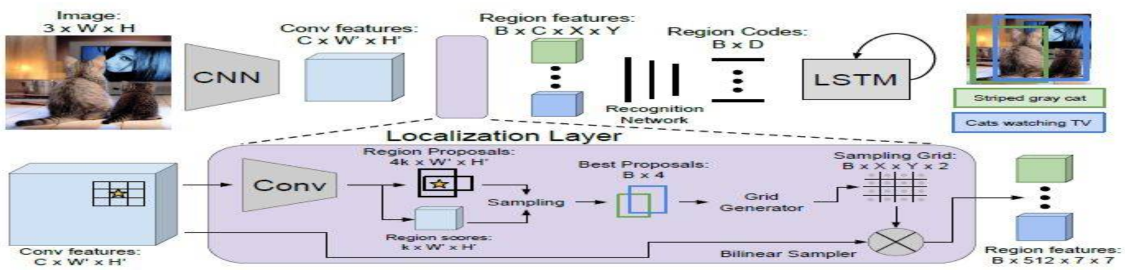

Figure 3: Model overview. An input image is first processed a CNN. The Localization Layer proposes regions and smoothly extracts a batch of corresponding activations using bilinear interpolation. These regions are processed with a fully-connected recognition network and described with an RNN language model. The model is trained end-to-end with gradient descent.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

The following are the limitations of Convolutional Neural Network. Convolutional Neural Network are tend to be applied indiscriminately to problems for which they are not well suited. ConvolutionalNeural networks are too much of a black box**.** This has several consequences. It makes them difficult to train: the training outcome can be nondeterministic and depend crucially on the choice of initial parameters it makes them hard to determine how they are solving a problem, because they are opaque. It makes them difficult to troubleshoot when they don't work as you expect, and when they do work, you will never really feel confident that they will generalize well to data not included in your training set because, fundamentally, you don't understand how your network is solving the problem, or what idiosyncrasies of the training data the network has over fit to.

III. PROPOSED CONCEPT

The following section gives a brief description of the proposed system. Recursive neural networks comprise a class of architecture that can operate on structured input. Deep connectionist architectures involve many layers of nonlinear information processing. This allows them to incorporate meaning representations such that each succeeding layer potentially has a more abstract meaning. Recent advancements in efficiently training deep neural networks enabled their application to many problems, including those in natural language processing (NLP). A key advance for application to NLP tasks was the invention of word embedding that represent a single word as a dense, low-dimensional vector in a meaning space, and from which numerous problems have benefited. Recursive neural networks, comprise a class of architecture that operates on structured inputs, and in particular, on directed acyclic graphs. A recursive neural network can be seen as a generalization of the recurrent neural network, which has a specific type of skewed tree structure. They have been applied to parsing, sentence-level sentiment analysis, and paraphrase detection. Given the structural representation of a sentence, e.g. a parse tree, they recursively generate parent representations in a bottom-up fashion, by combining tokens to produce representations for phrases, eventually producing the whole sentence. The sentence-level representation (or, alternatively, its phrases) can then be used to make a final classification for a given input sentence — e.g. whether it conveys a positive or a negative sentiment. Under the recursive neural networks framework the perceived structure of a problem is captured and expressed by using graphical models. In particular, the patterns used for the learning and recall phases not only encode the fragments of information (e.g.: information that can be characterized by specific attributes that are quantifiable and/or measurable) which play an important role in the machine learning problem but also the logical relationships between them. The nature of such relations is determined by the application context and attempts to explicitly model the logical correlations between fragments of information.

Figure 4: Example captions generated and localized on test images. The bottom row depicts the contrast amount of information RNN produces to that of CNN

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

There are basically seven modules, which are discussed in detail in the following section.

1.Data Acquisition: Data acquisition is defined as the process of collecting and organizing information. The data here is a set of images. It is important to make sure that the image contains the required objects. The accuracy of the output is determined mainly by the objects in the image. The objects that are included in an image must also be classified according to the level of importance. Not all objects in an image are required for further analysis. We now describe how the candidate images are selected in order to develop the training dataset. There are various places to acquire the images from like MS-COCO [3] or we can collect them ourselves.

2. Image Pre-Processing: Images are often degraded by noises. Noise can occur during image capture, transmission, etc. Noise removal is an important task in image processing. In general the results of the noise removal have a strong influence on the quality of the image processing technique. . Linear filters are not able to effectively eliminate impulse noise as they have a tendency to blur the edges of an image. On the other hand non-linear filters are suited for dealing with impulse noise. Several non-linear filters based on Classical and fuzzy techniques have emerged in the past few years. For example most classical filters that remove simultaneously blur the edges, while fuzzy filters [4] have the ability to combine edge preservation and smoothing. Keeping this in mind we make use of Gaussian Filter [5] in order to remove the Gaussian noise present in our input images. Gaussian noise is statistical noise having an $Probability\ Distribution\ Function\ (PDF)$ [6] equal to the normal distribution, which is also known as Gaussian distribution. Gaussian Filteris used to blur the image and remove noise.

3. Image Segmentation: The next phase after Pre-Processing the image is Image Segmentation [7]. Segmentation is one of the most important process in image processing. It consists of constructing a symbolic representation of the image: the image is described as homogeneous areas according to one or several a priori attributes. Keeping this in mind we fell Watershed technique is best suited for Life-logging. The watershed transform can be classified as a region-based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is flooded by water, watersheds being the divide lines of the domains of attraction of rain falling over the region. An alternative approach is to imagine the landscape being immersed in a lake, with holes pierced in local minima. Basins (also called 'catchment basins') will fill up with water starting at these local minima, and, at points where water coming from different basins would meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins separated by dams, called watershed lines or simply watersheds.When simulating this process for image segmentation, two approaches may be used: either one first finds basins, then watersheds by taking a set complement; or one computes a complete partition of the image into basins, and subsequently finds the watersheds by boundary detection. To be more explicit, we will use the expression 'watershed transform' to denote a labelling of the image, such that all points of a given catchment basin have the same unique label, and a special label, distinct from all the labels of the catchment basins, is assigned to all points of the watersheds. An example of a simple image with its watershed transform is given in Fig. 6(a-d).

Figure 5: Examples of watershed segmentation by immersion. (a): synthetic image; (b): watershed transform of (a); (c): natural image; (d): watershed transform of (c). Different basins are indicated by distinct grey values.

4. Object Detection: The next phase after segmenting the image is object detection [8]. Object recognition is an important task in image processing and computer vision. Extracting the points from an image that can give best define from an object in image namely key points is very important and valuable. These points have many applications in image processing like object detection, object and shape recognition, image registration and object tracking. By extracting the key points, we can use them for finding objects in the images. The Gabor filter is a multi-channel filtering which makes use of a bank of even-symmetric filters to functionally characterize the different channels. Those

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

filtered images that possess a significant component of the original images are selected, and the rest are discarded. The select filtered images are then subjected to a non-linear transformation that behaves like a blob detector. An energy measures is defined on the transformed images in order to compute different texture features for the different blobs. The measure calculates a sum of texture energies over a small windows around each transformed pixels in the select filtered images. Using the energy measure as features, a vector of Gabor features is defined for each pixel in the original image. The Gabor feature, and the spatial co-ordinates of the corresponding pixels, are subsequently inputted to a squared-error based clustering algorithm. This clustering algorithm yields a segmentation of the original image it assigns to each pixel in the original image a cluster label. This label identifies the amount of texture energy that a pixel possess across the different spatial orientations and frequencies (see Figure—for an overall schematic of the proposed technique).

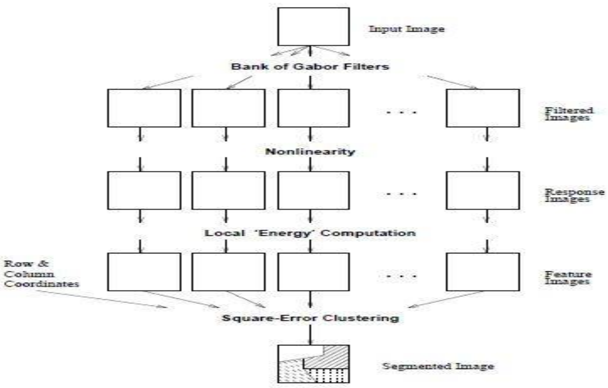

Figure 6: An overview of the texture segmentation algorithm

5. Image Captioning: Image caption are also known as cutline, are a few lines of text used to explain the image. In some cases captions and cutline are distinguished, where the caption is a short (usually one-line) title for the photo, while the cutline is a longer, prose block under the caption, generally describing the photograph, giving context, or relating it to the article. Automatically generating captions of an image is a task very close to the heart of scene understanding one of the primary goals of computer vision. Not only must caption generation models be powerful enough to solve the computer vision challenges of determining which objects are in an image, but they must also be capable of capturing and expressing their relationships in a natural language. For this reason, caption generation has long been viewed as a difficult problem. It is a very important challenge for machine learning algorithms, as it amounts to mimicking the remarkable human ability to compress huge amounts of salient visual information into descriptive language. Hence we make use of LSTM (Long Short Term Memory) [9] using Neural Network Captioning.

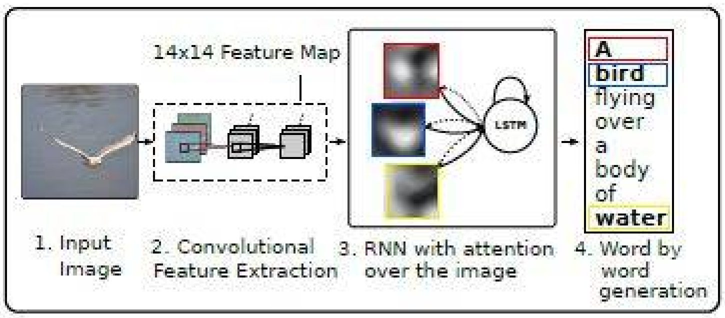

Figure 7: Architecture for Image Captioning Using LSTM

6. Geo-tagging: Geo-tagging is essential for visual life- logging. Geo-tagging is the process of adding geographical information to various media in the form of metadata [10].The data usually consists of coordinates like latitude and longitude, but may even include bearing, altitude, and distance and place names. When tagging

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

photos, the tagging may be done by the camera after processing the shot, or it can be applied when the photo is posted online. Some high-end cell phones have built in GPS, so they can geo-tag any photos taken using the phone's camera. Geo-tagging is most commonly used for photographs and can help people get a lot of specific information About where the picture was taken or the exact location of a friend who logged on to a service.

Geo-tagging location services can be used to find location-specific websites, news and other information. It is based on positions and co-ordinates and is often directly taken from a global positioning system (GPS). And if you are aware of your location data and are using it wisely there are many advantages of geo-tagging.

- Sharing locations with specific friends (such as with path) to help meet up at concerts or festivals.
- Documenting holidays by having your holiday photos arranged by location or uploaded to Google Maps automatically.
- Aiding disaster recovery teams.
- Sophisticated augmented reality App.

One of the biggest ways in which geo-tagging is currently being pushed is for location based marketing. This started off with companies such as Starbucks offering free coffee to people who became Mayor of their coffee shop on Foursquare. Now you can get push notifications from high street stores when you check-in somewhere nearby. With the advent of Facebook Places, and all the personal data that Facebook stores about individuals, you can bet that soon you will be receiving more and more accurately targeted marketing notifications. If you list photography as your main hobby on Facebook then you will probably be told of camera offers whenever you are near a photography shop. And because they will have access to your trends over time it will help the adverts be more refined to your movements and habits and also provide ads over mixed media depending on your tastes. While not all cameras capture the geo-tagging information, a lot of cameras do, such as the Canon brand; and such cameras will include the metadata (shutter speeds, flash, and other settings) embedded into the individual photo. In GPS devices like iPhones and other smart- phones, additional information is automatically captured, such as longitude and latitude of where the image was taken. Using particular tools and programs, a person determined to find out this type of information can decipher any geo-tagged image and retrieve the information for their own purposes. Geo-tagging websites such as Yelp, Foursquare and Facebook's Places are also able to pinpoint your location when you check into public places such as cafes, restaurants, and theatres. Photo sites such as Flickr and Picasa can provide geo-tagging information but this is not an automatic function.

7. Text to Speech Conversion: Life-Logging can be made effective by adding text to speech conversion. Text to speech (TTS) [11] is a natural language modeling process that requires changing units of text into units of speech for audio presentation. This is the opposite of speech to text, where a technology takes in spoken words and tries to accurately record them as text. Text to speech is now common in technologies that seek to render audio output from digital text to assist those who are unable to read, or for other kinds of uses. They should not be confused with voice response systems. Voice response systems synthesize speech by concatenating sentences from a database of prerecorded words and are used for different purposes than such systems, which form sentences and/or phrases based on a language's graphemes and phonemes. Voice response systems are limited to synthesizing sentences that contain only words that have been predetermined by the system. Text to Speech in contrast, is theoretically capable of "reading" any string of text characters to form original sentences.

It is so simple you can adopt it as a tool to help read, proof read, and study a language or just for enjoyment. Text-to-speech software can:

1) Assist persons with disabilities who may be vision impaired by reading text
2) Save eye strain – relax, sit back and listen
3) Save time - listen while driving, exercising or enjoying nature (mp3 converter feature)
4) Help writers improve by listening to your work
5) Assist second language students by expanding their experience and understanding through listening to any text at any speed

This feature caters to the needs of different kinds of Users:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

•**Learning disabilities** – Some people have difficulty reading large amounts of text due to dyslexia and other learning disabilities. Offering them an easier option for experiencing website content is a great way to engage them.
•**Literacy difficulties** – Some people
Have basic literary levels. They often get frustrated trying to browse the internet because so much of it is in text form. By offering them an option to hear the text instead of reading it, they can get valuable information in a way that is more comfortable for them.
•**People who speak the language but do not read it** – Having a speech option for the foreign born will open up your audience to this under-served population. Many people who come to a new country learn to speak and understand the native language effectively, but may still have difficulty reading in a second language. Though they may be able to read content with a basic understanding, text to speech technology allows them to take in the information in a way they are more comfortable with, making your content easier to comprehend and retain.
•**Multitasking Users** – A busy life often means that people do not have time to do all the reading they would like to do online. Having a chance to listen to the content instead of reading it allows them to do something else at the same time. With the prevalence of smartphones and tablets, it also provides an option for content consumption on the go, taking con- tent away from the computer screen And into any environment that's convenient for the consumer.
•V**isual Impairment** – Text to speech can be a very useful tool for the mild or moderately visually impaired. Even for people with the visual capability to read, the process can often cause too much strain to be of any use or enjoyment. With text to speech, people with visual impairment can take in all manner of content in comfort instead of strain.
•**People who access content on mobile devices** – Reading a great deal of content on a small screen is not always easy. Having text-to-speech software doing the work is much easier. It allows people to get the information they want without a great deal of scrolling and aggravation.
•**People with different learning Styles**– Some people are auditory learners, some are visual learners, and some are kinesthetic learners most learn best through a combination of the three. Universal Design for Learning is a plan for teaching which, through the use of technology and adaptable lesson plans, aims to help the maximum number of learners comprehend and retain information by appealing to all learning styles.

IV. APPLICATIONS

The following sections summarizes the applications of life-logging.
Many of the Life-logging applications come from the healthcare or quantified-self domains. Monitoring of dietary intake, sleep monitoring is in demand in the market. One of the early targets for personal Life-logging, especially life-logging from wearable cameras, was helping to overcome the difficulties some people have with short term memory recall, especially for people with Alzheimer's and other dementias.

V. CONCLUSION AND FUTURE WORKS

It is difficult to predict whether life logging will become mainstream or not. The first generation of use cases are only beginning to be understood and the potential is not yet clear. We may find that wearable sensing that impacts on bystanders and society as a whole (specifically cameras) is a step too far for society, or the benefits of extreme life logging may not live up to the potential. On the other hand, the new life logging technologies could be embraced as society has embraced smartphones and social networking in recent years. We have yet to see. A lot is dependent on the first generation use cases, which require the input of information retrieval to develop effective life logging applications of surrogate memories. The effectiveness of the underlying information organization tools will be key in whether life logging takes off as a commonly employed technology. As the field of information retrieval inevitably helped to progress web search, a similar requirement would be made on information retrieval to help progress life logging from a topic for early adopters to become a widely used and beneficial technology. It may appear at present that the future uptake of extreme life logging would be very small; however we can already see the uptake of the initial quantified-self and health-based applications and if this trend continues to increase, then life logging may be following a few years behind. If that is the case, life logging will surely keep the Information Retrieval scientist busy for the foreseeable future.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Life logging is a technology in its infancy. There are numerous challenges each of which demands attention. Not with standing these, and with the caveat that we are only beginning to explore the concept of life logging, we suggest some potential future directions for life logging to follow:

• **Enhanced Capture:** Currently our sensing technologies are still relatively rudimentary. We can sense aspects of a person's environment, their actions/interactions, a lot of what they see and hear, and to some extent their interests (via information interactions). However, this only provides a snapshot of life activities for the life log, and it remains for the surrogate memory to generate semantic interpretations of this data and make it useful for the user. More detailed human sensing (via next generation wearable sensors, Brain Computer Interfaces, etc.) would allow for much more detailed capture of the semantics of life, where concepts such as sentiment/mood/emotion could conceivably be captured into the life log. While this may seem far-fetched, continual advances in Brain Computer Interface technology slowly edge us closer to this point. The end point of this effort would be the modeling of both episodic and semantic memory within life logs, whereas currently life logging focuses on episodic memory logging;

• **Integration of life logging context information with search engines:** Life logging, as discussed earlier, provides a rich and novel source of data for contextual information retrieval. It is likely that new approaches to contextual understanding for omnipresent and intelligent search will emerge from both real-time life log streams as well as historical context from life log archives. Utilizing this data will provide both new challenges and new opportunities to the Information Retrieval researcher;

• **Use-cases:** As life logging becomes a more popular activity, the use cases will become clearer. This will inform development of capture technologies but most importantly, from the retrieval point of view, it will inform the information needs for real-world users, as opposed to relying on the early adopters or the five Rupees of memory access. The book "Total Recall" provides a good initial overview of the potential use cases.

• **Anonymity of Life logs:** With privacy concerns in mind, the issue of anonymity unknown passers-by in life logs will surely receive research attention. It is our belief that this should be an access-time, dynamic process, based on user access policies, as opposed to a non-reversible capture-time process.

• Recreating the Person: A more far-fetched concept, but one that is receiving attention and funding at present is the idea that the life log can be used to recreate the individual in digital form (an avatar) by using the detailed trace from the life log as source data. Some futurists have gone so far as to suggest that the human is an information processing machine, the memory of which can be replicated / enhanced indefinitely inside of an information processing machine that relies on life log data as the source of memory data.

• **Humanizing Technology:** A final and perhaps more down-to-earth concept is that life log principles can be applied to inanimate items to provide a semblance of episodic memory or personality; an early prototype which uses life logging technology to put a human-like interface to a standard coffee machine. This has most potential when combined with research into the development of humanoid robotics.

REFERENCES

[1] Marc Bolanos, Mariella Dimiccoli, and Petia Radeva, "Towards Storytelling from Life-logging: An Overview", *Journal of Transactions on Human-Machine Systems,* July 2015.

[2] Rahul Mohan, "Deep De-convolutional Network for Scene Parsing", *International Conference on Machine Learning",* November 2014.

[3] Tsung-Yi Lin, Micheal Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramana, C. Lawerence Zitnick, and Piotr Dollar, "Microsoft COCO: Common Objects in Context", *ECCV,* 2014.

[4] R. Saranya Pon Selvi, and C. Lokanayaki, "A Survey on Fuzzy Image Segmentation Techniques", *Int. Journal of Engineering Research and Applications www.ijera.com ISSN: 2248-9622, Vol. 4, Issue 3(Version 1) March 2014*

[5] Angus Leigh, Alexander Wong, David A. Clausi, and Paul Fieguth, "Comprehensive Analysis on the Effects of Noise Estimation Strategies on Image Noise Artifact Suppression Performance", *Computer Vision, IEEE International Conference,* August 2015.

[6] Mrudula Karande, and D. B. Shirsagar, "Probabilistic Model Based Image Segmentation", *International Journal of Multimedia and Its Applications [IJMA] Vol.6 No.2, April 2014*

[7] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik, "Contour Detection and Hierarchical Image Segmentation", CVPR, July 2008.

[8] Reza Oji, "An Automatic Algorithm For Object Recognition And Detection Based on ASFIT Key Points", *Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.5,* and October 2012.

[9] Justin Johnson, Andrej Karpathy, Li Fei-Fei, "Densecap: Fully Convolutional Localization Networks for Dense Captioning", *International journal of computer vision,* March 2015.

[10] Peng Wang, Alan F. Smeaton, Yuchao Zhang, Bo Deng, "Enhancing The Detection of Concepts For Life logs Using Contexts Instead of Ontologies", *International Journal of Multimedia Information Retrieval,* 2014.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

[11] Mrunmayee Patil, Ramesh Kagalkar, "An Automatic Approach for Translating Simple Images into Text Descriptions and Speech for Visually Impaired People", *International Journal of Computer Applications*, May 2015.

BIOGRAPHY

Aadithya G Bharadwaj is a Junior Networks Engineer in Avaya Inc. He received Bachelor of Engineering degree in Information Science and Engineering in 2016 from The Oxford College of Engineering from VTU, Belgaum, India. His interests are in the field of Computer Networks and Telecommunications.

Kedaar R Raois a Junior Software Engineer in Analytics Quotient Services India Pvt Ltd**.** He received Bachelor of Engineering degree in Information Science and Engineering in 2016 from The Oxford College of Engineering from VTU, Belgaum, India. His interests are in the field of Operating System and Internet of Things.

Mahesh Kumar is a Software Engineer in Cerner Healthcare. He received Bachelor of Engineering degree in Information Science and Engineering in 2016 from The Oxford College of Engineering from VTU, Belgaum, India. His interests are in the field of Machine Learning and Computer Vision.

Vijay Vardhan Tis a Trainee Engineer in IMS Health. He received Bachelor of Engineering degree in Information Science and Engineering in 2016 from The Oxford College of Engineering from VTU, Belgaum, India. His interests are in the field of Analytics and Web Development.