

A Smart Algorithm That Enables an Artificial Eye into an Intelligent Artificial Eye

Ankita Tiwari^{*1}, O. P. Singh¹, Arun Kumar Tiwari²

¹Department of Electronics and Communication Engineering Amity School of Engineering and Technology

²Department of Computer Science Engineering Shri Ram Murti Smarak College of Engineering and Technology

ABSTRACT

In this paper there is an introduction of an algorithm which converts a camera, which is known as an artificial eye, into an intelligent eye. Here the use of deep learning algorithm which is most commonly used in the field of artificial intelligence. In this given paper there is an embedded system which detects the presence of a human being and recorded the video at the required time and when nobody is in the range of camera then the recording system will automatically OFF. Raspberry pi3 is used as a processing block of the system and there is an embedded algorithm which helps the camera to identifying the human presence and also identify him. This algorithm is using the probability concept based on pLSA technique.

Keywords: Deep Learning Algorithm, Artificial Intelligence, pLSA.

I. INTRODUCTION

In modern era as the technologies increases, facilities increases, and network increases simultaneously there is also an increase in crime in the same ratio. There are a lot more new techniques are available to execute a criminal activity. So there is a need of smart eye which is available on every street, road, public places and homes are also needs to be protected from that. This smart eye is nothing but a camera which is placed everywhere now a days. These cameras recorded the data and it would be useful at the time of need. But there is a problem related to the storage of data it means until the camera is on it will record every data which is present in that area or in the area which is in its range. So for removing this issue there is a system proposed which only records the data until the objects which are available in the range are moving when there is a time at which the objects stops moving the camera will stopped recording. This will be executed for streets and roads but when homes

are coming in this range there is an additional issue is comes around that is security, and privacy.

As there is known that the privacy concern is very important for homes so there should be an identical system which detects the presence of the persons in home create a database with the time and date entry and when any person leaves the home that will also be noted but the other private talks and basic daily works which are not so much important are not recorded.

So here is a design of an embedded system [1] which has certain algorithms of deep learning [2]-[4] system seems something close to Artificial Intelligence [5]-[6]. It detects the persons face present in a home record that and when it come again in a certain time range it will not record the content when it detects any new face again it records that one and process for the same.

Here the embedded system is used for the ON and OFF mechanism of the recording system. The deep

learning algorithm used for identifying the known face or an unknown face and respond according to the result to the microcontroller who turns ON or OFF the recording of that camera. That algorithm is uploaded in the form of an operating system in a processing system known as raspberry pi.

Here is a block diagram shows in figure 1 which has the complete process. This complete system is divided in three parts one is caption from camera next is the Algorithm which matches the available result and the last one which allows the camera to record the frames.

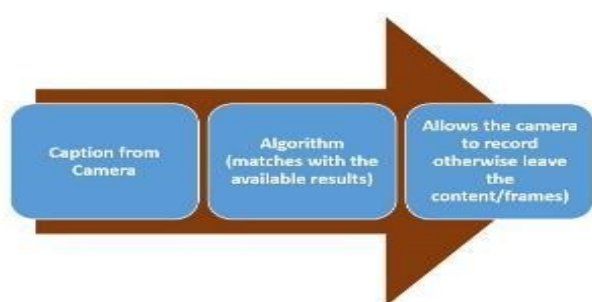


Figure 1. Block Diagram of the system

This system will be helpful in the privacy concerns of the people and also the security issues which are also quite important in these days. Both the issues can be managed simultaneously.

II. DEEP LEARNING ALGORITHM

In this paper as there is a need of identifying the face of a human being. If we concentrate on normal face recognition system then there is a problem in differentiating when the position of the face is changed. There is a possibility that may be once the subject is showing his left portion of face after that he will show his front face or move his face in different style. When these types of problems are arrived there is a solution available which is known as machine learning. Machine learning is a part of AI, here a very small portion of the algorithm, which is used in machine learning, known as Deep Learning Algorithm (DLA).

DLA is used in cascading of various layers in form of nonlinear processing units use for features extracting and also transformation. This algorithm is based on unsupervised learning of representation or these multiple levels of data representation. They are also a part of broader machine learning of the representation of data and learning various multiple levels of the representation which corresponds to multiple levels of abstraction (levels received from the hierarchy of given concepts).

In this algorithm a network is being created which is known as a deep belief network (DBN) [7] is a generative model and made up of various layers of embedded units, it is probabilistic model.

A. Probabilistic Latent Semantic Analysis

It is a technique from category of topic models. Main goal of this model is to co-occurrence of information under the probabilistic framework for discovering underlying structure of data. This pLSA technique was developed in 1999 by scientist Th. Hofmann [8] and used for text application in early days. After that this was shortly spread in various fields such as audio processing [9] and computer vision [10]-[12]. The structure of the pLSA model is showing intermediate layers of the latent topic which links documents and word. Every document is being represented as the mixture of weighted with concepts by different probabilities $P(z|d)$ and every word expressing a topic with probability $P(w|z)$.

B. Algorithm for Image Identification

There are various different phases given to a system for learning and recognition. Here is a model of an image which is a collection of the local patches. Every patch represents a code from large vocabulary of various codes. Goal of this learning system is to achieve the model which representing distributions of the codes in every category of scene.

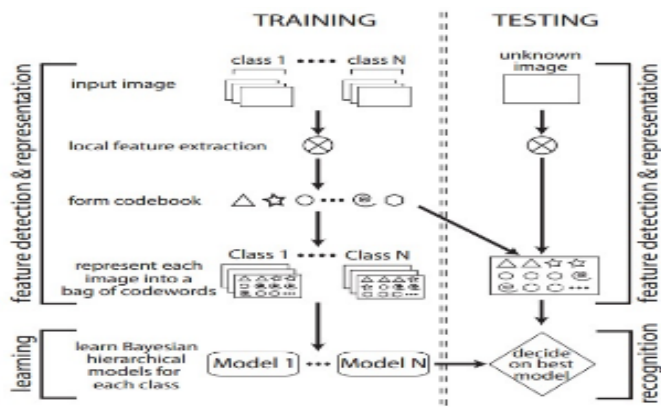


Figure 2. Algorithm for Image Identification

- ✓ Given algorithm is providing a principle approach for learning a relevant representation of scenes with and without supervision.
- ✓ Its approach is a principle probabilistic model for learning the texture via codes [13]-[15].
- ✓ This model is able in categorization of group images in the sensible hierarchy.

III. SYSTEM ARCHITECTURE

The architecture of the system can be classified in two sections. These are as:

- ✓ Intelligent Unit
- ✓ Embedded Unit

A. Intelligent Unit

This unit contains all the data retrieving section which helps in identifying the subject and gives the correct output. It has basically three sections one is raw input vector, second is algorithm section and the last one is output part.

Input can be any input picture which is captured by the camera and after capturing it processed by the above algorithm and there are lots more probabilities given for a single picture so after identifying the correct one the result will be considered by the embedded unit that whether it is to be stored or not.

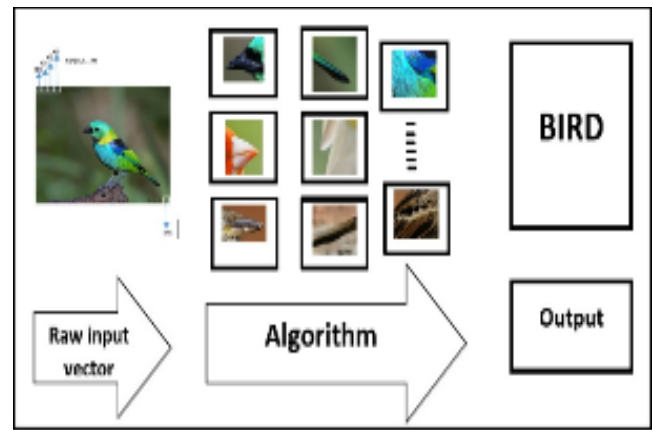


Figure 3. Intelligent Unit

As the above figure 3 showing the complete processing block of intelligent unit here is an example of a bird in which first it is divided into pixels after that it predicts the complete identity. This processing is being done with the help of above given deep learning and pLSA algorithms. In our practical there is a use of human picture as shown in figure 4.

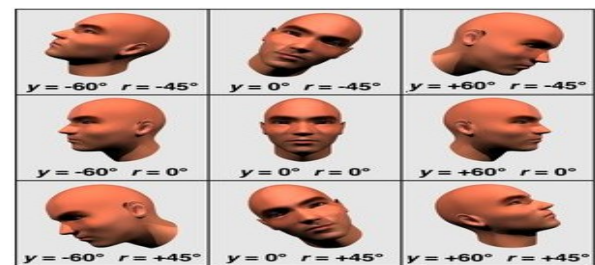


Figure 4. Human face positions

B. Embedded Unit

This unit gives the information to the controller block which forward the task that whether the camera should record the video. There is a complete system which identifies that whether the human face [21]-[22] which is available in the region is recorded or not if response would be negative it sends a signal that it should be recorded and the recording unit will be active.

The figure 5 which is shown below is showing the step by step response of how the controller receives a signal by the intelligent unit and sends to the controller and further the controller which is

controlling the recording unit [16] ask the system whether to record or not.



Figure 5. Embedded Unit

Embedded unit further contains a controller unit in this system we are using a Raspberry pi3 board as a processing unit. In this we uploaded the algorithm in form of an operating system which operates the complete task.

● Raspberry pi3

It is a series of single board computers. In this system the program, which is based on above given algorithm, is uploaded on the board as the operating system uploaded on the system. This uploaded program will identify the person who is present in the range of camera with an image clearing unit [20] and send the priority bit to the recording unit that whether it is to be recorded or not. Its general purpose input output pins are used for this sensing system [17]-[19].

IV. RESULT ANALYSIS AND DISCUSSION

After the complete analysis of the established setup we have recorded the results which we want. During the testing we place the camera in the classroom and detect the faces coming inside it and when the new face is passing the range it will record it with the time. The pictures which are recorded are shown in the figure below.



Figure 6. Caption from Intelligent Artificial Eye

This result shows that when this feature was not included in the cameras which are placed in homes, they take a lot more recording space as this system takes only 0.46 GB space and the normal system will take 3.99 GB in memory.



Figure 7. Caption of Artificial Eye

There is a survey shown in chart 1 which a relation shown between recording of data at normal time and in the recording by the designed system.

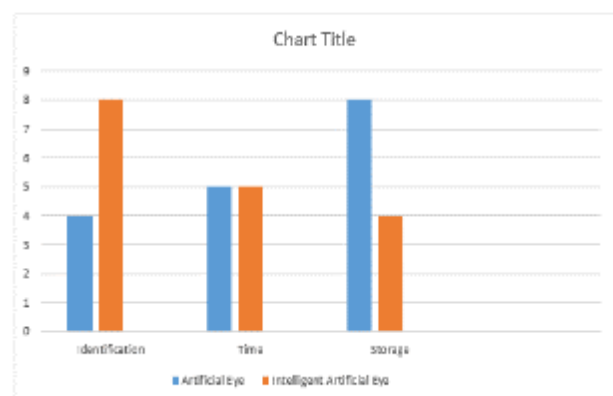


Chart 1. Survey Result

V. CONCLUSION

Here a proposal of a system in which there is an embedded AI Unit with a camera and a recording unit with a controller set. Controller unit decides whether the images should be recorded or not. It decides the priority on the basis of some set of images which are already stored and repeated frequently. It will not record the data when only prior set images or frames are continuously coming. Recording will be active when an unknown person is present in the range of the camera by this the storage is also reduced. There is a frequent database will be created in the system automatically with the time and name of the person, who were available at that place.

In normal system there would be no possibility of this much data storage and that was a system which records the data continuously but when any of the object is not moving in the range of the camera that time these system creates a lot more trash data. This given system solves the problem of trash data and also identify the person available in the range. But the design of the system is more complex than the traditional system.

VI. FUTURE SCOPE

The proposed system enables many doors for introducing AI with embedded applications in terms of security purpose or other identification methods. Now a days we are using AI in phone camera's for detection and in social media where the automatic face detection is available. When this will introduced with embedded system there are more ideas which are going to be implemented in this research field as well as in practical field. This system can be explored by adding a database of the known person identity and also the data can be updated by attaching it with cloud. If we create a cloud server where the identity of every person is uploaded who are present in the range, the system would become more precise and accurate.

VII. ACKNOWLEDGEMENT

The authors are thankful to Hon'ble C – VI, Mr. Aseem Chauhan (Additional President, RBEF and Chancellor AUR, Jaipur), Maj. General K. K. Ohri (AVSM, Retd.) Pro-VC Amity University, Uttar Pradesh Lucknow, Wg. Cdr. Dr. Anil Kumar, Retd.(Director, ASET), Prof. S. T. H. Abidi (Professor Emeritus), Brig. U. K. Chopra, Retd.(Director AIIT), Prof. H K Dwivedi (Director, ASAP), Prof O. P. Singh (HOD, Electrical & Electronics Engg.) and Prof. N. Ram (Dy. Director ASET) for their motivation, kind cooperation, and suggestions.

VIII. REFERENCES

- [1]. https://www.tutorialspoint.com/embedded_systems/es_overview.htm
- [2]. Y. Bengio, H. Schwenk, J.-S. Senecal, F. Morin, and J.-L. Gauvain. Neural probabilistic language models. In *Innovations in Machine Learning*. Springer, 2006
- [3]. Ian Goodfellow, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. MIT Press. Online
- [4]. Bengio, Y.; Courville, A.; Vincent, P. (2013). "Representation Learning: A Review and New Perspectives". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35 (8): 1798-1828. arXiv:1206.5538. doi:10.1109/tpami.2013.50
- [5]. Hutter, Marcus (2005). *Universal Artificial Intelligence*. Berlin: Springer. ISBN 978-3-540-22139-5.
- [6]. Neapolitan, Richard; Jiang, Xia (2012). *Contemporary Artificial Intelligence*. Chapman & Hall/CRC. ISBN 978-1-4398-4469-4.
- [7]. https://en.wikipedia.org/wiki/Deep_learning
- [8]. Thomas Hofmann. Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information*

- retrieval, SIGIR '99, pages 50-57, New York, NY, USA, 1999. ACM.
- [9]. Matthew D. Hoffman, David M. Blei, and Perry R. Cook. Finding latent sources in recorded music with a shift-invariant hdp. In International Conference on Digital Audio Effects (DAFx), 2009.
- [10]. Florent Monay and Daniel Gatica-Perez. PLSA-based image auto-annotation: constraining the latent space. In Proceedings of the 12th annual ACM international conference on Multimedia, MULTIMEDIA '04, pages 348-351, New York, NY, USA, 2004. ACM.
- [11]. J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and W. T. Freeman. Discovering object categories in image collections. In Proceedings of the International Conference on Computer Vision, 2005.
- [12]. Rainer Lienhart, Stefan Romberg, and Eva Horster. Multilayer pLSA for multimodal image retrieval. In Proceeding of the ACM International Conference on Image and Video Retrieval, CIVR '09, pages 9:1-9:8, New York, NY, USA, 2009. ACM.
- [13]. T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *IJCV*, 43(1):29-44, June 2001.
- [14]. M. Varma and A. Zisserman. Texture classification: are filter banks necessary? In *CVPR03*, pages II: 691-698, 2003.
- [15]. J. Portilla and E.P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *IJCV*, 40(1):49-70, October 2000.
- [16]. M. Szummer and R. Picard. Indoor-outdoor image classification. In *Int. Workshop on Content-based Access of Image and Video Databases*, Bombay, India, 1998.
- [17]. S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381:520-522, 1996. 13] A. Torralba, K. Murphy, and W.T. Freeman. Sharing features: efficient boosting procedures for multiclass object detection. In *Proc. of the 2004 IEEE CVPR*. 2004.
- [18]. A. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12:97-136, 1980.
- [19]. A. Vailaya, M. Figueiredo, A. Jain, and H. Zhang. Image classification for content-based indexing. *IEEE Trans. on Image Processing*. 10, 2001.
- [20]. M. Varma and A. Zisserman. Texture classification: are filter banks necessary? In *CVPR03*, pages II: 691-698, 2003.
- [21]. J. Vogel and B. Schiele. A semantic typicality measure for natural scene categorization. In *DAGM'04 Annual Pattern Recognition Symposium*, Tuebingen, Germany, 2004.
- [22]. J. Winn. Variational Message Passing and its applications. PhD thesis, University of Cambridge, 2003.