Deep Learning from Human Experience to Enhance the Performance for Robotics

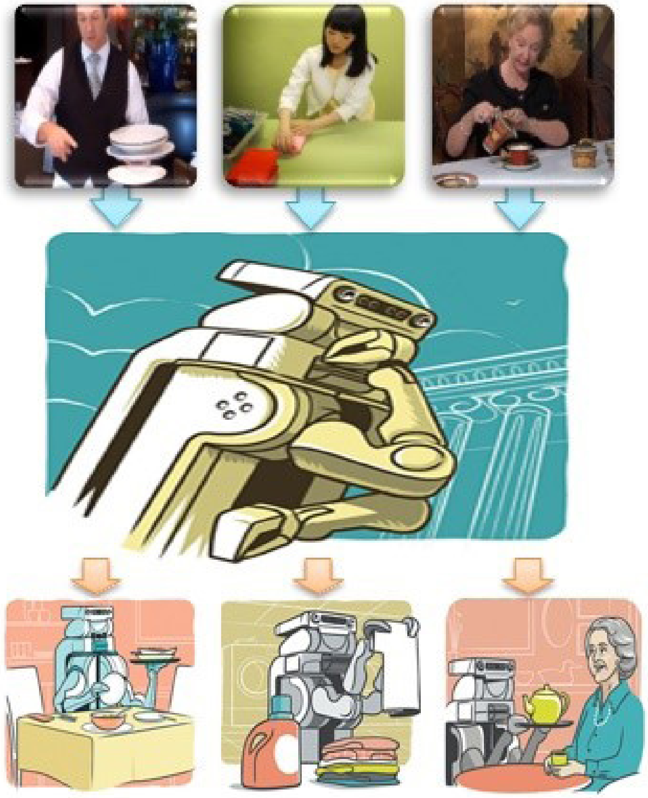
LogDynamics
International Graduate School

Humans are able to achieve a high level of performance while performing complicated tasks based on their experience. Experience is a kind of knowledge that we recall from the stored memory. It is the collection of past personal activities that occurred at a particular time and place, also known as Episodic Memories (EM). We achieve perfection in our task from our experience of knowledge. To obtain experience for any particular task human goes through observation over the time until we get the satisfying result. When we reach a certain level of performance we also gain a high level of experience for that particular task stored as episodic memory. This experience of knowledge is currently not reflected in the intelligence of robotic agents. As a result, they fail to reach human like perfection even when they are built with sophisticated hardware to execute tasks very precisely. It will make a great impact while transferring human experience from episodic memory to robotic agents. There is a huge collection of videos available on the internet which reflects the different activities perfectly compiled by human expertise. The challenge remains on analyzing these videos to extract knowledge as experience and let the robot learn from it. This transformation will allow the robotic agents to reach human-like performance by saving the time that humans spend through complicated evolving procedure to achieve a certain level of experience.

Knowledge Generalization Problems

However the main challenges remain on extracting the generalized knowledge from different types of episodic memory to learn a particular task. The same task can be performed by different experts in different ways. It is quite complicated to learn in a general way from different examples of episodic memories that reflect the same task with different solutions. For example, there are many videos for making pancakes. However, the way of making a pancake is not the same in all videos. There are differences in how humans perform a specific action, how they grasp or use tools and even the order of actions to perform a task.

The challenge is, by observing all different types of videos for making pancakes, how the system extracts general information about how to make a pancake.



Learning from Episodic Memory

A knowledge-based learning model will be developed to generate relational information for robotic agents to perform particular tasks. The model will learn human activities by analyzing the videos referred as episodic memory containing different human expertise activities. Learning human experience from videos will make the robots more proficient in uncertain environments and improve the overall performance. The episodic memory can be thought of as a

Feroz Ahmed Siddiky
M.Sc.

Mathematics /
Computer Science

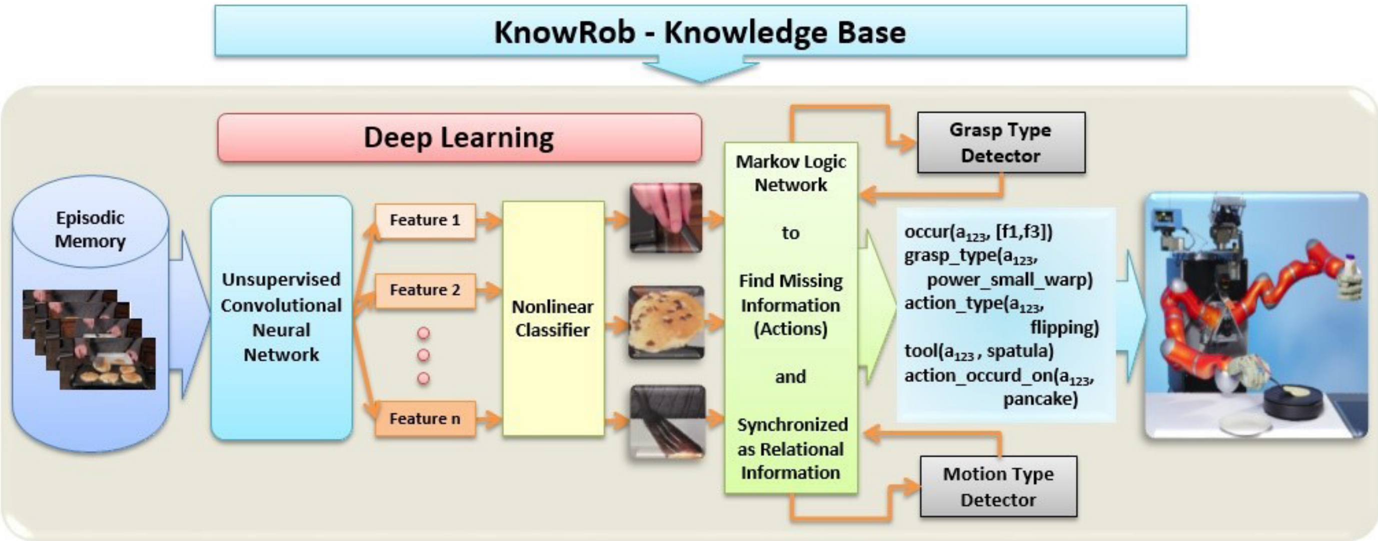
Dhaka, Bangladesh
siddiky@cs.uni-bremen.de

International Graduate School for Dynamics in Logistics
Contact: Dr.-Ing. Ingrid Rügge

Universität Bremen
Hochschulring 20
28359 Bremen, Germany

www.logistics-gs.uni-bremen.de
info@IGS.Log*Dynamics*.de

University of Bremen



process with several different steps, each of which relies on a separate system of the brain. The recollection of experiences is contingent on three steps of memory processing: encoding, storage and retrieval. The designed model will be more or less correlated with these three stages. The videos will investigate in depth level to extract relational features where higher level features are defined in terms of lower level features, detect local objects, find relations between the objects to perform a particular action and finally generate relational information for the robot to learn the particular kind of job.

The proposed model will consider two subsequence of steps to process the information:

- Lower-Level Processing: extract raw features from video analysis
- Higher-Level Processing: receive uncertain raw feature data from lower-level to generate missing information and organize as relational information for particular action plan processing

To extract information from the lower level an unsupervised learning model will be developed by inheriting the 2D layer concept from well-known Convolutional Neural Network (CNN). Once the invariant features have been learned successfully with unlabeled data, a nonlinear classifier can use these features to classify objects through supervised learning system. The information from the lower level is not adequate and consistent for visual event recognition.

A possible solution would be integrating extensive collections of common sense knowledge into robots' knowledge bases to enable the efficient infer control decisions under various undetermined environmental conditions. A probabilistic first order logic based graphical model such as Markov Logic Network (MLN) will be used with knowledge base to observe the missing information and uncertainty among extracted features from the lower level and generate synchronized information for various robotic events.

Comprehensive Learning Model for Dynamic Real World Environment

Learning from episodic memory for robotic agents will not be limited to a certain environment. It will enhance invariant learning. The comprehensive learning model will provide a solution not only for a particular environment, but also for dynamic environment like warehouse unloading and mixed-case depalletizing or autonomous vehicles for transportation in logistics. For this, it is important to obtain a sufficient and efficient knowledge base for the procedure of invariant event recognition. Integrating openEASE, a cloud-based knowledge base of robot experience data could provide the sufficient information for the comprehensive learning model on dynamic environment. The knowledge from openEASE will provide sufficient information to understand the environment, possible location of the particular objects for a particular task and infer preciously when what type of actions have to be performed.