Monitoring and Maintaining Student Online Classroom Participation Using Cobots, Edge Intelligence, Virtual Reality, and Artificial Ethnographies

Ana Djuric[1], Meina Zhu[1], Weisong Shi[1], Thomas Palazzolo[1], Robert G. Reynolds[1]
[1] Wayne State University, Detroit MI 48202, USA

Abstract

In this project Virtual World technology and Edge Intelligence to produce a shared social landscape for the society of learners. The idea is to create a Virtual World in which learners can participate and interact. One that is parallel to the learning environment or classroom. This can be viewed as an online multi-user environment such as "second-life" where on-line learners can interact and construct their own spaces. Their ability to work in that space is governed by input from their robot mentor (Human Robot Learning Unit). Skills in the Classroom Virtual World are provided as a result of a student's behavior in the learning environment. The Virtual World can persist after the learning session is concluded so it provided an incentive for learners to do well in the learning session so that they can acquire points that translate into skills in the corresponding Virtual World. That Virtual World can be shared by several learning sessions or classes to provide a more comprehensive learning environment. An online ethnography of the interactions of learners and instructors can be produced as suggested by McCarthy and Wright (3).

Keywords: Cobots, Human-Robet Learning Units, Edge Computing, Artificial Ethnographies, Virtual World, Learner Focus.

Motivation and Vision

A Cobot is a robot intended for direct human interaction within a shared space. Unlike traditional industrial robots that whose actions are isolated from their human counterparts. [1]. Cobots were invented in 1994 by J. Edward Colgate and Michael [8]. Cobots can be used in variety of situations including public spaces Plishkin providing informational services. [5]. This is the context in which we view them here. The International Federation of Robotics [4] has identified 4 different categories of Cobots [7]:

1. Coexistence: The human and the robot work along each other with a partition but have no shared workspace.
2. Sequential collaboration: The human and the robot are both active within a shared workspace but their actions are sequential and they don't work at the same time.
3. Cooperation: The human and robot work on the same task at the same and are both in motion.
4. Responsive collaboration: The robot responds in real-time to the actions of its human counterpart.

It is this latter category that is of concern here. This project is concerned with the development of a Human Robot team (Human Robot Learning Unit) that is able to participate in a society of online learners. The motivation behind this that one way to maintain a learner's attention is to have a "paraprofessional" monitor their activity online. However, it is difficult for a single human to closely monitor a large group of learners especially since individuals have different learning styles and learning rates. In addition, a learner can simply turn off their audio and visual and fly under the radar. The classic case is where a student thought that they had switched off their audio and video so the observer was able to see them playing video games in the background the entire session.

In this project the use of Virtual World technology and Artificial Intelligence to produce a shared social landscape for the society of learners. The idea is to create a Virtual World Classroom in which learners can participate and interact. One that is extension of the learning environment or classroom. This can be viewed as an online multi-user environment such as "second-life" where on-line learners can interact and construct their own spaces. Their ability to work in that space is governed by input from their robot mentor. Skills in the Virtual World are provided as a result of a student's behavior in the learning environment. The Virtual World can persist after the learning session is concluded so it provided an incentive for learners to do well in the learning session so that they can acquire points that translate into skills in the corresponding Virtual World. That Virtual World can be shared by several learning sessions or classes to provide a more comprehensive learning environment. The experiences can be combined to produce an online

In T. Kido, K. Takadama (Eds.), Proceedings of the AAAI 2022 Spring Symposium "How Fair is Fair? Achieving Wellbeing AI", Stanford University, Palo Alto, California, USA, March 21–23, 2022.

ethnography similar to that generated by McCarthy and Wright for the online game "Second Life" [3]. They proposed a four-part framework through which to interpret the users' subjective experiences:

1. The impact of the experience on the senses. The experiences concrete and visceral impact.
2. The emotional and affective impact of the experience.
3. The compositional of the sequence of actions that comprise an event.
4. The spatial and temporal context of the experience.

Although there are many dimensions to the learning activity that can be studied the system described here addresses the most fundamental aspect of learning, how a learner maintains focus in their environment. Other qualities can be added in down the road. The challenges that online learners face in terms of focus will be discussed in the next section.

Challenges to the Focus of Online Learners

The loss of online students' attention to learning is a common and severe problem. Due to the COVID-19 pandemic, more than 200 million students, consisting of 12.5% of total enrolled students worldwide, were influenced by the university and school closures in December 2020 [1]. It is clear that the Pandemic has accelerated the process of shifting courses from a traditional face-to-face format to an online one [2]. Online learning offers students more choices and flexibility in necessary coursework, which requires increased skills to plan, monitor, and manage learning [4] [5]. However, online education is challenging for both students and teachers. The loss of focus of attention and engagement in online learning is one of the primary challenges of online education [6]. Given that attention comes prior to cognitive learning, staying focused and engaged is vital to cognitive learning activities [7]. Losing focus affects lectures, labs, tests, quizzes, group activities, and projects in online education (see Figure 1).

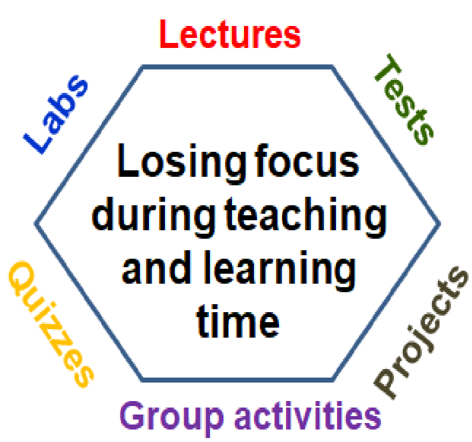

Fig. 1. Losing focus affects online learning activities

The factors that can possibly lead to losing focus of attention can be categorized into the external state (see Figure 2) and internal state (see Figure 3). The students' external state reflects the impact that their learning environment has on their cognition. engagement, tiredness, overload, loneliness, and lack of communication with classmates and instructors [8] (see Figure 3).

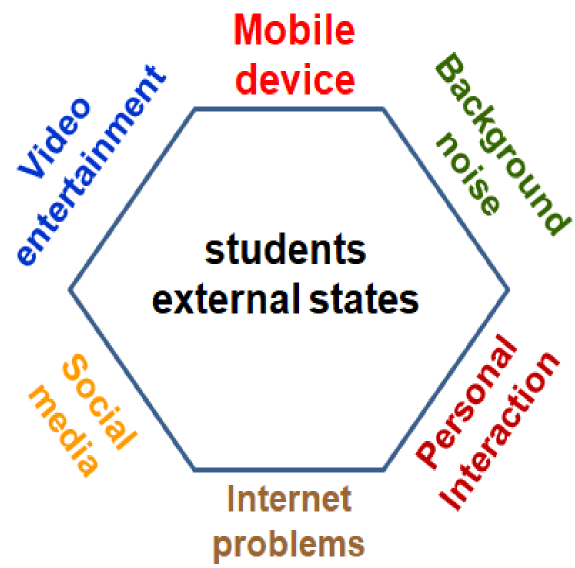

Fig. 2. External state

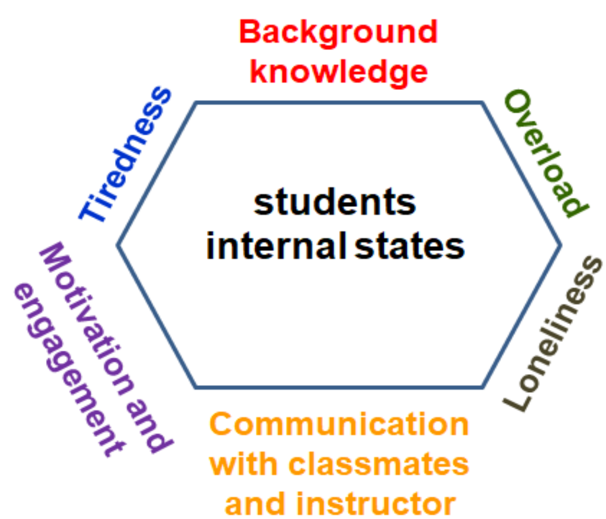

Fig. 3. The Internal state.

Robotic technologies have played a significant role in education. Research has indicated that online pedagogical agents can promote effective instruction [9] [10]. For example, robots have taken diverse roles in education, such as addressing absenteeism [11], enhancing motivation [12], supporting students' emotions [13], triggering productive conversation in language education [14], promoting collaboration [15] [16], fostering computational thinking [17] [18], and enhancing creative thinking and problem-solving skills [19]. However, a majority of the agents were virtual robotics or physical robotics for classroom teaching. Little research has focused on the use of physical robots as participants in an online students' learning environment. In other words, each student would have a robot mentor that will help monitor the student's progress and provide feedback to the instructor. The instructor can then use that information at a meta-level to make strategic decisions about class trajectories.

The vision of this project is to exploit the synergistic potential of the robot student team. That is, humans can perform certain tasks better than robots and vice versa. The goal is to exploit the complementarity nature of their relationship in order to produce a true marriage of minds. This Human-Robot-Learning-Unit (HRLU) is the fundamental building block upon which to scaffold a new framework for online learning. In the next section the basic structure of the HRLU will be discussed along with the information that can be passed to the Supervisor. The Supervisor will then use that information to update the Virtual World based on learner's performances and update their ethnography. The updated ethnography will be the basis for adjusting the HRLU components for the next learning session.

HRLU Methodology

Robots are used as teaching and learning tools to be manipulated and operated by students in many schools. For online teaching, the robot assistants will be located at online learner's homes. Because of that, we made a comparison between different teaching robots based on their suitability for such an application. The factors that are compared include their functionality, price, weight, software, hardware, etc. Therefore, this research uses a robot, like Misty (https://www.mistyrobotics.com/), to facilitate students' self-regulation in the online learning environment (see Figure 4).

Fig. 4. Misty robot (https://www.mistyrobotics.com/)

The Robots contribution to the HRLU can be as follows:

5. First, robots can provide pre-scheduled learning activities during the entire semester in order to support students' time-management.
6. Second, the robots can monitor the students' learning behavior through eye-tracking and monitoring facial expressions and gestures during synchronous classroom and related meeting sessions. Based upon learned patterns in the students' behavioral data, the robot can track students' learning progress and provide interventions to facilitate students' cognition and meta-cognition.
7. Third, robots can facilitate formative assessment and provide immediate feedback to students in online learning.
8. Fourth, robots can communicate not only with students but also with the Supervisor. The Superviso Unit (ILRU) will facilitate communication between the HRLUs and with the Virtual World.

The HRLUs communicate in the Virtual Classroom with other HRLUs. See Figure 5. The communication will be arranged such that each student is communicating with a personalized robot, while all robots are communicating in the network, and the instructor (IRLU) is communicating with all robots in the network. It is possible that the Instructor will have their own intelligent agent learning unit.

In order to control the online HRLU classroom, instructor(s) (IRLU) will be using provide scripts for interactions (e.g., questions) prepared using their previous teaching experience. See Figure 5. The control flow can be as following:

1. Input - scripted interactions that are designed to get information about the students' internal state)
2. Output - Collecting answers from students
3. Output - Analysis of student's answers

4. The Supervisor (IRLU) updates the Virtual World parameters based upon the student robot interactions. The Virtual World is referred to as the Virtual Classroom Matrix in Figure 5 as a reference to the "Matrix" in the corresponding films
5. Data Analytics of the updated VCM in order are performed by the IRLU to adjust the state of the Virtual World.
6. The Supervisor ILRU updates the Ethnography Classroom Matrix (ECM) of the Virtual World using the adjusted VR parameters from 5 above.
7. Express ECM and VCM parameters in a graphical update using a GUI. This GUI will be used for generating a virtual classroom map using Machine Learning techniques such as Evolutionary and Deep Learning. The interface represents an indicator for controlling students' focus of attention. The instructor(s) will use this display to improve students' self-regulation skills, motivation, and learning outcomes.
8. Calculate the error between expectations and outcomes in order to produce new scripts for the HRLUs and repeat the cycle.

This two tiered framework is ideally suited for an Edge Computing framework How the framework can be used to support the workflow above will be the subject of the next section.

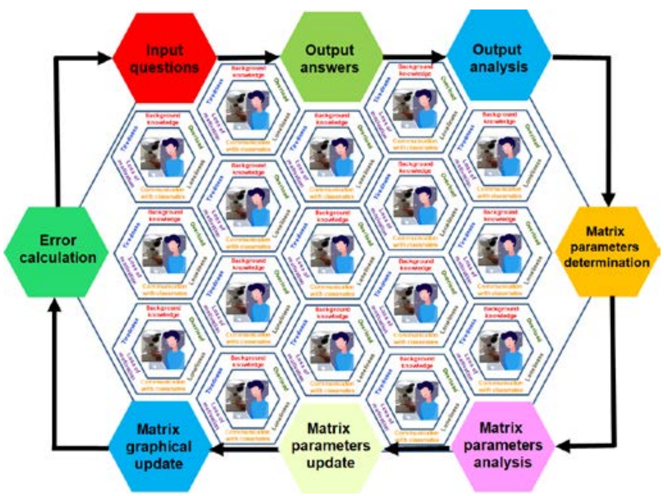

Fig. 5. Graphical representation of the dynamic virtual classroom matrix (VCM).

Using Edge intelligence to Support the HRLU and IRLU Cycle.

Capturing students' real-time learning status is vital to effective online learning. Sensor technology can objectively gather students' learning behaviors. Prior research and educators (Daniel & Kamioka, 2017; Hwang et al., 2011; Krithika & GG, 2016; Su et al., 2014; Sharma et al., 2019) have utilized sensor technology to capture students' behavior, including eye movement, facial expressions, and body movement. Through students' learning behavior, we can detect and indicate to what extent students stay focused on online learning scenarios. Prior research was primarily focused on traditional face-to-face education settings or capture the video data only. For intelligent agent-based approaches, prior studies used to train one single model and deploy it for all users without considering the personalized factors. In the early detection phrase in our system, we move forward to include two factors that are usually neglected by the community: one is environmental noise, which is a passive factor that can affect the concentration; another is personalized behavior, as different students will demonstrate different distraction behavior and expression. To this end, we propose a cloud-edge collaborative system to provide personalized detection based on multi-dimensional data. We jointly combine video and audio data for Focus Index (FI) detection. Our pro system encapsulates detection objects in module units and provides APIs for third-party integration. Beyond that, we propose the idea to leverage edge intelligence for personalized model training and serving.

Edge computing (Shi et al., 2016) has become the most popular computing paradigm with the development of the Internet of Things and other devices located at the edge of the network. Statistics show that these devices will generate 60% of the data in the future, reaching PB level data volume. One typical data generation scenario is HRLU, where cameras are highly used to help detect the distraction degree of one student. Each camera generates a considerable volume of video data every day (in GB level). In cloud computing, all the video data has to be sent to the cloud for processing, which poses considerable pressure on the bandwidth and workload of the data center. Edge computing can offload data from the cloud to the process units near the data source or even offload tasks to the camera itself.

There are two main factors that inspire us to leverage edge computing in HRLU: (a) Large data volume. Uploading all the generated data to the cloud is impossible and is also a waste of bandwidth, transmission resources, and cloud storage resources. Edge computing can help to pre-process and filter the valuable data before sending it to the cloud for centralized control or offload the whole task. (b) Reliable performance. The distraction of students is expected to be detected in a timely fashion. If the detection relies on cloud processing, its performance will be affected by many uncertainties: network connection, data center status, to name a few. Especially when online learning already takes a considerable bandwidth, edge computing is more reliable to guarantee near-real-time processing with capable hardware equipped.

Artificial Intelligence (AI) has been greatly developed in this decade thanks to hardware development. The convolutional neural network (CNN) promotes the development of Computer vision (Krizhevsky et al., 2017), and the Transformer network promotes the development of Natural Language Processing (Vaswani et al., 2017). **Spoken** language processing is also accelerating its momentum with deep neural networks (Amodei et al., 2016). AI-related services usually rely on the computation resources on the cloud to provide service. Recently, with the development of lightweight AI models, edge-oriented hardware and software, edge devices, and platforms gain the capability to execute AI algorithms, i.e., Edge Intelligence (Zhang et al., 2019).

Edge intelligence not only inherits the advantages from edge computing, where offloading the processing from the cloud; it also brings intelligence to the edge devices and demonstrates a huge potentiality to serve the real world. In the HRLU, we propose an edge intelligence system for the robot, which is designed to detect the student's state and intervene when necessary for online learning. Considering the functionality of the robot, which is equipped with a microphone array, 4K camera, HIFI speakers, it is capable of capturing input data in different dimensions and deploying different types of AI models to make decisions jointly.

The following section describes how the Focus detection HRLU prototype can be expressed in terms of the Edge Computation Environment.

HRLU System Design on the Edge.

To quantify the distraction degree of the students, we develop a **Focus Index (FI)** to represent the focus degree of a student that ranges from 0 to 100. This score is translated into points that can be used by the learners in order to participate in the Virtual World Classroom. The points can be exchanged for tools and objects that allow them to interact with others in the Virtual Classroom.

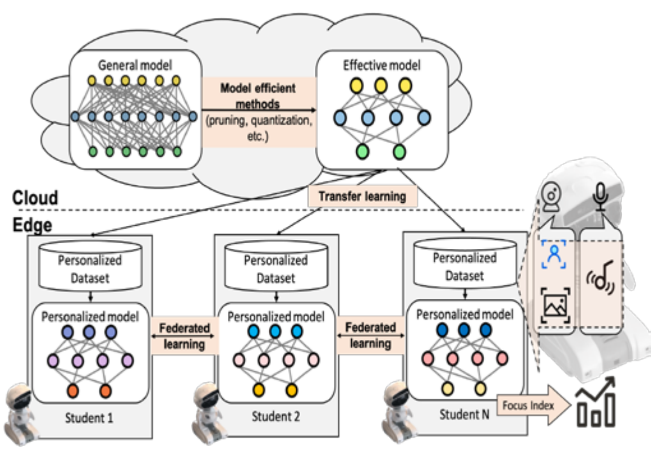

Figure 6. Cloud-edge collaborated early detection system.

The early detection system is presented in Figure 6. It is a cloud-edge collaborative system for FI prediction based on personalized multi-dimensional data. To provide a reliable and solid detection for a valid intervention, the cloud is responsible for training a general detection model with a large amount of labeled data. The cloud collects the video and audio data in oder to obtain the students' focus information and predict FI scores based on the trained detection model. Considering the scale of the dataset, the intelligent model generated by the cloud will be expensive for edges to compute and store. To fit the developed intelligent model to resource-constraint edge nodes, some model efficiency methods will be taken (Han et al., 2015a). For example, model pruning (Han et al., 2015b), quantization (Gong et al., 2014), knowledge distillation (Hinton et al., 2015), network architecture search (Cai et al., 2018) can all contribute to effective pruning of the model. The processed efficient F1 evaluation model is then deployed on each robot through transfer learning. With the built-in camera and microphone array, each robot can capture video and audio as the input of the efficient model to compute the FI for the students and assessed. Every so often the models performance in F1 detection can be assessed and the data used to update the model in the cloud.

Conclusion

In this paper the use of Virtual World technology and Artificial Intelligence are employed to produce a shared social landscape for the society of learners. The idea is to create a Virtual Classroom World in which learners can participate and interact. One that is parallel to the learning environment or classroom. This can be viewed as an online multi-user environment such as "second-life" where on-line learners can interact and construct their own spaces. Their ability to work in that space is governed by input from their robot mentor. Skills in the Virtual World are provided as a result of a student's behavior in the learning environment. The Virtual World can persist after the learning session is concluded so it provided an incentive for learners to do well in the learning session so that they can acquire points that translate into skills in the corresponding Virtual World. That Virtual World can be shared by several learning sessions or classes to provide a more comprehensive learning environment. This shared experience can be documented as an online ethnography of the Virtual Classroom.

References

1. UNESCO, "COVID-19 Educational Disruption and Response," 22 December 2020. [Online]. Available:

https://en.unesco.org/themes/education-emergencies/coronavirus-school-closures.

2. L. Gardener, "COVID-19 Has Forced Higher Edu to Pivot to Online Learning. Here Are 7 Takeaways So Far," *The Chronicles of Higher Education,* 2020.
3. J. McCarthy and P. Wright, "Technology as Experience," *Interactions,* vol. 11, no. 5, pp. 42-43, 2004.
4. M. Ally, "Foundations of Educational Theory for Online Learning," in *Theory and Practice of Online Learning*, Athabasca, Athabasca University, 2004.
5. J. C. Y. Sun and R. Rueda, "Situational Interest, Computer Self-Efficacy and Self-Regulation: Their Impact on Student Engagement in Distance Education," *British Journal of Educational Technology,* pp. 191-204, 2021.
6. J. Y. Wu, "The Indirect Relationship of Media Multitasking Self-Efficacy on Learning Performance Within the Personal Learning Environment: Implications from the Mechanisms of Perceived Attention Problems and Self-Regulation Strategies," *Computers & Education,* pp. 56-72, 2017.
7. S. E. Petersen and M. I. Posner, "The Attention System of the Human Brain: 20 Years After," *Annual Review of Neuroscience,* pp. 73-89, 2012.
8. Y. C. Kuo, A. E. Walker, K. E. Schroder and B. R. Belland, "Interaction, Internet Self-Efficacy, and Self-Regulated Learning as Predictors of Student Satisfaction in Online Education Courses," *The Internet and Higher Education,* pp. 35-50, 2014.
9. B. Heller and M. Procter, "Embodied and Embedded Intelligence: Actor Agents on Virtual Stages," *Intelligent and Adaptive Learning Systems: Technology Enchanged Support for Students and Teachers,* pp. 280-292, 2012.
10. A. Pereira, C. Martinho, I. Leite and A. Paiva, "iCat, The Chess Player: The Influence of Embodiment in the Enjoyment of a Game.," *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems,* pp. 1253-1256, 2008.
11. M. A. Mac Iver and D. J. Mac Iver, ""STEMming" the Swell of Absenteeism in Urban Middle Grade Schools: Impacts of a Summer Robotics Program," *Society for Research on Educational Effectiveness,* 2014.
12. A. Gomoll, C. E. Hmelo-Silver, S. Sabanovic and M. Francisco, "Dragons, Ladybugs, and Softballs: Girls' STEM Engagement with Human-Centered Robotics," *Journal of Science and Education and Technology,* pp. 899-914, 2016.
13. M. Dennins, J. Masthoff and C. Mellish, "Adapting Progress Feedback and Emotional Support to Student Personality," *International Journal of Artificial Intelligence in Education,* pp. 877-931, 2016.
14. S. Tegos, S. Demetriadis and T. Tsiatsos, "A Configurable Conversational Agent to Trigger Students' Productive Dialogue: A Pilot Student in the CALL Domain," *International Journal of Aritificial Intelligence in Education,* pp. 62-91, 2014.
15. W. Y. Hwang and S. Y. Wu, "A Case Studet of Collaboration with Multi-Robots and its Effect on Children's Interaction," *Interactive Learning Environments,* pp. 429-443, 2014.
16. M. Menekse, R. Higashi, C. D. Schunn and E. Baehr, "The Role of Robotics Teams' Collaboration Quality on Team Performance in a Robotics Tournament," *Journal of Engineering Education,* pp. 564-584, 2017.
17. G. Chen, J. Shen, L. Barth-Cohen, S. Jiang, X. Huang and M. Eltoukhy, "Assessing Elementary Students' Computational Thinking in Everyday Reasoning and Robotics Programming," *Computers & Education,* pp. 162-175, 2017.
18. J. Leonard, A. Buss, R. Gamboa, M. Mitchell, O. S. Fashola, T. Hubert and S. Almughyirah, "Using Robotics and Game Design to Enhance Children's Self-Efficacy, STEM Attitudes, and Computational Thinking Skills," *Journal of Science Education and Technology,* pp. 860-876, 2016.
19. E. Z. F. Liu, C. H. Lin, P. Y. Liou, H. C. Feng and H. T. Hou, "An Analysis of Teacher-Student Interaction Patterns in a Robotics Course for Kindergarten Children: A Pilot Student," *Turkish Online Journal of Educational Technology-TOJET,* pp. 9-18, 2013.
20. W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal,* pp. 637-646, 2016.
21. A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Communications of the ACM,* pp. 84-90, 2017.
22. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez and I. Polosukhin, "Attention is All You Need," *Advances in neural Information Processing Systems,* pp. 5998-6008, 2017.
23. D. Amodei, S. Ananthanarayanan, R. Anubhai, J. Bai, E. Battenberg, C. Case and Z. Zhu, "Deep Speech 2: End-to-End Speech Recognition in English and Mandarin," *International Conference on Machine Learning,* pp. 173-182, 2016.
24. X. Zhang, Y. Wang, S. Lu, L. Liu and W. Shi, "OpenEI: An Open Framework for Edge Intelligence," *2019 IEEE 39th International Conference on Distributed Computing Systems,* pp. 1840-1851, 2019.
25. X. Zhang, M. Qiao, L. Liu, Y. Xu and W. Shi, "Collaborative Cloud-Edge Computation for Personalized Driving Behavior Modeling," *Proceedings of the 4th ACM/IEEE Symposium on Edge Computing,* pp. 209-221, 2019.
26. S. Han, J. Pool, J. Tran and W. Dally, "Learning Both Weights and Connections for Efficient Neural Network," *Advances in Neural Information Processing Systems,* pp. 1135-1143, 2015.
27. S. Han and H. D. W. J. Mao, "Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding," 2014. [Online]. Available: https://papers.nips.cc/paper/2015/hash/ae0eb3eed39d2bcef4622b2499a05fe6-Abstract.html.
28. Y. Gong, L. Liu, M. Yang and L. Bourdev, "Compressing Deep Convolutional Networks Using Vector Quantization," 2014. [Online].
29. G. Hinton, O. Vinyals and J. Dean, "Distilling the Knowledge in a Neural Network," 2015. [Online].
30. H. Cai, L. Zhu and S. Han, "Proxylessnas: Direct Neural Architecture Search on Target Task and Hardware," 2018. [Online].
31. X. Zhang, Y. Wang, S. Lu and L. S. W. Liu, "OpenEI: An Open Framework for Edge Intelligence," *2019 IEEE 39th*

International Conference on Distributed Computer Systems,
pp. 1840-1851, 2019.