Dynamic Reconfiguration Using Signal Processing for Artificial Intelligence

Deageon Kim[1], Dongoun Lee[2*]

*[1]Department of Architectural Engineering, Dongseo University,
Republic of Korea, E-mail: gun43@hanmail.net*
*[*2]Department of Architectural Engineering, Dongseo University,
Republic of Korea, E-mail: ldu21@dongseo.ac.kr*

A newly developed meta-surface reconfigurable intelligent surface has been singled out as a top choice for 6G mobile communications networks. By reflecting signals, RIS might effectively install new communication conduits. Meanwhile, it has been widely accepted in recent years that AI may be a useful tool. With the help of artificial intelligence (AI), RIS communications will be more capable of adapting to new situations and more resilient to interference. Artificial intelligence (AI) may tackle contemporary signal processing problems in a data-driven manner by extracting the inherent features from the real data, in addition to the conventional model-driven methodologies. Because of this, AI is well-suited for the digital signal problems through RIS networks under less-than-ideal settings such as modeling mismatch, a shortage of resources, hardware deterioration, and dynamic transmissions. As one of the first survey articles on the topic, we will discuss the application of AI and RIS, or artificial intelligence and radio frequency identification systems, to a wide range of signal processing tasks, including environmental sensing, channel acquisition, directional antenna design, bandwidth allocation, etc. In addition, we'll discuss the challenges of AIRIS and provide some promising future directions.

Keywords: Reconfigurable intelligence; artificial intelligence; deep learning; deep reinforcement learning; Signal processing.

1. Introduction

Applications for signal and image processing need a lot of computational power. Many of these were previously accomplished using digital signal processors (DSP). It offers low-cost flexibility and computational power. However, since the start of the decade, processor core frequency has stopped increasing while power consumption has increased significantly with performance [1, 2]. This power wall causes embedded systems, particularly those used for battery-powered applications, to have uncontrollable temperature and autonomy problems. The application roadmaps nonetheless, call for a significant increase in computer capacity notwithstanding these restrictions. Consumer electronics applications employ the system-on-chip (SoC) technique to get around such restrictions, where customized accelerators deliver DSP and GPP cores to match the system requirements within a constrained power envelope [3]. The primary disadvantage of this strategy is that it is quite costly and results in

complicated systems that seem to be difficult to test. As a result, they are used in applications that need extremely large numbers, such as mobile phone handsets [4]. Applications of professional electronics that are distinguished by a costly SoC solution cannot be used because to low quantities and a high design count compared to consumer ones. It's worth noting that plant and animal studies, particularly in the realm of agricultural and life sciences (AL), face analogous computational challenges. In AL research, tasks such as genomic analysis, phenotyping, and ecological modeling require substantial computational resources [5, 6]. Additionally, machine learning algorithms are increasingly being employed for real-time fraud detection and risk assessment, leveraging the power of signal processing techniques to analyze vast amounts of financial data [7].

To boost compute capabilities and match roadmap needs, they would prefer mix DSP or GPP with FPGA. In contrast to a pure DSP approach, FPGA solutions are static and lack flexibility. However, the growth of roadmaps is moving towards multi-standards or even software-defined applications, necessitating a virtualization layer to allow for dynamic behavior changes or adaptation. The software-defined radio (SDR) application serves as an example [8]. Dynamic partial reconfiguration (DPR) of Field Programmable Gate Arrays (FPGA) may be employed to address this flexibility problem [9, 10]. It adds virtualization to static hardware, allowing hardware functional blocks to be managed similarly to software components. SDR has been identified by Xilinx as the primary target application for DPR as a result of this capability [11].

DPR has received a lot of academic attention during the last ten years [12–14]. Numerous publications have written about the problem of component programming, as well as case studies for potential applications [15–17]. However, all of those trials are performed for research purposes, and only a small number of them consider their implications for actual applications. In fact, there are currently relatively few practical applications of DPR, despite all the research that has been done. In order to increase its applicability, there is a dearth of input on its application in actual goods that would identify pertinent research difficulties.

This paper's contribution is to assess the benefits and drawbacks of employing DPR in practical professional electronics applications and to provide suggestions for expanding its usage. First, it provides a fair assessment based on tests performed on a group of seven signal and image processing programs in real-world circumstances. In addition, it exposes the challenges created by SDR and indicates the missing components for its implementation in professional electronics applications based on a detailed study of the current flow for real-world use. Third, it lists many benefits of using DPR in commercial electronics applications. Fourth, it offers suggestions for future study to enhance its use. Fifth, compared to the vendor solution, it adds a rapid reconfiguration manager that is employed in the trials and offers an 84-time improvement. According to our knowledge, it is the first reconfiguration manager on the Virtex 4 operating at this speed. In fact, with an 8-bit ICAP, the maximum speed has just been tested in Virtex II [18-21].

2. Literature Review

Molefe et al. (2022) said that a potential backbone network technical option to support the

numerous dynamic bandwidth applications is flexible and combines all photonic and wireless transport networks. In this study, we discuss energy-efficient ways for mitigating approaches targeted at maximizing existing funds in a mutual and wirelessly transport network. In doing so, we consider the difficulties given by transmission impairments, since they often degrade signals and shorten their optical range. Therefore, we suggest a cutting-edge and creative network design that can effectively handle a lot of data. This architecture is meant to support bandwidth-demanding and bandwidth-dynamic services and applications. To further promote greater spectral efficiencies and minimise blocking in both the optical and wireless portions of the composite network, we further present a load-aware energy-efficient resource optimization approach utilizing the LERA algorithm. Performance research demonstrates that, when compared to older benchmark methods, the suggested architectural scheme based on the LERA algorithm outperforms them in achieving minimum blocking while effectively optimizing spectrum usage. As a result, this plan significantly improves the usage of the available resources and network performance as a whole [1].

In a paper by Tej et al. (2020), the authors offer a real-time development of a video shot border detection system on FPGA using the Local Histogram (LH) and Color Structure Descriptor (CSD). These identifiers have been implemented in a few distinct ways, with two for the LH and three for the CSD (32, 16, and 8 quantization levels respectively) (8 and 4 quantization levels). The purchased hardware modules' versions may be loaded into one of the FPGA circuit's many Partial Reconfigurable Regions (PRR), which can be rearranged to meet the application's requirements. Each PRR must be tailored to the specifics of the hardware project to make the most efficient use of the available hardware. The recommended partitioning accounts for the different use cases and component versions already in place. They used the ZedBoard, an affordable function generator for the Xilinx Zynq-7000 SOC. This strategy dramatically increased the load placed on hardware resources.

Saeedi et al 2020 point out that Machine learning methods, which generally create a mathematical model from a set of training data, are used by wearables to detect interesting occurrences. On the other hand, the precision of the related computational model degrades whenever the configuration or context of a wearable device is altered. Through the use of transfer learning as the unifying force, we provide a novel design paradigm that enables the independent reconfiguration of wearable devices. Particularly, we investigate cases when the sensor nodes (SNS) or subject criteria are different from those utilized to generate the training data. They provided two new techniques for creating maps of data in this paper. The first data-mapping strategy was preoccupied with finding signal themes, whereas the second was predicated on effective techniques for tracking down signal analogs inside a network. Transfer learning is where our design shined, and that's where the data mapping algorithms come in. We demonstrated the value of the data mapping methods by applying them to two publicly available datasets concerning human activity detection. Data mapping approaches were found to be up to two orders of magnitude faster than a brute-force approach. Further, they concluded that the proposed framework improves activity detection accuracy by an average of 15% for the initial dataset and 32% for the second.

Venere et al. (2020) introduced a seven-port modulator that could change the polarization of the outgoing signal. The device modulated the carrier wave by using the reflection coefficients produced by four variable loads, much as the six-port modulator does. A pair of antennas, or a dual-port antenna, may be fed by the seventh port. It was the loads that determine the polarization of the received carrier signal and the complexity of the signal's envelope. This research used the device's mathematical model to investigate its functionality and operational modes. The 1575.42 MHz frequency was the focus of the prototype that was being created for testing purposes. It was put to work in the production of multipolar M-quadratic amplitude modulation. Good agreement could be seen between the signal values measured at the prototype's output and the values anticipated by the model [4].

According to Ma et al. (2020), Real-time applications are becoming less viable due to the computational complexity and energy usage of current telecommunication decoding techniques. As a means of resolving the aforementioned tension, they introduced a reconfigurable decoding architecture that could operate in real-time. The proposed architecture was dynamically reconfigurable in part, allowing for the deployment of many decoding techniques on a single physical partition of reconfigurable hardware and their subsequent swap without incurring severe performance or complexity penalties. Experiment results demonstrated that their method was capable of rendering many decoding techniques incorporated in a single system with practicality, simplicity, and power-saving capabilities; the decoding system itself was developed on an FPGA evaluation board. Their three-LDPC-code decoder is built using their architecture and development methodology. Additionally, decoders may switch in milliseconds in real time while saving approximately half the hardware logic required [5].

Wu et al. (2020) were of opinion that over the last several years, developers have seen a rising disparity between the complexity and efficiency of application development. To make up for this discrepancy, new design methodologies attempt to automate much of the designer's work. As a bonus, modern Models of Computation (MoCs) like dataflow-based MoCs make it simpler to define parallelism inside programs, hence increasing designer efficiency. Design technologies that allow for rapid prototyping also provide quick estimates of whether or not a given design decision is a good one. Key to the success of any application prototype is the use of meaningful performance indicators to gauge the soundness of the design decisions made. Information about the hardware may be used to measure many characteristics and new libraries like the Performance Application Programming Interface (PAPI) which make it simpler to collect this data. Automatic PAPI-based instrumentation of dynamic dataflow applications was introduced here using the PAPIFY toolkit. It brought together the PAPIFY parameterization language with the PREESM dataflow Y-chart-based design framework and the Synchronous Parameterized and Interfaced Dataflow Embedded Runtime System run-time reconfiguration manager (SPiDER). The PAPIFY toolkit includes a specialized library to handle run-time monitoring, an automated code generation for static and dynamic programs, and two User Interfaces (UIs) to facilitate the setup and analysis of collected run-time data. Its primary benefits were 1) the capacity to adjust monitoring based on the current state of the system, and 2) the ability to adjust monitoring based on the dynamic redistribution of application workloads.

PAPIFY's run-time monitoring had an overhead of up to 10%, as shown by a detailed characterization utilizing the Sobel-morpho and Stereo-matching dataflow applications [9].

3. Signal Processing Problems

To enhance the performance of RIS communications, four major challenges in signal processing must be addressed: environmental sensing, channel acquisition, beamforming design, and resource scheduling. To accomplish their aims, environmental sensors must acquire highly accurate coordinates or images for use with the RIS. As a result of needing to acquire either the cascaded channels or the individual channels on both sides of the RIS, channel acquisition in an RIS is more complex than in ordinary MIMO systems. Bear in mind that the first CSI obtained will affect the beamforming design. Since it is crucial to steer the reflected beam in the appropriate direction, beamforming design is essential for regulating the phase shifter of the RIS with high accuracy. To create a productive network layout, resource scheduling takes into consideration not only the users' allocations but also the BS power allocation. Careful scheduling of user-RIS pairs or even of user-RIS-BS connections allows for numerous BSs to serve users near one another. AIRIS was designed with the goal of improving upon current RIS communication systems by using artificial intelligence (AI) based techniques to the most critical signal processing issues. Figure 1: Using Tried and True Methods Gaining accuracy in environmental sensing and channel acquisition is facilitated by training on a large dataset. It might also be used to address the nonlinear problem introduced by hardware failures and contaminated pilots. Additionally, AI's diverse learning methodologies may be used for resource scheduling to improve network topology, and the beamforming architecture can be tweaked to achieve different objectives. This suggests that AI may be used to enhance beamforming designs and resource allocation, leading to more reliable systems. The concept of AIRIS is explained in the figure 1. Figure 2 explains a few scenarios of how RIS communications are used in environmental sensing and channel acquisition.
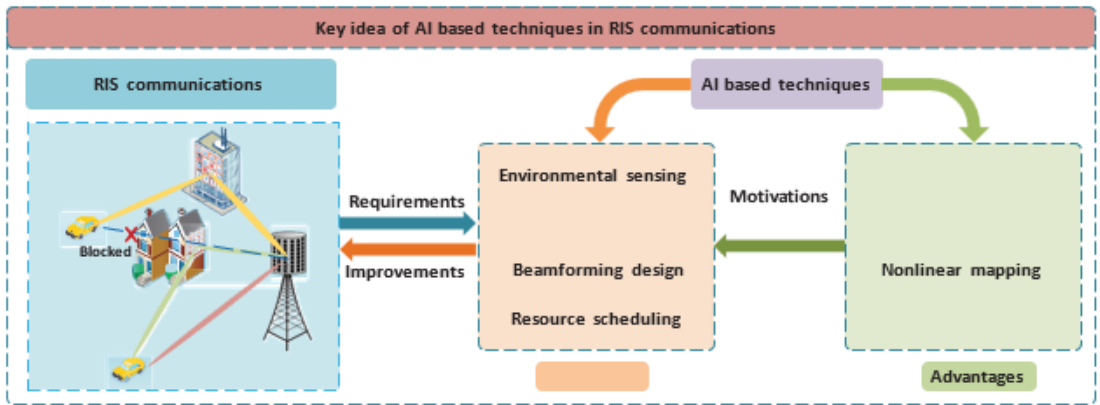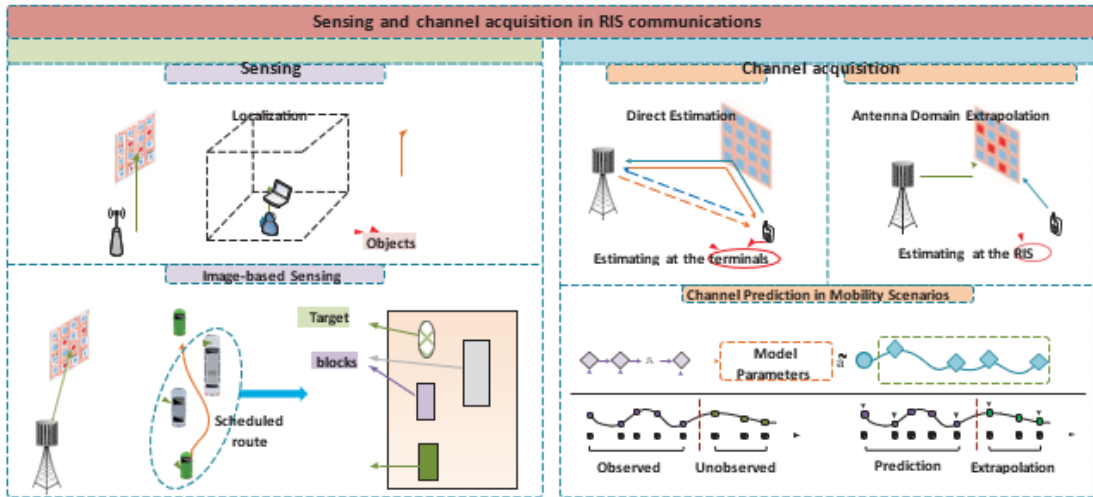

Fig 1 Concept of AIRIS

Fig. 2. Scenarios of environmental sensing and channel acquisition for RIS communications

4. AI Assisted Sensing and Channel Acquisition

For RIS communications, the signal first goes from the transmitter to the RIS, and then from the RIS to the user. There are two different sorts of channel, the feedforward channels and also the individual ones, in the same way as there are two fundamentally different kinds of relay networks. The most prominent applications of artificial intelligence to help with environment monitoring and channel acquisition are shown in Figure 2. Recent research has shown that it is possible to recover the entire channel from the partial one using a technique that is assisted by AI. For completely passive RIS, extrapolation of the whole cascaded channel may be performed at the terminals, which can refer to either the base stations (BSs) or the users. By constructing links between different data layers in a neural network convolutional architecture, the authors of [22] can more accurately estimate the transfer between the sub-sampled channel and the full channel (CNN). In the recently created hybrid active/passive RIS design some RIS components have been equipped with the ability to do signal processing [23][26]. This makes it possible to do channel-by-channel estimation directly at the RIS [22]. In Liao & Hsu (2019)'s paper, the authors use a convolutional neural network (CNN) to activate some of the RIS nodes, then restore channels one by one at those nodes, and finally extrapolate whole missing data from those nodes. In addition, the authors propose a probabilistic sampling hypothesis antenna selection network as a means of positioning these passive RIS components inside the environments that provide the greatest potential benefits [25]. Both the Theodoridis et al. (2007) and Liao & Hsu (2019) works are characterized by their high levels of efficiency and resilience [22][25][27].

5. Methodology

To reduce the cost of cascaded channel estimation, the element-grouping technique has recently been developed. All RIS components in all groups are kept visible under this plan

must be in the same CSI condition, has the same reflection coefficient, and is considered to have the same CSI. However, contrary to popular belief, there are subtle differences across channels even within the same element group.

Although the element-grouping approach improves channel information, it does so only in part. In addition, there would be interference between the groups due to the accomplished channels' connections with various RIS parts.

Antenna extrapolation may inspire designing an NN that can reduce or eliminate interference within individual groups of elements and so obtain more refined partly cascaded channels. The remaining channels from the cascade may be extrapolated using yet another DL method. The good mobility scenario, in which antenna domain channel extrapolation (Liao, & Hsu, 2019) may be transplanted into the time domain, should also be included in the channel estimate over RIS communications.

We can verify that the time-domain extrapolation of channels is, in essence, a time series reconstruction and prediction issue, which the RNN is well-suited to solve. However, problems arise with gradient vanishing and explosion in the RNN model when time-varying RIS channels have long-term dependencies.

A different approach is to create a system based on long short-term memory (LSTM) that partitions its memory from its time-continuous state. Nonetheless, some of the regularly sampled points may lose their efficacy owing to hardware damage and environmental interferences, and only sporadically sampled channels may be acquired. The ODE structure provides inspiration for building a model of a time-varying, continuously sampling channel by connecting the points collected at irregular intervals using coefficients and linear computations.

To achieve best-in-class channel extrapolation performance in the time domain, the newly introduced latent ODE may be used here . In contrast to the standard ODE model, the latent variable framework in the latent-ODE allows the system variation, the probability of observations, and the recognition model to be independently investigated and determined

As a result, the difficulty of dealing with the time domain prediction issue using randomly sub-sampled observations will diminish. When comparing latent-ODE based channel prediction to RNN and traditional ODE, the latter two methods' predictive MSEs are drastically reduced.
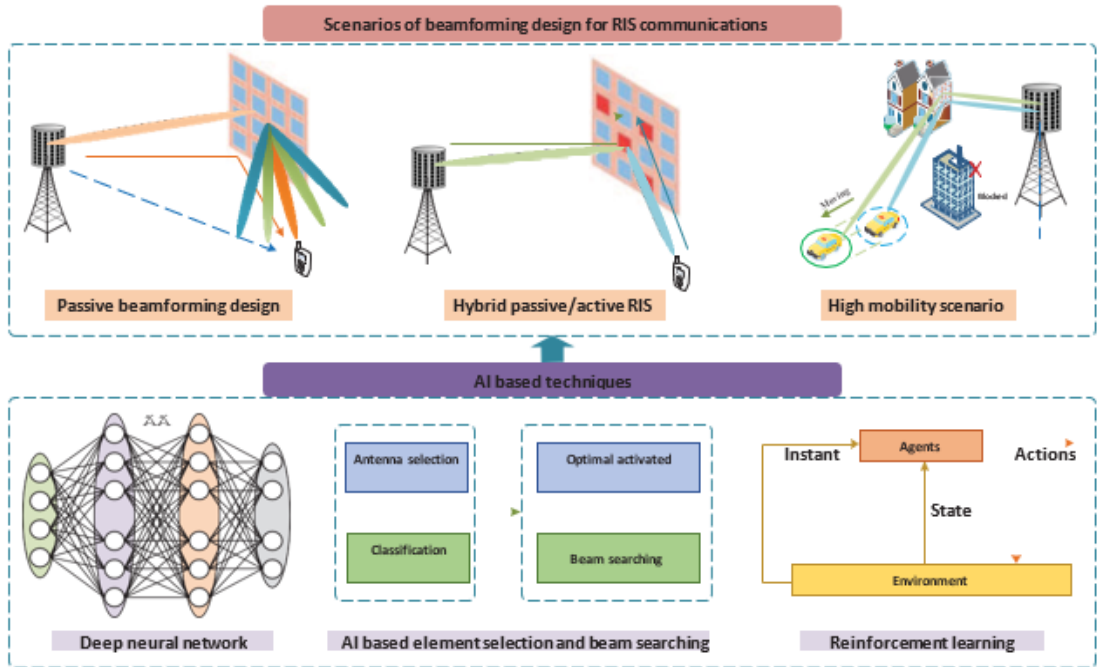
Fig.3. AI-based beamforming design for RIS communications

The deep learning (DL) based strategy to implementing the RIS setting for the indoor communications scenario is proposed, and the DNN is trained to accurately map the user's location as well as the installation of the RIS's components. Therefore, improving the strength of the user's received signal may lead to the optimal RIS beamforming matrix. However, this supervised learning approach needs a lot of data and time to train before it can be useful. By having a single agent learn the best course of action via experimentation and error interactions with its environment, the authors are able to produce optimal beamforming matrices using a reinforcement learning (DRL) based technique. This approach can adapt to its environment and improve its performance, much like weighed Mean square error MSE and equal to 0 beamforming-based proportional programming.

Many papers advocate switching from a phase array-based analog modulation scheme at the BS to RIS beamforming. Optimizing the sum rate of the system is challenging since the constraints on the discrete phase difference are not convex. The authors propose a hybrid precoding strategy based on DL and MDC to tackle the parallel DNN classification problem. In this setup, a large number of DNNs are employed, and their respective outputs are mapped to diagonal cells in a RIS-based analog beamforming matrix.

6. Conclusion

In this work, we provided an overview of the most up-to-date studies on combining AI with RIS communications (AIRIS). Using AI, RIS communications may improve in all areas, including environmental sensing, channel acquisition, beamforming design, and resource allocation, all to a greater extent than is possible with conventional signal processing

methods alone. The convolutional neural network (CNN), the recurrent neural network (RNN), and the deep neural network (DNN) are only a few of the neural network topologies that have been presented. We also give various possible future paths for AIRIS communications, including the assimilation of communication and sensing applying the theory for model generalizations, and AI-based RIS which was before design. The AIRIS is predicted to greatly aid in transmission stability and play a crucial part in the actualization of medium-supplemented and 6G.

Acknowledgments

Authors' contributions
All authors contributed toward data analysis, drafting and revising the paper and agreed to be responsible for all aspects of this work.

Declaration of Conflicts of Interests
The authors declare that they have no conflict of interest.

Data Availability Statement
The databases generated and /or analyzed during the current study are not publicly available due to privacy but are available from the corresponding author on reasonable request.

Declarations
Author(s) declare that all works are original and this manuscript has not been published in any other journal.

References
1. Molefe, M., & Nleya, B. (2022). An energy-efficient resource provisioning scheme for joint all photonic and Wireless Networks. 2022 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (IcABCD) 7(3), 318–415.
2. Tej, B., Hannachi, M., & Abdelali, A. B. (2020). Partial dynamic reconfiguration for efficient adaptive implementation of a video shot Boundary Detection System. 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET). 5(7), 68–148
3. Saeedi R. and Gebremedhin. A. H. (2020). A signal-level Transfer Learning Framework for autonomous reconfiguration of Wearable Systems. IEEE Transactions on Mobile Computing, 19(3), 513–527
4. Venere, A. J., Fernandez Michelli, J. I., Hurtado, M., & Muravchik, C. H. (2019). Design of a multiport microwave modulator for dynamic polarization reconfiguration. IEEE Transactions on Microwave Theory and Techniques, 67(5), 1937–1945.
5. Cho, O.H. (2024). Machine Learning Algorithms for Early Detection of Legume Crop Disease. Legume Research. 47(3), 463-469. https://doi.org/10.18805/LRF-788
6. Gayal, S.D., Sakhare, M.P., Suman, P., Shafi, T.A., Siddiqui, M.F.M.F. and Yeotikar, P.V. (2022). Molecular Detection and Seroprevalence of Lumpy Skin Disease in Cattle. Indian Journal of Animal Research. https://doi.org/10.18805/IJAR.B-4933

7.   Vainshnav, S. P., & Dave, K. K. (2022). A study on Artificial Intelligence and Machine Learning in Banking Sector with special reference to term loan. Pacific Business Review (International), 15(5), 34-53.

8.   Ma, L., Sham, C.-W., Sun, J., & Valencia Tenorio, R. (2020). A real-time flexible telecommunication decoding architecture using FPGA partial reconfiguration. IEEE Transactions on Circuits and Systems II: Express Briefs, 67(10), 2149–2153.

9.   Wu, Q., & Zhang, R. (2020). Towards smart and reconfigurable environment: Intelligent Reflecting Surface Aided Wireless Network. IEEE Communications Magazine, 58(1), 106–112.

10.   Abu-Shaban, Z., Keykhosravi, K., Keskin, M. F., Alexandropoulos, G. C., Seco-Granados, G., & Wymeersch, H. (2021). Near-field localization with a reconfigurable intelligent surface acting as Lens. ICC 2021 - IEEE International Conference on Communication 14(4),14287-14685

11.   Lin, S., Zheng, B., Alexandropoulos, G. C., Wen, M., Chen, F., & Mumtaz, S. (2020). Adaptive Transmission for reconfigurable intelligent surface-assisted OFDM Wireless Communications. IEEE Journal on Selected Areas in Communications, 38(11), 2653–2665

12.   Ying, K., Gao, Z., Lyu, S., Wu, Y., Wang, H., & Alouini, M.-S. (2020). GMD-based hybrid beamforming for large reconfigurable intelligent surface assisted millimeter-wave massive MIMO. IEEE Access, 8, 19530–19539.

13.   Cai, C., Yuan, X., Yan, W., Huang, Z., Liang, Y.-C., & Zhang, W. (2021). Hierarchical passive beamforming for reconfigurable intelligent surface aided communications. IEEE Wireless Communications Letters, 10(9), 1909–1913.

14.   Yang, Z., Liu, Y., Chen, Y., & Al-Dhahir, N. (2021). Machine learning for user partitioning and phase shifters design in ris-aided Noma Networks. IEEE Transactions on Communications, 69(11), 7414–7428.

15.   Yang, Y., Gao, F., Zhong, Z., Ai, B., & Alkhateeb, A. (2020). Deep transfer learning-based downlink channel prediction for FDD Massive Mimo Systems. IEEE Transactions on Communications, 68(12), 7485–7497.

16.   Zhou, Q., Yang, L., & Yan, X. (2018). Reconfigurable instruction-based Multicore parallel convolution and its application in real-time template matching. IEEE Transactions on Computers, 67(12), 1780-1793

17.   Al-Hilo, A., Samir, M., Elhattab, M., Assi, C., & Sharafeddine, S. (2022). Reconfigurable intelligent surface enabled vehicular communication: Joint user scheduling and passive Beamforming. IEEE Transactions on Vehicular Technology, 71(3),2333-2345.

18.   Liu, B., Qin, H., Gong, Y., Ge, W., Xia, M., & Shi, L. (2018). EERA-ASR: An Energy-Efficient Reconfigurable Architecture for Automatic Speech Recognition With Hybrid DNN and Approximate Computing. IEEE Access, 6, 52227–52237.

19.   Wu, Y., Lam, C., Wong, M., Sin, S., & Martins, R. P. (2018). A Reconfigurable and extendable digital architecture for mixed signal power electronics controller. IEEE Transactions on Circuits and Systems II: Express Briefs, 65(10), 1480-1484.

20.   Zhang, S., Zhang, S., Gao, F., Ma, J., & Dobre, O. A. (2021). Deep learning optimized sparse antenna activation for Reconfigurable intelligent surface assisted communication. IEEE Transactions on Communications, 69(10), 6691-6705.

21.   Cao, Y., Jiang, S., Guo, S., Ling, W., Zhou, X., & Yu, Z. (2021). Real-time SAR imaging based on Reconfigurable computing. IEEE Access, 9, 93684- 93690

22.   Theodoridis, G., Soudris, D., & Vassiliadis, S. (2007). A survey of coarse-grain Reconfigurable architectures and cad tools. Fine- and Coarse-Grain Reconfigurable Computing, 89-149

23.   Shang, B., Bentley, E. S., & Liu, L. (2022). UAV swarm-enabled aerial Reconfigurable intelligent surface: Modeling, analysis, and optimization. IEEE Transactions on

Communications, 1-1.

24.　He, G., Ding, X., Zhou, M., Liu, B., & Li, L. (2022). Background noise adaptive energy-efficient keywords recognition processor with reusable DNN and Reconfigurable architecture. IEEE Access, 10, 17819-17827

25.　Liao, Y., & Hsu, T. (2019). A cost-effective adaptive overlapped cluster-based MIMO detector in a frequency domain Reconfigurable modem. IEEE Access, 7, 36103-36121.

26.　Osama, M. (2021). Behavior Recognition based on Signal Processing Technology. International Journal of Hybrid Innovation Technologies, 1(2), 75-90, doi:10.21742/IJHIT.2021.1.2.05

27.　Bellotti, W., Davies, D. N., and Wang, Y. H. (2022). Development and Replication of Improved Recommendation Algorithm for Network Representation Learning. Journal of Science and Engineering Research, 1(2), 61-75, DOI: http://dx.doi.org/10.56828/jser.2022.1.2.5