Design of High-Speed SAR ADC based on 40nm CMOS Process

Chonghui Liu
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Tianfu Li
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Fengbo Wang
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Abstract: In this paper, a high-speed successive approximation analog-to-digital converter is designed based on a 40 nm CMOS process. A split-capacitor Vcm-Based switching strategy is designed, which is able to keep the output common-mode voltage constant for the DAC capacitor array without using Vcm level drive, while saving half of the capacitance compared to the conventional switching strategy and greatly reducing the area consumption. It also combines the features of binary redundancy technology and non-binary redundancy technology by splitting the highest bit capacitor to the lower bit, which enables the DAC array to generate two additional redundant bits without adding other capacitors, reducing its requirements for noise and establishment accuracy; and it adopts asynchronous timing control without external clocks, which is conducive to improving the speed of SAR ADCs and reducing the design complexity. The designed SAR ADC achieves 9.83 bit effective bits, 60.9 dB signal-to-noise distortion ratio, 77.2 dB spurious-free dynamic range, 1.68 mW overall power consumption, and 18.46 fJ/conv-step superiority at a supply voltage of 1.1 V and a sampling frequency of 100 MHz through simulation.

Keywords: high-speed digital-to-analog converter; redundant bits; capacitive splitting; asynchronous logic

1. INTRODUCTION

With the growing demand for ultra-low power systems on chip (SOC), which involve a large number of sensing with analog to digital signal conversion needs and need to be powered, such as mobile and Internet of Things (IOT), medical instruments and other devices. Low-power, low-area ADCs with medium resolution (8-16 bits) and speed (10-100 MS/s) are well suited for these applications, and advances in this type of ADC not only reduce power consumption and extend power cycles, but also provide significant space savings due to their small chip area. Among these common types of ADCs, SAR ADCs use a binary search algorithm, have a simple circuit structure, low power consumption, low requirements for analog circuits, high robustness, and are not process sensitive, which fits well with the trend of process size reduction.

The literature points out [1] that the current SAR ADC design faces many challenges. (1) Achieving a high linearity DAC array design [2], which is used to generate the internal reference voltage during conversion. (2) The maximum achievable number of effective bits is limited by the noise of the comparator [3]. (3) The signal-to-noise ratio of the ADC is limited by the front-end sampling circuit KT/C noise [4]. (4) Unstable supply voltages may lead to unstable reference voltages causing the circuit not to operate properly, especially in low-power SAR ADC designs [5]. There are some trade-offs in SAR ADC implementation, for example, some additional circuitry is essential to achieve low noise, but this may introduce additional area and energy consumption, and high robustness design has the same problem. For high-resolution designs of SAR ADCs, more comparison cycles are inevitable, which may lead to poor speed conditions or more severe energy consumption. Although trade-offs are important, it should be noted in the design that the specific design must be coordinated with the system design requirements and the trade-offs must serve the core metrics, so the core

optimization metrics for SAR ADC design are currently focused on the following three points: low power consumption, high speed and high accuracy.

In this paper, a "capacitance splitting" Vcm-Based switching strategy for DAC arrays is designed based on the 40nm process, so that the DAC arrays can be switched without the need of Vcm level driving and the output common mode voltage can be maintained constant. At the same time, a non-binary capacitance reorganization method is adopted to add two extra redundant bits to the DAC capacitance array without increasing the total amount of capacitance, which increases the tolerance of SAR ADC to comparator noise and reduces the circuit requirement for DAC array setup time, thus realizing the high-speed SAR ADC design.

2. SAR ADC STRUCTURE DESIGN

As shown in Figure 1, the overall structure of the SAR ADC consists of five parts: sample-and-hold circuit, DAC array, comparator, SAR asynchronous logic, and digital error correction circuit. The DAC array adopts an all-capacitor structure and adds two extra redundant capacitors by non-binary capacitor reorganization, which reduces the requirement of establishing accuracy of DAC capacitor array. The switching strategy of the DAC array uses a Vcm-Based switching strategy combined with a "capacitor splitting" technique to keep the output common-mode level constant without the need for Vcm-Based levels. The main function of SAR asynchronous logic is to generate the high-frequency clock inside the SAR ADC and output the digital code word of the comparator comparison result, and to generate the control signal of the switch in the DAC array according to the comparison result. The digital error correction circuit is to convert the 12-bit digital code word output from the SAR asynchronous logic into a 10-bit binary number and output it.

The sampling clock in the figure is an externally provided clock with a frequency of 100MHz and a duty cycle of 20%, which means that during the whole system operation, the sampling time is 2ns and the conversion cycle is 8ns, and because two redundant bits are added, the conversion cycle is divided into 12 comparison cycles. The working process is as follows: when the sampling clock is high, the SAR ADC is in the sampling stage, and the input signal is sampled through the sampling circuit to capture the signal into the DAC capacitor array. When the sampling clock is low, the SAR ADC is in the conversion stage and the comparator starts to compare the sampled input signal and transmits the result to the SAR logic, which outputs the comparison result and generates a control signal to control the DAC array, and the DAC array generates a new voltage input comparator until the end of 12 comparison cycles, and finally the digital error correction circuit converts the 12-bit digital code word The final digital error correction circuit converts the 12-bit digital code word into a 10-bit binary number and outputs it.

Figure 1. SAR Structure Diagram

3. CIRCUIT STRUCTURE DESIGN

3.1 Bootstrap Circuits

Figure 2 shows the gate voltage bootstrap switching circuit in this design. From the figure, CLK is the clock provided externally, and CLKB and CLK are opposite clocks. In the hold phase, CLK is low, node D is high, M5 tube is off, at this time CLKB is high, so that M3 is on, node A is pulled down to ground, while M8 is also on, passing the low level to node G. M4 is on so that node B is charged to VDD, at this time the upper end of capacitor C1 is connected to VDD, the lower end is grounded, and sampling switch M9 is off. In the sampling stage, CLKB is low, M8 is off, and M3 is also off, CLK is high, M2 conducts node D also from high level discharge to low level, so that M5 conducts, at this time, node B and node G are connected, so M7 conducts, node A changes from low level to Vin, because there is no discharge loop, so the charge at both ends of capacitor C1 is conserved, and the voltage of node G is raised to VDD+Vin At this time, the Vgs of sampling tube M9 remains unchanged, so its conduction resistance does not change with the change of voltage. It should be noted that the substrate of M4 tube needs to be shorted with node B. If M4 is connected to VDD according to the conventional connection, M4 is PMOS, and PMOS is generally made in n-well in the layout, there is a forward diode from node B to VDD, so that the voltage at point B can only reach VDD+0.7V at the highest, and when there is no secondary tube, it can reach 2VDD, which will seriously limit the swing..
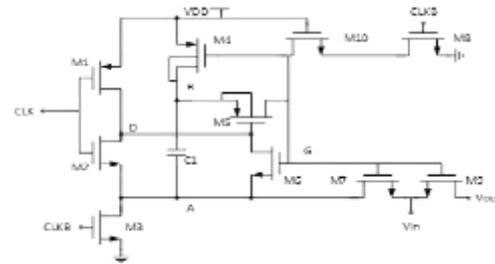

Figure 2. Bootstrap Switch Circuit Diagram

3.2 DAC Array

In this design, the Vcm-Based switching strategy is based on splitting the capacitor, which is difficult to provide Vcm in a low-voltage process. The capacitance of each bit is divided into two equal parts, one part of the lower pole plate is connected to Vcm and the other part is connected to ground. The capacitor connected to Vref is the high level capacitor, and the one connected to ground is the low level capacitor. Since the size of Vcm is half of VDD, the charge stored in the total capacitor is the same after the "capacitor splitting" technique. Meanwhile, the non-binary capacitance reorganization method is used to achieve a certain amount of redundancy by splitting the high level capacitance and assigning it to the low level capacitance. This method does not need to add additional redundant bit capacitors compared to binary redundancy, which greatly saves area, and the base of the weight is no longer 2, but smaller than 2, which can bring a certain amount of fault tolerance to the ADC when an establishment error occurs. The switching method is shown in Figure 3. The design is a 10bit SAR ADC, which introduces two redundant bits, and its MSB capacitance weight is 512, of which 64 capacitors are now split and assigned to the lower capacitor, and in the previous design, due to the "capacitance splitting technique", the lower capacitor is assigned to ensure that In the previous design, due to the "capacitance splitting technique", the lower capacitors are assigned to the even number of the corresponding bits of the DAC capacitor array.
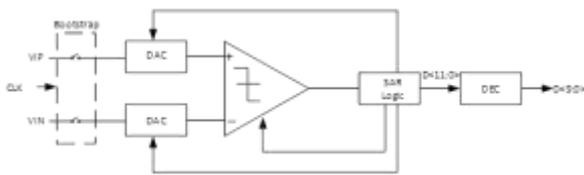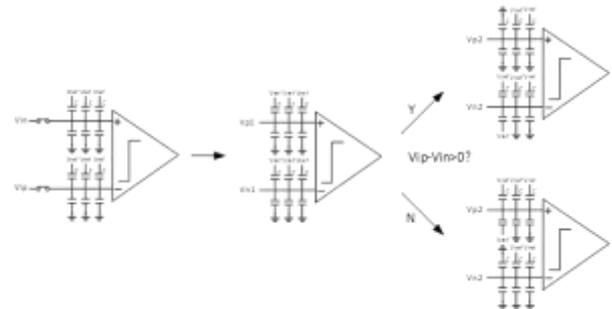

Figure 3. Vcm-Based Switching Based on Split Capacitor

3.3 SAR Logic

The entire SAR ADC asynchronous timing basic logic circuit is shown in Figure 4, where SOC is the sample clock simple inverse generation, used to control the on and off of the asynchronous loop, VALD is used to indicate the working state of the comparator, when the comparator completes the comparison of one bit, VALD = 1. CLKC is the asynchronous SAR ADC internal self-generated high-speed clock, used to control the working state of the comparator. When CLKC=0, the comparator is reset, and when CLKC=1, the comparator compares the comparator. CLKi is the data acquisition clock, which is used to control the acquisition of comparator output results and switches in the DAC array. OVER is the output signal after the last 12 comparisons are completed.
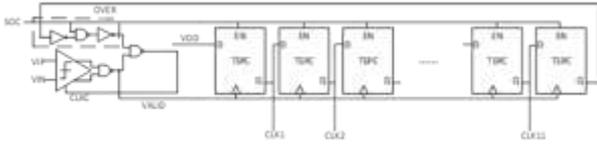
Figure 4. SAR ADC Asynchronous Timing Logic

3.4 Digital Error Correction

Since two redundancies are added to the DAC capacitor array, the total number of bits becomes 12, so the digital error correction（DEC） module is now needed to convert the 12-bit digital code word into a 10-bit binary number output. Since the method of non-binary capacitor reorganization is used, in splitting the MSB capacitors are disassembled with a power of two put, and the weight of each bit is known above, so the weight of each bit can be expressed in a power-of-two summation, and the conversion process can be represented in Figure 5.
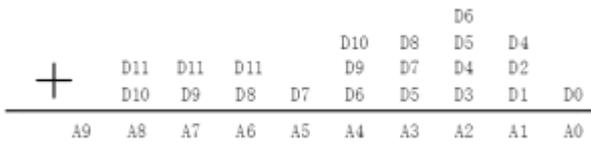

Figure 5. Conversion process

4. SIMULATION RESUITS AND ANALYSIS

The entire 10bit 100Ms/s SAR ADC has been built by verifying the feasibility of each of the above modules. In order to verify whether its overall performance meets the design specifications, a signal with close to full swing is input with a frequency of about 7.52MHz at a supply voltage of 1.1V, a temperature of 25°C, a clock frequency of 100MHz under TT model and a duty cycle of 20%, and the output code word is output through the ideal DAC to obtain the output waveform after sampling and conversion. The output waveform is exported to Cadence simulation software in CSV file format, imported to MATLAB for simulation, and 1024 points are taken for FFT analysis, and the final result is shown in Figure 6.
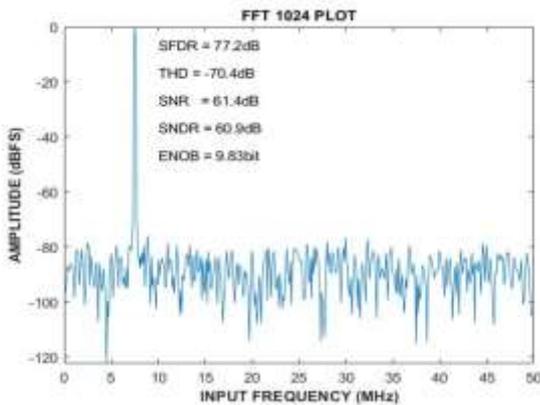

Figure 6. Conversion process

The performance of the final SAR ADC designed in this thesis is shown in Table 1.

Table 1. SAR ADC performance parameters

Design Parameters	Performance Indicators
Supply Voltage	1.1V
Resulotion	10bit
Sampling Rate	100Ms/s
ENOB	9.83bit
SFDR	77.2dB
SNDR	60.9dB
Power	1.68mW
FOM	18.46 fJ/conv-step

In order to design the reliability of the SAR ADC under different environments, the dynamic performance and power consumption under two process angles of ff and ss were simulated at 80°C and -20°C, respectively, as shown in Table 2.

Table 2. Simulation of performance at different corner

	ff -20°C	ff 80°C	ss -20°C	ss 80°C
ENOB(bit)	9.869	9.820	9.841	9.873
SFDR(dB)	77.53	77.91	77.7	78.3
SNDR(dB)	61.17	60.88	61.01	61.2
Power（mV）	1.703	2.369	1.617	1.681

Table 3 compares the performance of the SAR ADC designed in this thesis with that of 10-bit precision SAR ADCs designed in other literature.

Table 3. Simulation of performance at different corner

	[7]	[8]	This work
Process	65nm	20nm	40nm
ENOB	9.51 bit	9.19 bit	9.83 bit
Sampling Rate	100 Ms/s	320 Ms/s	100 Ms/s
Power	1.13 mV	1.52 mV	1.68 mV

5. CONCLUSIONS

In this paper, we design a 10-bit 100MS/s high-speed asynchronous SAR ADC in 40nm process, analyze the energy consumption of different switching strategies, and combine the "capacitance splitting" technology to design a Vcm-based switching strategy based on "capacitance splitting". Based switching strategy of DAC array is designed, so that the DAC array can be switched without Vcm level drive and the output common-mode voltage can be maintained constant. At the same time, the non-binary capacitor reorganization method is adopted to add two extra redundant bits to the DAC capacitor array without increasing the total amount of capacitors, which increases the tolerance of the SAR ADC to the comparator noise and reduces the circuit's requirement for the DAC array build-up time, and improves the speed of the SAR ADC to a certain extent. Completed the overall circuit building of SAR logic and realized the asynchronous timing control of SAR ADC. Completed the construction of the digital error correction circuit module and realized the conversion of 12-bit redundant output codes to 10-bit binary numbers. Simulation verification and layout design of the overall circuit were performed at different process angles. The simulation obtained an effective bit count of 9.83 bits, a spurious-free dynamic range of 77.2 dB, a signal-to-noise distortion ratio of 60.9 dB, and a power consumption of 1.68 mW in a 40 nm

process with a 1.1 V supply voltage. the final optimal value is 18.46 fJ/conv-step.

6. REFERENCES

[1] Wang R. A Review on the Key Optimization Techniques of SAR ADC Design[C]//2021 International Conference on Electronic Information Engineering and Computer Science (EIECS). IEEE, 2021: 951-957.

[2] T. Rabuske and J. Fernandes, "A SAR ADC With a MOSCAP-DAC", in IEEE Journal of Solid-State Circuits, vol. 51, no. 6, pp. 1410-1422, June 2016. Vittoz E,Degrauwe M G R,Bitz S.High-performance crystal oscillator circuits: theory and application[J].IEEE Journal of Solid-State Circuits,1988,23(3): 774-783.

[3] S. Choi, H. Ku, H. Son, B. Kim, H. Park and J. Sim, "An 84.6-dB-SNDR and 98.2-dB-SFDR Residue-Integrated SAR ADC for Low-Power Sensor Applications", in IEEE Journal of Solid-State Circuits, vol. 53, no. 2, pp. 404-417, Feb. 2018.

[4] Y. Shu, L. Kuo and T. Lo, "An Oversampling SAR ADC With DAC Mismatch Error Shaping Achieving 105 dB SFDR and 101 dB SNDR Over 1 kHz BW in 55 nm CMOS", in IEEE Journal of Solid-State Circuits, vol. 51, no. 12, pp. 2928-2940, Dec. 2016.

[5] M. Liu, K. Pelzers, R. van Dommele, A. van Roermund and P. Harpe, "A106nW 10 b 80 kS/s SAR ADC With Duty-Cycled Reference Generation in 65 nm CMOS", in IEEE Journal of Solid-State Circuits, vol. 51, no. 10, pp. 2435-2445, Oct. 2016.

[6] A. Olmos, "A temperature compensated fully trimmable on-chip IC oscillator," 16th Symposium on Integrated Circuits and Systems Design, 2003. SBCCI 2003. Proceedings., Sao Paulo, Brazil, 2003, pp. 181-186.

[7] Hedayati R. A study of Successive Approximation Registers and implementation of an ultra-low power 10-bit SAR ADC in 65nm CMOS technology[J]. 2011.

[8] Liu C C, Kuo C H, Lin Y Z. A 10 bit 320 MS/s low-cost SAR ADC for IEEE 802.11 ac applications in 20 nm CMOS[J]. IEEE Journal of Solid-State Circuits, 2015, 50(11): 2645-2654.

Visibility Detection Based on Object Region Selection

Tao Xu
School of Communication
Engineering Chengdu
University of Information
Technology
Chengdu, China

Jie Zhang
School of Communication
Engineering Chengdu
University of Information
Technology
Chengdu, China

Rong Fu
School of Communication
Engineering Chengdu
University of Information
Technology
Chengdu, China

Abstract: In response to the shortcomings of the dark channel prior and optimization methods in window selection and object depth variation, this paper proposes a visibility detection algorithm based on object region selection. The algorithm introduces superpixel segmentation to divide the image into superpixel blocks and extracts the superpixel blocks containing the target object for visibility calculation. In the visibility calculation, the dark channel and transmission rate extraction can be obtained only from the target object region, without the need to process the entire image region. This improves the accuracy of the target object parameter calculation and speeds up the computation.

Keywords: dark channel prior; visibility detection; superpixel segmentation; object region

1. INTRODUCTION

In recent years, with the popularization of cars, traffic safety has received increasing attention. The reduced visibility caused by haze is an important factor leading to traffic accidents. At the same time, atmospheric transparency is an external manifestation of air pollution. Therefore, detecting atmospheric visibility is an important meteorological observation index, which can guide transportation and environmental management. Currently, visibility detection instruments are relatively expensive and require professional personnel for installation and debugging, making it difficult to be widely promoted. Based on the dark channel prior(DCP), the visibility detection method is simple to use and low in cost, and has great research potential. In the visibility calculation, The dark channel algorithm optimization involves selecting a fixed-size filtering window and optimizing it in the subsequent processing. After two rounds of minimum filtering, a rough estimate of the dark channel image is obtained. The transmission rate image of the scene has a linear relationship with the dark channel image and is usually not directly processed. After obtaining the transmission rate image from the dark channel image, refinement is performed on the transmission rate image. And the dark channel and transmission rate extraction can be obtained only from the target area, without processing the entire image area. Therefore, superpixel segmentation is introduced to segment the image into superpixel blocks, and extract the target object superpixel blocks for visibility calculation. This improves the rigor of the target object parameter calculation and enhances the calculation speed.

2. DCP ALGORITH

Fog imaging model[1][2] is

$$I = Jt + A(1-t) \qquad (1)$$

Here, J is haze-free image, I is haze image, t is transmission and A is airlight.

He[3] counted 5000 outdoor haze-free images, and the results showed that in most non-sky areas of the images, there would always be at least one color channel (in the RGB color space) with a very low value. In other words, applying a minimum filter on the RGB color channels of outdoor haze-free images results in a grayscale image where the pixel values are very low. This is the dark channel prior theory. The dark channel refers to the grayscale image obtained by applying the minimum filter on the color channels of any image, which can be represented by formula (2):

$$\begin{cases} I^{dark} = \min_{c \in \{r,g,b\}} \left(\min_{x \in \Omega} \left(I^c(y) \right) \right) \\ J^{dark} \to 0 \end{cases} \qquad (2)$$

Combining formula (1) and (2) gives the expression for the transmission t:

$$t = 1 - \left(\frac{I}{A} \right)^{dark} \qquad (3)$$

The distance d between the observation location and the target object is measured. The target is located in the image and the center pixel position of the target is marked as x. The visibility can be calculated this way[4]:

$$Vis = -\frac{1}{\beta} \ln \varepsilon \qquad (4)$$

Where, ε is Visual contrast threshold, β is atmospheric extinction coefficient, which can be represented by formula (5):

$$\beta = \frac{\ln t}{d} \qquad (5)$$

The traditional algorithm has a global nature that blurs the boundaries between the target object and neighboring pixels. Processing the input image with a window-based approach actually blends the features of the target object with those of the neighboring pixels. As the window size increases, the interference from neighboring pixels becomes more severe, diluting the features of the target object. Although the traditional algorithm uses optimized transmission[5] to remedy this issue, experimental results show that this optimization only improves but does not solve the problem. It even increases the uncertainty of the results, making the visibility detection results more unreliable.

3. VISIBILITY DETECTION ALGORITHM BASED ON OBJECT REGION SELECTION

3.1 Superpixel segmentation

Image segmentation is a fundamental step in image processing applications, aiming to extract the parts of the image that people are interested in. Depending on the processing task, image segmentation methods include threshold segmentation, region segmentation, etc. In region segmentation, methods such as greedy-based, clustering-based, and graph-based approaches are typically used to segment the image into useful regions for people. Ren et al.[6] proposed the concept of superpixel segmentation, which is a method of image segmentation that groups pixels into superpixels with similar semantic information. By using some features of pixel points, the image is first divided into a large number of irregular pixel blocks, so that the pixels in each region have similar colors, textures, and spatial positions. These pixel blocks are called superpixels. Superpixel segmentation provides a method of image pre-processing that can reduce the complexity of subsequent image processing tasks and has good performance in many computer vision tasks.

Due to the importance of superpixel segmentation, many algorithms have been proposed, but their models are complex and their running times are generally long. To address these issues, Achanta et al.[7] proposed a simple linear iterative clustering (SLIC) algorithm based on the K-means clustering algorithm for superpixel segmentation. The SLIC algorithm has a clear principle, is easy to use, and has a faster processing speed than other superpixel segmentation algorithms. It is a simple and efficient algorithm, and therefore, we chose to use SLIC for superpixel segmentation in this paper.

3.2 Visibility detection algorithm based on object region selection

This paper proposes a visibility detection algorithm based on the selection of target regions. By selecting specific target regions, this algorithm reduces computation and errors, and improves the accuracy and efficiency of detection. Visibility detection is performed specifically on the target region. Due to the selection of target regions, the algorithm only needs to process that region, which can avoid the computational complexity of processing the entire image. The original image and target block are involved in the algorithm calculation, reducing the number of pixels, saving time, and reducing interference from irrelevant pixels in the neighborhood. The algorithm steps are as follows:

(1)Scene setting

In a scene, we need to select a target object, which can be either a colored or a dark object. In the image, the tree trunk is a suitable object, with a dark color and a relatively regular shape. Therefore, we have chosen the tree trunk as our target object and recorded the distance d between the tree trunk and the observation point.
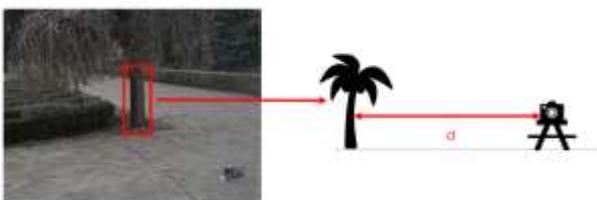
Figure 1. Scene Setting

(2)Segmenting the image.

By performing superpixel segmentation on the image to extract the target object, the superpixel blocks belonging to the target object are labeled as Ω_0. The segmentation result is shown in the figure
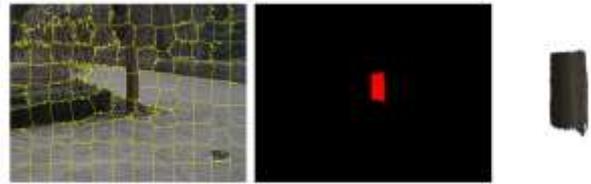
Figure 2. Target Area Segmentation

(3)Calculating the dark channel:

Applying the dark channel prior to Ω_0, by substituting into Equation (2), we obtain the dark channel values within the window of Ω_0.

$$I^{dark} = \min_{c \in \{r,g,b\}} \left(\min_{x \in \Omega_0} \left(I^c(y) \right) \right) \qquad (6)$$

In the equation, only the Ω_0 window exists in the image. Therefore, for any pixel within the Ω_0 window, its corresponding dark channel value is consistent. At this point, I^{dark} is no longer an image of the same size as the input image I, but rather a single numerical value.

(4) Calculating the airlight

For the calculation of airlight value, since the calculation of the dark channel for the entire image is not required in this algorithm, the method of indirectly obtaining the brightest pixel through the dark channel is no longer suitable. Therefore, a quad-tree segmentation method is adopted. The quad-tree segmentation method divides the image into four regions, calculates the mean value of each region, and then divides the region with the highest mean value into four regions again, and calculates the mean value of each region. This process is repeated until the number of pixels in the segmented region reaches a given threshold, and the mean value of this region is taken as the atmospheric light value A. For an RGB image, the quad-tree segmentation method needs to separately obtain A_r, A_g, A_b for each color channel. The illustration below shows the process of obtaining the airlight value A for a single channel (the sky region is not included in the figure, and is only used for illustrative purposes).

Figure 3. Quad-tree segmentation

(5) Calculating transmission

When we know a from fourth step, the formula (3) can be transformed into:

$$t = 1 - \frac{I^{dark}}{\min_{c \in \{r,g,b\}} A^c} \qquad (7)$$

Here, t is also a numeric value, not an image.

(6) Calculating visibility

By measuring the distance d in the first step and the transmission rate in the fifth step, we can obtain the value of visibility in this scene by formula (4).

3.3 Experimental results

The algorithm proposed in this paper extracts the target object and places it in a separate window, strictly controlling the window boundary to avoid crossing objects or backgrounds with different depths of field. This theoretically ensures the accuracy of the detection results, and the consistency of the target object itself can also reduce the interference from picture noise.

Assuming the input image size is M×N, and the number of pixels in the target object window is n, in practical visibility detection scenarios, value of n is much smaller than M×N. The following table lists the orders of magnitude of the data volume used in parameter calculations for the two algorithms:

Table 1. Pixels of calculating

Algorithm	I^{dark}	A	t
DCP	M×N	M×N	M×N
This paper	n	M×N	1

The table shows that the traditional dark channel algorithm requires more computational resources, and as the image size increases, the computational cost of the traditional algorithm increases linearly, greatly reducing its efficiency. The algorithm proposed in this paper is a fast approximation algorithm that is independent of the input image size. As the image size increases, this advantage becomes even more prominent.

4. CONCLUSIONS

The traditional visibility detection algorithm based on the dark channel prior has difficulty in selecting the scale of the filtering window, and optimization methods cannot truly solve this problem, only refining the rough estimated image. This paper introduces superpixel segmentation to solve the window selection problem, making the algorithm results more reliable. Additionally, the algorithm proposed in this paper has lower computational complexity than traditional algorithms, making it more efficient.

5. REFERENCES

[1] E. J. McCartney Optics of the Atmosphere: Scattering by Molecules and Particles New York:John Wiley and Sons 1976.

[2] S. K. Nayar and S. G. Narasimhan, "Vision in bad weather," Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 1999, pp. 820-827 vol.2, doi: 10.1109/ICCV.1999.790306.

[3] He K, Sun J, Tang X. Single image haze removal u- sing dark channel prior[J]. Pattern Analysis and Ma- chine Intelligence, IEEE Transactions on, 2011, 33 (12) : 2341-2353.

[4] Zhao Rui，Hu Liaolin, Zhao Jiangtao. Atmospheric Visibility Measurement Based on Dark Channel Prior [J]. Laser & Optoelectronics Progress,2016,53(6):72-79

[5] HE K M, SUN J, TANG X O. "Guided image filtering."IEEE transactions on pattern analysis and machine intelligence35.6 (2012): 1397-1409.

[6] Ren X, Malik J. Learning a classification model for segmentation[C]// Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. IEEE, 2003.

[7] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.

Design and Implementation of an Accelerated DMA Controller based on AXI Bus

Deqing Cai
School of Communication
Engineering

Chengdu University of
Information Technology
Chengdu, China

Muyao Ge
School of Communication
Engineering

Chengdu University of
Information Technology
Chengdu, China

Hashi Wang
School of Communication
Engineering

Chengdu University of
Information Technology
Chengdu, China

Abstract: Because Direct Memory Access (DMA) hardly consumes processor resources when carrying high-speed data, an accelerated DMA controller based on Advanced eXtensible Interface（AXI）bus protocol is proposed in this paper. The accelerable part of the controller is to replace the CPU for descriptor splitting processing by hardware, which greatly improves the CPU computing power. At the same time, the controller is equipped with eight deeply configured channels, which are used to process different types of tasks. The design supports single transmission of data and linked list transmission type, which can verify the transmitted data and ensure the security of data transmission. In this design, the working frequency of the controller can reach 500M, the power consumption is 1.3mw, and the data throughput rate can reach 40Gbps, which greatly improves the data moving efficiency of the system.

Keywords: AXI bus; High bandwidth; Accelerable; DMA controller; Data verification; Low power consumption;

1. INTRODUCTION

Direct Memory Access is a high-speed data transmission technology that hardly consumes processor resources. In this technology, data is directly read and written without additional intervention of the Central Processing Unit (CPU). The traditional microprocessor data transfer is usually controlled by the CPU [1-3]. The CPU first copies the data of the source address to the internal register, and then writes the data of the register to the destination register, and the amount of data copied has an obvious length limit [4]. With the rapid development of large-scale integrated circuits and the rapid increase of market demand, the above methods severely limit the performance of CPU. In the existing SoC products, the performance of CPU has been greatly improved, which is no longer the key factor for the low data exchange rate in SoC design [5-6]. Therefore, how to improve the data transmission efficiency of the bus and the working efficiency of each slave device has become another key factor affecting the design of a SoC [7].

DMA technology is more and more widely used in the existing SoC design. This technology is to carry out data transfer through hardware, and does not require CPU to participate all the time. During the data transfer of hardware, CPU can handle other tasks [8]. In this way, the data migration mode can be flexibly adjusted and the CPU performance can be greatly exercised. After the CPU transfers the bus right to DMA controller, DMA controller can independently complete the data transfer and processing, and the bus right can be returned to the CPU after the transfer is completed. In SoC chip design, IP core reuse technology has been widely used in order to reduce design cost and design cycle. In order to better use the IP core reuse technology to achieve the enhancement and expansion of the chip, the way of using the on-chip bus in the design process is gradually recognized by the designer [9]. AMBA bus is performance microcontroller communication standard proposed by ARM company. Due to its characteristics of strong universality and high performance

and low power consumption, AMBA bus has become a widely used data transmission bus protocol in SOC design.

In this paper, the design of an accelerated DMA controller based on AXI bus adopts high performance, low power consumption, low delay, high bandwidth in the data transmission [10]. The design realizes single transmission and linked list transmission data types, and the data can be checked in the process of data transmission. Multiple channels are designed inside the controller to store different commands sent by the CPU. Hardware is used instead of CPU to accelerate the analysis of DMA descriptors in the channel. The data transmission granularity can be configured in the DMA descriptor, which well meets the requirements of versatility and flexibility of DMA controller.

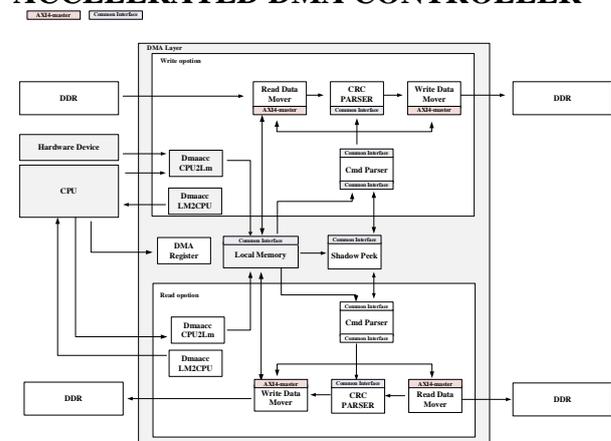## 2. OVERALL STRUCTURE DESIGN OF ACCELERATED DMA CONTROLLER

Figure 1. Overall structure of DMA controller

The DMA controller in this paper is an accelerated circuit based on AXI bus, which mainly implements hardware acceleration processing on the process of CPU parsing descriptor, which replaces the traditional idea of using

software to process descriptor and further improves the efficiency of data moving. As shown in the figure 1, the controller can be divided into write operation and read operation according to command types. It can be divided into external interface circuit and internal logic circuit of controller by function. Therefore, the accelerated DMA controller peripheral interface circuit designed in this paper includes master interface module based on AXI bus, which is used for request and transmission in the process of data moving. The internal circuit of the controller mainly includes the hardware module which can accelerate the parsing descriptor, the multi-channel monitoring and arbitration module, the DMA descriptor processing module, the data request and processing module, the data verification module, the data move status information module and the register module.

3. CIRCUIT STRUCTURE DESIGN

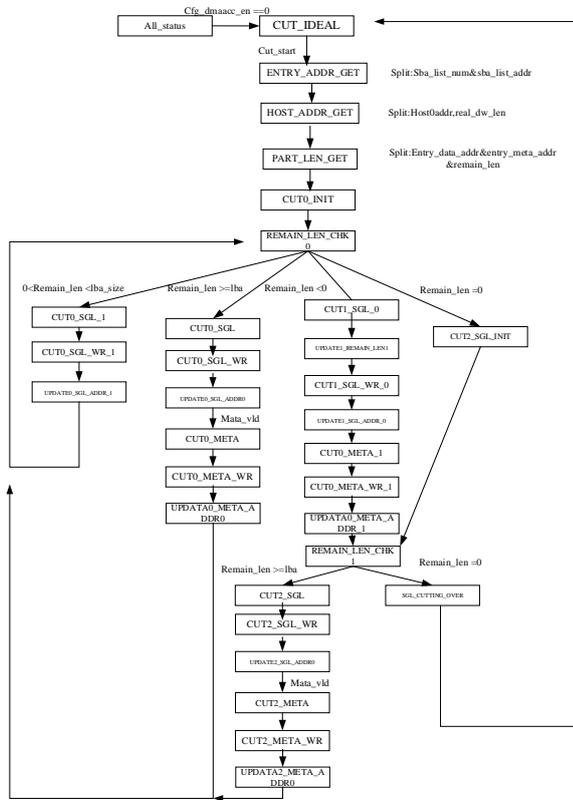## 3.1 THE CIRCUIT OF ACCELERATED PARSER DESCRIPTOR

Figure 2. Split the DMA command state transition graph

As shown in Figure 2, this module is the core circuit for accelerating analytic descriptors in the accelerable controller, and the logic for accelerating processing descriptors is implemented by the state machine logic in the figure above, and corresponding command processing is carried out according to different command types. We can find in Figure 2, the core idea of this module is to split and reassemble DMA descriptors by hardware on behalf of the CPU, allowing the CPU to handle other tasks in the meantime. The module has a FIFO data cache module, which is used to synchronize DMA requests from software or other hardware modules. Then, the internal state machine is triggered to read, parse, split and store the descriptors inside the FIFO. There are 8 independent channels in the module, which can be deeply configurable, used to store the DMA descriptors and Entry information after

splitting, in order to meet the needs of processing different types of DMA commands. The number of DMA requests received by each channel and the response after processing are maintained by the pointer inside the channel.

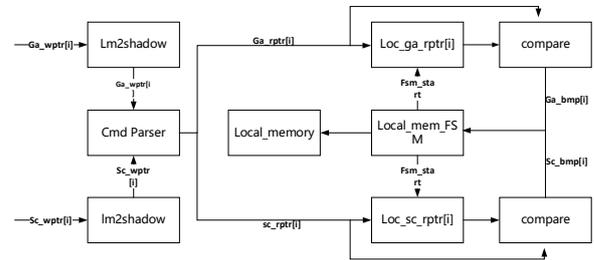## 3.2 MULTI-CHANNEL MONITORING CIRCUIT

Figure 3. Channel monitoring module design diagram

As shown in Figure 3, the core of this module is to design multiple monitoring circuits to work in parallel, which is used to monitor the change of all channel read and write Pointers in DMA controller. If the pointer changes, the state machine will be triggered to acquire tasks in the corresponding channel, and then the DMA controller will be further triggered to work, and the acquired commands and channel pointers will be analyzed and processed. We can find in Figure 3, at the same time, the monitoring circuit further calculates the remaining DMA commands in the channel according to the changes of the pointer, compares the pointer data maintained locally with that in the channel, and determines the state of all channels. If the Pointers maintained by both are consistent, the channel is in idle state; on the contrary, the corresponding channel is undergoing data processing. Finally, the read pointer of the channel is updated to Local Memory module.

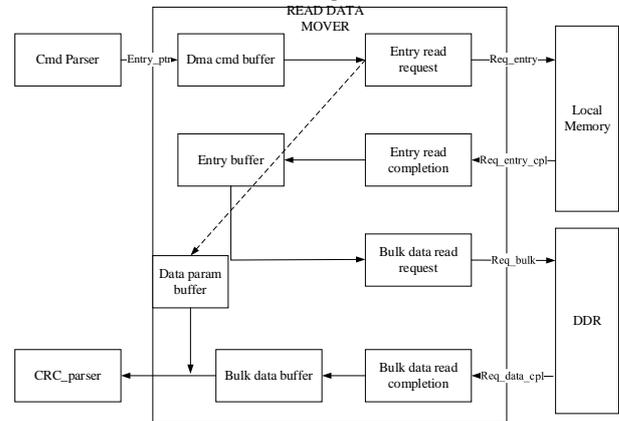## 3.3 READ DATA REQUEST CIRCUIT

Figure 4. Read data request state transition diagram

As shown in Figure 4, this module mainly caches the command information from other module into the internal cache, then parses the command information in the cache area, and initiates an Entry information request to Local Memory according to the storage address of the Entry information contained in the parsed information. Store the requested Entry information in the local Entry Buffer. When the internal state machine detects new data in the Entry Buffer, we can find in Figure 4, it will parse and process the entry in the local buffer. The internal logic of the module will obtain the actual address and data length of the data, and then trigger the control circuit of the state machine to send a request for origin data to the DDR. When the data request is complete, the data and control

information are synchronized, and then transmitted to the others module for processing.
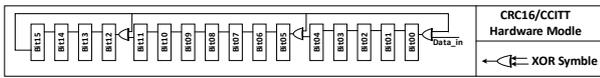
3.4 DATA VERIFICATION CIRCUIT

Figure 5. CRC16-CCITT check hardware model

As shown in Figure 5, the data verification method used in this design is Cycle redundancy Check (CRC) algorithm. The module through the hardware way to describe the circuit. The basic idea of CRC check is to use linear coding theory, that is according to the K-bit binary sequence transmitted at the sending end of data, to generate a set of check codes for check according to certain check rules, whose bit width is r, that is CRC codes. In the process of data transmission, CRC codes are added to the tail of user data to form a new set of binary sequences. At the receiving end, according to the data information and CRC code verification rules, to determine whether the data transmission in the process of error. If there is no remainder in the result of data verification at the receiving side, it indicates that the data transmission is correct. Otherwise, the error location of the transmitted data can be inferred according to the remainder. The figure above shows the hardware circuit model under the CRC16-CCITT protocol.
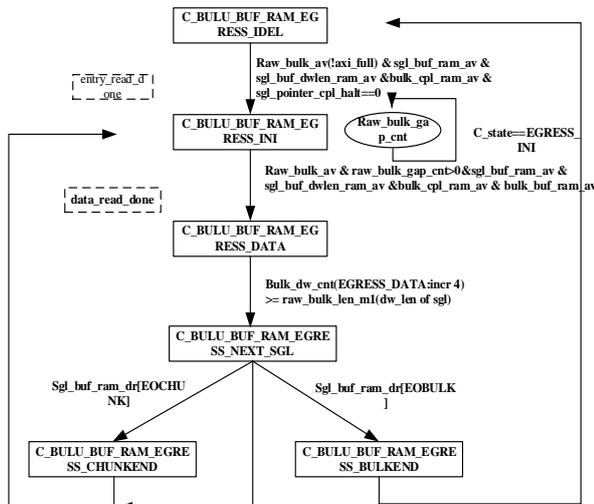
3.5 WRITE DATA TRANSMISSION CIRCUIT

Figure 6. Write data transfer state transition diagram

The working principle of this module is the same as that of the data reading request module. we can find in Figure 6, The core logic of this circuit is mainly to cache the command information from the command parser module into the internal cache area of this module, and then parse the command in the cache area. Then a request for Entry information is sent to the Local Memory according to the destination address information of entry stored in the DMA command. The verified data and the entry information returned by the request are synchronized and matched through the state machine inside the module, and the information is written to the destination through the AXI bus. The above state machine used for data transmission is shown in Figure 6. The data transmission process of the state machine can be roughly divided into two parts. First, according to the control

information in the entry, the data within each entry is transmitted; then, according to the effective identifier of the control information, the data transmission of the entire command is completed by the state machine.

3.6 FIXED PRIORITY ARBITRATED CIRCUIT

The monitoring circuit of this module will monitor the update of internal commands of all channels at the same time. If multiple channels receive new data moving commands at the same time, and the DMA controller only processes commands of a specific channel each time, the preemption of hardware resources by multiple channels will be involved. In this case, the quorum module is required to determine which channel uses the hardware resources. Because in this scenario, different channels correspond to different hardware devices, the arbitration module in this design adopts the arbitrator with fixed priority, that is, the priority of each channel is fixed. If multiple channels simultaneously initiate resource occupation requests, the channel with higher priority will be granted the channel use right.

4. ANALYSIS OF SIMULATION RESULTS

Figure 7. Data simulation results

The design follows the collaborative verification method of hardware and software, and uses the VCS simulation software to build the verification environment. It mainly verifies the single data transmission mode, linked list data transmission mode, accelerated analytic descriptor function, multi-channel monitoring function, arbitration function and interface transmission function of DMA controller. Through the report, we can find that the above functions are consistent with the expected hardware behavior. The data can be moved normally. This Design is based on SMIC's 0.18um process library and synthesizes the controller with Design compiler software. The working frequency of the controller can reach 500M, the power consumption is 1.3mw, and the data throughput rate can reach 40Gbps.

5. CONCLUSIONS

In this paper, an accelerated DMA controller interface circuit based on AXI bus is designed, which has the characteristics of high bandwidth and low power consumption. In the read and write transmission circuit, there are 8 independent channels, and the channel depth can be flexibly configured. In order to meet different command requirements, and based on the SMIC 180nm process comprehensive implementation. The performance of design satisfies the expected index.

6. REFERENCES

[1] Ahmed M A,Aljumah A,Ahmad M G． Design and Implementation of a Direct Memory Access Controller for Embedded Applications．International Journal of Technology,2019,10(2)：309-319 .

[2] Qiao Lufeng,Wang Zhigong. Design of DMA controller for multichannel PCI bus frame engine and data link manager [C].New York,USA.2002.P1481-1485.

[3] Zou Y Y,Chen M,Wei K L．Design of Custom AXI4 IP Based on AXI4 Proto-col.Applied Mechanics and Materials,2014, (3634):2326-2330.

[4] Markatos, E.P. and Katevenis,M.G.H. "User-level DMA without operating system kernel modification" High-Performance Computer Architecture, 1997. Third International Symposium on 1-5 Feb. 1997 P322-331.

[5] Kim, D.and Managuli, R.and Kim, Y."Data cache and direct memory access in programming mediaprocessors", Micro,IEEE,Volume:21 Academic Publishers, 1998.

[6] Kidav J U,Sivamangai N M,Pillai MP,et al．FPGA and prototype of cycle stealing DMA array signal processor for ultrasound sector imaging sys-tems Microprocessors and Microsystems,2019,64(64)：53-72.

[7] Saluja H,Grover N.Memory Controller and Its Interface using AMBA 2.0 ．IJEM-International Journal of Engineering and Manufacturing ,2019,9(4)：33-44.

[8] A. F. Harvey, *National Instruments. DMA Fundamentals on Various PC Platforms Application Note 011*, 1991.

[9] J. Corben, A. Rubini and G. Kroah-Hartman, "Linux device drivers", *O'Reilly Media*, 2005.

[10] C. Bohm, H. Kavianipour, D. Nygren, C. Robson, C. Wernhoff and G. Wikstrom, "A low energy muon trigger for icecube", *Proc. Conf. Rec. 2008 IEEE Nucl. Sci. Symp.*, pp. 2784-2787.

A Convolutional Neural Network Accelerator Based on FPGA

Jincheng Zou
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Qing Tang
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Congcong He
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Abstract: This paper analyzes and studies the hardware programmable logic resources on small-scale FPGA chips, providing reasonable hardware resource support for subsequent neural network accelerator designs. A flexible 32-bit instruction set is designed for control by the Processing System (PS) on the Programmable Logic (PL) side, making motion state detection flexible and controllable. When designing the hardware side, this paper uses a resource-sharing strategy, and most of the calculation modules are designed using on-chip DSP resources to reduce the resource consumption of the calculation module. An innovative strategy of partially not caching the data between layers of the neural network is applied to reduce the demand for on-chip cache. To optimize on-chip storage space, this article partitions the limited BRAM space on the chip in a reasonable manner and improves the efficiency of on-chip data reading and writing through parallel processing, thereby improving the real-time performance of the neural network.

Keywords: FPGA; Accelerator; Neural Network; Real-time; Instruction

1. INTRODUCTION

Currently, the hardware platforms for computing or accelerating convolutional neural network algorithms mainly consist of four types: Central Processing Units (CPUs), Graphics Processing Units (GPUs), Application-Specific Integrated Circuits (ASICs), and Field-Programmable Gate Arrays (FPGAs). CPUs are general-purpose processors that have strong single-threaded performance and a large amount of cache, which is very effective for processing control flow and serialization tasks in neural network models. However, for compute-intensive tasks, CPU performance may not meet the requirements. GPUs are highly parallel hardware computing platforms designed to accelerate graphics rendering, but due to their highly parallel nature, they are also widely used to accelerate the training and inference computation of neural networks. Compared to CPUs, GPUs have stronger computing capabilities, which can greatly reduce the training and inference time of neural networks. Although GPUs have an advantage in computing power, their power consumption is relatively high. Therefore, energy consumption and heat dissipation may need to be considered when performing large-scale neural network computations, and it may be difficult to apply GPUs on embedded platforms. ASICs are customized integrated circuits specifically designed for specific applications, and their advantages lie in their high performance and low power consumption. In some neural network computations that have high requirements for performance and energy efficiency, ASICs can provide very outstanding performance. However, the design and production cost of ASICs is high, and they typically need to be designed and manufactured for specific applications. FPGAs are programmable logic devices that can implement specific logic functions through programming. They can be designed and optimized according to specific applications, and therefore, they can adapt to new neural network algorithms, models, and tasks faster than ASICs. Compared to GPUs and CPUs, FPGAs have lower latency and higher computing performance, and their power consumption is lower, making it easier to apply them on embedded platforms. Currently,

accelerating neural network computations through FPGAs still faces some challenges, such as the high development cost of accelerators for specific convolutional neural network models, weak portability of accelerators designed for different FPGA models, and relatively few open-source materials related to accelerating neural network computations using FPGAs. These problems to some extent hinder the development of FPGAs as accelerators for neural network computations.

This article studies and designs a hardware accelerator for a convolutional neural network model that is used for computing motion state detection in a software-hardware collaborative manner. The accelerator has the characteristics of flexibility, high scalability, and low power consumption, which makes it easy to apply to FPGAs with limited hardware resources. A set of instruction sets based on this hardware computing platform is designed so that users can flexibly implement various convolution calculations through different instructions. This article optimizes the hardware design of the accelerator for a specific convolutional neural network model used for motion state detection, using on-chip DSP resources to build multiplier arrays and accumulators to complete convolution and pooling calculations, and designs an activation function calculation module to enable it to run smoothly on FPGA platforms with limited resources. Special hardware design is also carried out for the reading of image data by the hardware computing platform, with limited on-chip BRAM resources allocated as on-chip memory, and the on-chip cache is divided into multiple subspaces.

2. THE DEVELOPMENT OF FPGA-BASED NEURAL NETWORK ACCELERATORS

In 1996, Cloutier et al. first used FPGA to perform calculations for convolutional neural networks. However, due to the limited amount and variety of resources on the FPGA at that time, the speed of computing convolutional neural networks was very slow. In 2015, Microsoft researchers used Intel's Stratix 10 FPGA to accelerate the convolutional neural network inference of the deep learning platform Caffe. They

used a system called Project Brainwave, which uses a large number of FPGAs to achieve low-latency neural network inference. On the ImageNet dataset, the system achieved higher inference performance than GPUs and has been widely used in fields such as speech and image recognition. In 2022, the National Space Science Center of the Chinese Academy of Sciences proposed a convolutional parallel acceleration scheme based on FPGA to improve the speed and energy efficiency of convolutional neural networks running on resource- and power-limited embedded platforms. They used the fusion of convolutional layers and batch normalization layers to reduce the complexity of computation, and achieved a peak computing performance of 52.56 GFLOPS on the ZCU104 platform. The performance is 4.1 times that of CPUs, and the energy consumption is only 9.9% of GPUs..

3. ACCELERATOR BASED ON FPGA

3.1 Convolution Calculation Module

This article presents three parallel multiplier arrays designed for simultaneously computing convolution operations on three channels in a convolutional neural network. Each parallel multiplier array consists of 25 8-bit width multipliers generated by DSPs, with a one-clock cycle delay in the output of the multiplier. To prevent overflow in the output results, the multiplier's output width is set to 16 bits. With this design, the parallel computing module can output the results of 75 multiplication operations in two clock cycles at the highest speed.
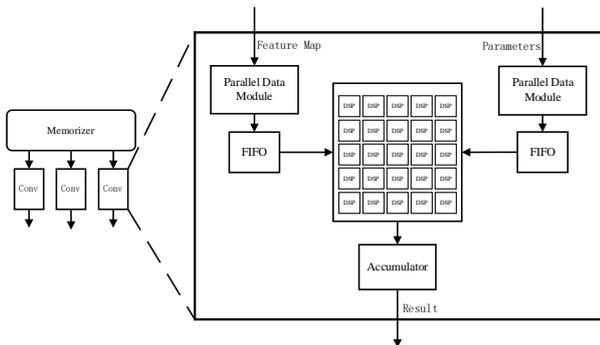

Figure. 1 Convolution computing module

In the convolution operation of a neural network, after completing the multiplication operation of a convolution kernel, the output results of the multiplier need to be accumulated to obtain a complete output result of the convolution kernel. This article designs a unified accumulator structure using DSPs to construct a stack of adders. The inputs of the first-layer adder consist of two data, while the inputs of the remaining adders come from the calculation results of the previous-layer adder and a new data. After the accumulation is completed, the result is output through a register to ensure stability, thereby saving the on-chip buffer space required for the accumulation process.
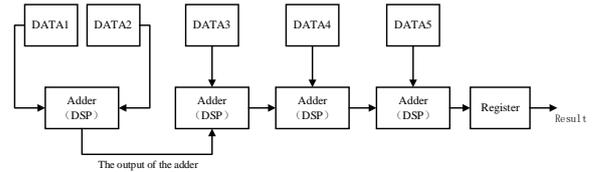

Figure. 2 Accumulator

In order to efficiently perform convolution operations, the feature maps and weight parameters need to be input into parallel multipliers. Since the data is output from BRAM in a serial manner, a separate parallel data generation module is designed in this paper for data parallelization. This module converts the data of the feature maps output from BRAM and the data of the convolutional neural network model parameters into parallel data by using shift registers. Specifically, when the data output from BRAM is valid, the control logic outputs a data valid signal. Upon receiving the data valid signal, this module stores the data and performs shifting operations. When this module has stored all the data to be calculated for a convolution kernel, it puts the parallel data into the corresponding FIFO.
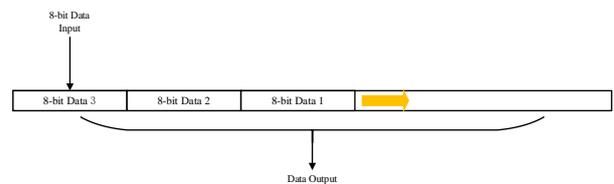

Figure. 3 Parallel data module

3.2 Activation Function Calculation Module

This article uses Hard-Sigmoid and Hard-Swish as activation functions. These activation functions are different piecewise linear functions. Compared with traditional activation functions, piecewise linear activation functions can not only retain the information of the original signal but also introduce non-linear characteristics, enabling the network to better adapt to various complex input data. Moreover, they do not involve complex exponential calculations that are unfriendly to FPGA. Therefore, the structure of the activation function calculation module in this article can be designed to be relatively lightweight.
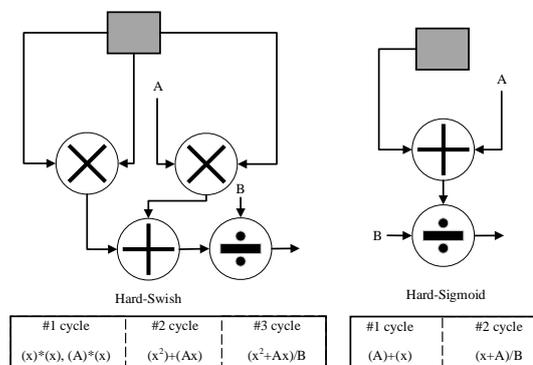

#1 cycle	#2 cycle	#3 cycle
(x)*(x), (A)*(x)	$(x^2)+(Ax)$	$(x^2+Ax)/B$

#1 cycle	#2 cycle
(A)+(x)	(x+A)/B

Figure. 4 Activation function calculation

3.3 Instruction

The instruction set designed in this article adopts a 32-bit width, which is the same as the register width of the ARM Cortex-A9 processor. Using one register can store all instruction information, thus reducing the hardware resource consumption on the programmable logic side. This instruction set covers the basic information required to perform one convolutional neural network calculation, including the calculation operation type, input feature map size, input channel number, output channel number, convolution kernel size, padding size, stride size, and so on. Among them, the bit width used for the calculation operation type is 4 bits.
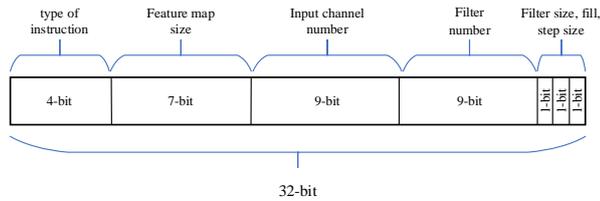

Figure. 5 Instruction composition

3.4 Memory on Chips

There is programmable BRAM memory in FPGA, which is a dual-port memory based on on-chip resources of FPGA. It can achieve high-bandwidth and low-latency storage access.
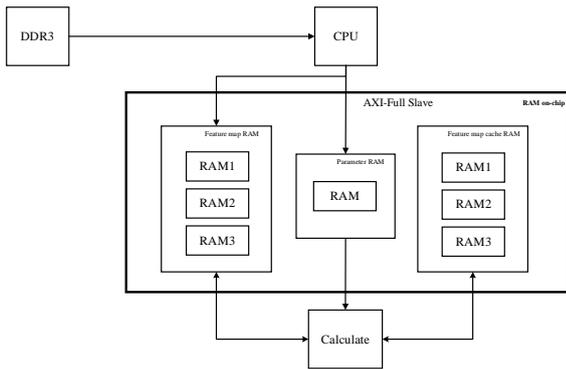

Figure. 5 RAM on-chips

The data structure of the on-chip RAM storage space is shown in the following table.

Table 1. The data structure for storing data in RAM

RAM	Width(bit)	Depth	Capacity(Mb)
FM-RAM1	8	102400	0.78125
FM-RAM2	8	102400	0.78125
FM-RAM3	8	102400	0.78125
P-RAM	8	20480	0.15625
FMC-RAM1	8	53248	0.40625
FMC-RAM2	8	53248	0.40625
FMC-RAM3	8	53248	0.40625

4. ANALYSIS OF RESULTS

The development environment for the convolutional neural network accelerator designed for motion state detection in this article is Xilinx's Vivado 2018.3. The experimental platform uses the XC7Z020-CLG400-2 chip from the ZYNQ-7000 series.
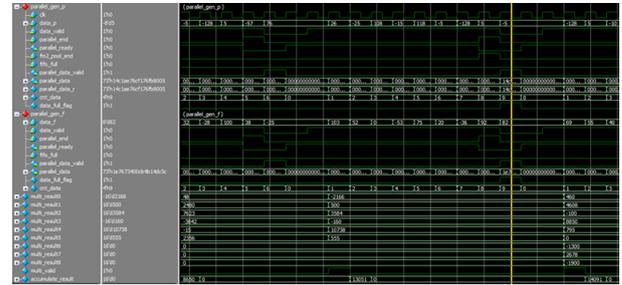

Figure. 6 Simulation waveform of the convolution calculation

In order to evaluate the error caused by quantized 8-bit parameters in this article, the following figure shows the error in each interval when performing operations on quantized 8-bit data for Hard-Swish and Hard-Sigmoid.
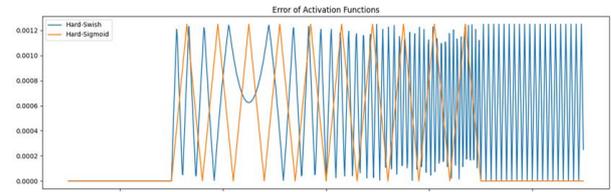
Figure. 7 Error of activation function

The Hard-Swish and Hard-Sigmoid activation functions have fixed value ranges, resulting in relatively small errors outside the range of [-3, 3]. Based on statistics, the average error caused by the Hard-Sigmoid within the range of [-3, 3] is 6.197×10^{-4}, while the average error caused by the Hard-Swish is 8.135×10^{-4}.

The improved LeNet-5 was tested on the MNIST dataset, and the inference results were validated using both hardware computation and simulation.

Table 2. The data structure for storing data in RAM

	LeNet-5
Original model	96.5%
Accelerator	96.1%

To ensure the operational efficiency of the convolutional neural network accelerator and improve the real-time performance of motion detection, it is necessary to increase the operating frequency of the accelerator within a reasonable range. In this article, the entire convolutional neural network accelerator was tested using clock frequencies of 150MHz, 180MHz, and 200MHz, respectively.

Table 3. The performance of accelerator

Clock frequency(MHz)	Performance	Power（W）
150	10.60GOPS	3.293
180	12.67GOPS	3.354
200	14.12GOPS	3.401

5. CONCLUSION

This paper investigates the acceleration of convolutional neural networks based on FPGAs, and proposes a motion state detection platform with high flexibility and lower power consumption, and runs on top of a chip of limited size, highly utilizing the on-chip BRAM and DSP resources, making it advantageous for embedded devices.

6. REFERENCES

[1] Krogh A. What are artificial neural networks?[J]. Nature biotechnology, 2008, 26(2): 195-197.

[2] Collobert R, Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning[C]//Proceedings of the 25th international conference on Machine learning. 2008: 160-167.

[3] Erhan D, Szegedy C, Toshev A, et al. Scalable object detection using deep neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 2147-2154.

[4] Li D, Chen X, Becchi M, et al. Evaluating the energy efficiency of deep convolutional neural networks on CPUs and GPUs[C]//2016 IEEE international conferences on big data and cloud computing (BDCloud), social computing and networking (SocialCom), sustainable computing and communications (SustainCom)(BDCloud-SocialCom-SustainCom). IEEE, 2016: 477-484.

[5] Qiu J, Wang J, Yao S, et al. Going deeper with embedded fpga platform for convolutional neural network[C]//Proceedings of the 2016 ACM/SIGDA international symposium on field-programmable gate arrays. 2016: 26-35.

[6] Choquette J, Giroux O, Foley D. Volta: Performance and programmability[J]. Ieee Micro, 2018, 38(2): 42-52.

[7] DiCecco R, Lacey G, Vasiljevic J, et al. Caffeinated FPGAs: FPGA framework for convolutional neural networks[C]//2016 International Conference on Field-Programmable Technology (FPT). IEEE, 2016: 265-268.

[8] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.

[9] Han S, Mao H, Dally W J. Compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv 2015[J]. arXiv preprint arXiv:1510.00149, 2015, 305.

[10] Zhang C, Li P, Sun G, et al. Optimizing FPGA-based accelerator design for deep convolutional neural networks[C]//Proceedings of the 2015 ACM/SIGDA international symposium on field-programmable gate arrays. 2015: 161-170.

[11] Ma Y, Suda N, Cao Y, et al. Scalable and modularized RTL compilation of convolutional neural networks onto FPGA[C]//2016 26th international conference on field programmable logic and applications (FPL). IEEE, 2016: 1-8.

Renewable Energy Implementing Artificial Intelligence: Applications, Problems, and Challenges

Omkar Singh[*]
National Institute of Fashion
Technology
Patna, India

Vinoth R
National Institute of Fashion
Technology
Patna, India

Abhilasha Singh
National Institute of Fashion
Technology
Patna, India

Navanendra Singh
National Institute of Fashion Technology
Patna, India

Abstract: World transformation has meaningfully abridged the chief gas, diesel, and coal power sources. Therefore, substitute power sources based on renewable energy mainly focus on fulfilling the energy demand of the world and avoiding global warming. Among different energy sources, solar energy is the critical substitute energy source for producing electricity using photovoltaic (PV). Conversely, energy engendering performance is highly dependent on cyclical and environmental factors. The changeable version of the environment shakes energy productivity and tends to create a disapproving influence on constancy, dependability, and the grid process. Therefore, a precise prediction of PV productivity is critically required to guarantee the endurance and reliability of the grid. The detailed study reviews the perilous techniques based on PV forecast using machine learning techniques. This paper summarizes different types of renewable energy with their merits and demerits. This paper also demonstrates the key challenges existing in renewable energy using machine learning and various artificial intelligence applications in real-life scenarios. Finally, an inclusive analysis of machine learning techniques in renewable energy through a detailed literature survey is presented to forecast energy production shortly better.

Keywords: Artificial intelligence, machine learning, renewable Energy

1. INTRODUCTION

Electrical energy plays a significant role in today's era due to the demand for globalization in the world. Earlier, oil and coal natural gas were remnant energies used to produce electrical power [1]. Though these cradles of energies are contenting electricity petition, its enormous custom has instigated massive reduction of relic energies and ecological glitches [2]. The cohort of electricity from remnant energy power shrubberies has produced foremost effluence in CO2 release and conservatory gas (GHG) emanation, therefore foremost to foremost environment vicissitudes everywhere of the world [3]. Custody, in the opinion of this evidence, the custom of substitute foundations of energy to encounter electrical petition has been discovered intensively [4]. Amid these substitute possessions, renewable energy has increased primary attention universally. Energy creation from RESs is ecologically approachable, with precise GHG production, long-lasting and fewer charge than conservative dynamism [5] sources. Among various kinds of RESs PV plays a dictate role in diffusion degree in energy arcades because of solar energy of sun-provided ended atmosphere [6]. As per the study, the amount of received energy from solar radiation is 1.8×10^{11} MW. It is probable maintainable energy created from solar with the attention of investors, policymakers, environmental technologies, and governments [7]. Therefore, PV has a large prospective in town and rural electrification [8]. Solar Energy is being used in various forms, such as cooling, heating, energy creation, mutual power, and unreceptive system [9]. PV system also includes minimum preservation cost,

prolonged lifetime, and connection cost compensation within a specific time [10].

2. REALTED WORK

Gautam et al. [11] proposed a framework to enhance solar energy power to the grid and solar on a similar node by using similar switching technology with the help of a decision tree ML-based algorithm in a python framework. Hybrid ML technique with extensive data analysis proposed by Sharmila et al. [12] for optimization distribution of existing energy sources targeted to intelligent power management. Alkhnadari et al. [13] developed a machine learning-based hybrid model called MLSHM, which merges ML techniques using the statistical technique for further precise forecasting of future solar energy production using energy plants based on renewable energy. Traditional functions are used to develop MLSHM model, and these functions are given below:

$$r(t) = \sigma(w_r x(t) + u_r h(t-1) + b_r) \tag{1}$$

$$z(t) = \sigma(w_z x(t) + u_z h(t-1) + b_z) \tag{2}$$

$$\hat{h}(t) = \sigma(w_h x(t) + u_h h(r(t) * h(t-1) + b_h) \tag{3}$$

$$h(t) = (1 - z(t)) * h(t-1) + z(t) * \hat{h}(t) \tag{4}$$

The appropriate symbols used in Eq. (1)-(4) are symbolized by author in [13]
An intelligent system model MERIDA is implemented by Marijana et al. [14], which assimilates the collection of big

data and forecast model of power utilization for every power source in public construction. The computation function used to develop this model is as follows:

$$y_c = f(\sum_{i=1}^{n} w_i x_i) \quad (5)$$

where y_c is the calculated output and x_i, and w_i represents the elements of input and weight vectors X and W. Som et al. [15] developed heuristic methods called SOC to assess the formal of responsibility of lead–acid succession through optimum energy cohort in a separate amalgam wind-solar renewable energy distribution method. The specific vector used to design the process is as follows:

$$v_i^t = \frac{(m_{i+\frac{n}{2}} + \varepsilon(m_{i+\frac{n}{2}})) v_{i+\frac{n}{2}}}{m_i + m_{i+\frac{n}{2}}} \quad (6)$$

The author in [15] has elaborated on specific notations and symbols used in Eq. (6). Two various ML-based algorithms ELMS and MOGAs, have been proposed by Ronay et al. [16]. Developed techniques are implemented for immediate airstream rapidity forecast from basic information customary of hourly airstream rapidity capacities for Canada and Regina in Saskatchewan. The basic predictions intervals used to develop both ML-based techniques are as follows:

$$P_r(L(x)) < y(x) < U(x) = 1 - \alpha \quad (7)$$

$$PICP = \frac{1}{n_p} \sum_{i=1}^{n_p} c_i \quad (8)$$

$$NMPIW = \frac{1}{n_p} \sum_{i=1}^{n_p} \frac{(U(x_i) - L(x_i))}{y_{max} - y_{min}} \quad (9)$$

The appropriate symbols and notations used in Eq. (7)-(9) have been described by the author in [16]. Musaylh et al. [17] developed an ANN model that applied forecast variables for 6-hour (h) and daily prediction power demand (G) conjecturing. The response parameters comprised 6 pertinent weather parameters SILO and 51 Re-investigation parameters attained from ECMWF models. To develop ANN model, a particular technique is used:

$$y(x) = F(\sum_{i=1}^{L} w_i(t).x_i(t) + b) \quad (10)$$

The author in [7] has summarized symbols and notations mentioned in Eq. (10). REMS technique has been proposed by Prakash et al. [18] that efficiently changes conceivable lots to renewable thrilled native power storing centered based on

charge-discharge dealings and grid obtainability thus plunging energy ingesting from the grid. Perera et al. [19] proposed HOA technique by uniting Replacement and AEM in directive to haste up optimization procedure though upholding correctness. The efficiency of the solar panel used in the proposed technique can be calculated as follows:

$$\eta_n^{SPV} = p^{SPV} \left[q^{SPV} (\frac{G_t^{\beta}}{G_0^{\beta}}) + (\frac{G_t^{\beta}}{G_0^{\beta}})^{m^{SPV}} \right]$$

$$\left[\begin{array}{l} 1 + r^{SPV} (\frac{\theta_t^{SPV}}{\theta_0^{SPV}}) + s^{SPV} (\frac{AM}{AM_0}) + \\ (\frac{AM}{AM_0})^{u^{SPV}} \end{array} \right], \forall t \in T \quad (11)$$

The author in [19] describes specific notations and symbols used in Eq. (11). CRO approach has been developed by Sanz et al. [20], which syndicate dissimilar exploration appliances into a solitary technique, giving a worldwide exploration process of extraordinary excellence. A wrapper method is used as:

$$\sigma^o = \arg\min_{\sigma,\alpha}(\int v(y, f(x^*\sigma,\alpha)) dp(x,y)) \quad (12)$$

The author in [20] has elaborated on notations and symbols used in Eq. (12). Zhang et al. [21] proposed IESs technique for monitoring the power adaptation has developed an operative technique for refining grid litheness and plummeting functioning charge of IESs. A specific function is used to design IES technique.

$$\theta_{th}(t) = COP_{ave}\Delta P_{HP}(t) \quad (13)$$

Ahmad et al. [22] developed diverse regions of definite conservational and power ingesting statistics that are acquired for effort specifications assortment and demonstrating examination. The specific formula is used for developing diverse regions:

$$b^n = g^n (X^n g^{n-1} g^{n-2} (X^2 g^1 (X^1 q + c^1) + c^2 + c^{n-1}) + c^n) \quad (14)$$

The author in [12] has elaborated on precise symbolizations used in Eq. (14). Assessment of various ML techniques used in renewable energy is given in Table 1.

Table 1: Assessment of various ML techniques used in renewable Energy

ML Technique	Advantages	Disadvantages
ML	Enhances solar power	When employees on

Framework [11]	usage, collects real-time data, exports data prediction strategy, and reduces electricity cost.	large scale affect other environmental factors.
Hybrid ML [12]	Substantial gain ensures innovative energy management leads and relevant data collecting.	Electricity streamlining data demand is very high.
MLSHM [13]	Achieves higher accuracy, and collects data very efficiently.	Training set parameters need to be tested.
MERIDA [14]	Improves energy efficiency, enables reconstruction management plan, and minimizes power utilization.	Potential intelligent power management using macro and micro ML techniques should be enabled.
SOC [15]	Utilizes standalone renewable power and optimized renewable sources.	Enhances electricity and design costs.
ELMS and MOGAs [16]	Employee's good precision prediction predicts airstream.	Need help finding out the methodical analysis impact on hidden neurons.
Hybrid ANN Model [17]	Calculates forest indecision provides high accuracy, and covers multiple horizons.	Need to improve forecasting accuracy. Consumes more electricity.
REMS [18]	Minimizes power utilization, and provides better accuracy.	On large scale performance of REMS decreases.
HOA [19]	It provides better accuracy, reduces operational time, and saves energy.	Unable to provide energy necessity generation on a regional scale.
CRO [20]	It improves wind speed direction and gives an admirable prediction.	Need to explore new methods of deep learning in big FSP.
IES [21]	Solves DM problems, and adapts energy conversion ratio.	Need help to work with a multi agent.
CA & CN [22]	Improves prediction accuracy, and reduces operational cost.	Performance degrades with the selection of ultra-short and ST energy.

3. RENEWABLE ENERGY AND ITS TYPES

Renewable energy, frequently mentioned as spotless energy, originates from natural cradles or continually replaced procedures. Such as airstreams or sunlight preserve admirable and blustering, and their obtainability is contingent on weather and period [23, 24].

3.1 Solar Energy

Hominids have been binding solar power for a whole year to produce harvests, stay earnest and parched foods. As per NREL, new power from the sun sprays on the ground in a single hour and is used by everybody in the ecosphere throughout the year. In today's era, sun rays are used through different types, including warmth hospices and trades, warm marine, or energy devices [25].

3.2 Wind Energy

We have returned from windmills of old fashioned, now day turbines as high as towers using turbines closely as extensive in thickness standpoint at courtesy everywhere the ecosphere. Breeze power chances a turbine's knife-edges, which fodders a power-driven producer and harvests power. Airstream, which explanations for slightly more than 6 % of U.S. cohort, has developed an inexpensive power source in various portions of the nation [26].

3.3 Hydroelectric Power

Hydropower is a prevalent renewable power cradle for power in the United States. However, airstream power is rapidly predictable to revenue over the chief. Hydropower trusts on water characteristically rapid marine in an extensive waterway or quickly sliding marine from a high theme and changes the strength of marine into power by rotating a generator's turbine knife-edges [27].

3.4 Biomass Energy

Carbon-based biomass derives from florae and faunae and comprises trees, waste wood, and crops. Once the biomass is scorched, the biochemical energy is unconfined as warmth and can produce power with a vapor turbine. Biomass is frequently incorrectly pronounced as spotless, renewable firewood and an olive green substitute to firewood and further relic coals for creating power [28].

3.5 Geothermal energy

If you are always unperturbed in the warm mainspring, you use geothermal energy. The ground's core is warm as the sun's superficial because of the sluggish deterioration of harmful atoms in pillars at the earth's midpoint. Puncturing bottomless shafts carries subversive hot water to superficial as a hydrothermal source, formerly impelled over a turbine to generate power [29].

3.6 Ocean

Tidal and wave power energy are quiet in the growing stage, but oceanic will continually be lined by the moon's magnitude, which brands yoking its influence and a beautiful

selection. Specific tidal energy methods may damage the environment, e.g., tidal bombardments that seem like barriers and are situated in a marine inlet or cove [30].

3.7 Solar Power

At the lesser gauge, we can yoke the sun's emissions to influence the entire community, whether over PV cubicle plates or inert solar home-based enterprises. Inert solar households are intended to be comfortable in the sun over south-facing spaces and then recollect the balminess through concrete tiles, bricks, and other resources that stockpile warmth [31].

3.8 Geothermal Heat Pumps

A geothermal system is a novel revenue on an identifiable procedure. The loops at the spinal of your fridge are a slight warmth impel, eliminating warmth from the inner to save nourishments renewed and unruffled. At home-based, geothermal impels custom the ground's constant temperature to calm households in the summertime and earnest households in wintertime and smooth to warm water [32].

3.9 Hydrogen

Hydrogen desires to be joined with extra rudiments, such as oxygen, to create water, not to happen logically to gas on its individual. While hydrogen is unglued from alternative components, it can be second handled together with petroleum and Energy [33]. Merits and demerits of different renewable energies have been shown in Table 2.

Table 2: Renewable energy merits and demerits

Types of Energy	Merits	Demerits
Solar	Infinite functionality of sunshine, unlimited supply of solar energy.	Solar Energy is an impractical cost for some households.
Wind	Its unsoiled energy source avoids air pollution; it does not produce CO_2 and other harmful products.	It increases the cost of transition lines. Some cities oppose rising noise pollution; certain birds are killed by striking the turbine while flying.
Hydroelectric	It is used for projects working on a large scale. Such as hoover dam, and also covers small projects, including small dams on small rivers.	It creates disturbance and negative distress for animals and living lives. It also changes water status and ecosystems.

Geothermal	It signifies the potential of energy; it avoids footprints on the earth.	It takes maximum cost to build infrastructures and susceptibility to the earthquake.
Ocean	It is predictable energy, more reliable, and plentiful; it is a clean energy source.	It disturbs the ocean environment, habitats, and sea life, and rough weather creates lower energy.
Hydrogen	It is used as a clean scorching fuel and creates low pollution.	It is incompetent when it originates to stop contamination.
Biomass	It is used for personal use at home, in business, and in our daily lives.	Though fresh florae require carbon dioxide to cultivate, florae revenue time to cultivate.

4. AI APPLICATIONS IN RENEWABLE ENERGY

Investigators and corporations are sightseeing customs to influence AI and recover the competence and convenience of predictable power technology. These tools work within three predictable energy compasses: Energy Forecasting, Energy Efficiency, and Energy Accessibility [34].

4.1 Energy Forecasting

4.1.1 Nnergix

It touches the energy source produced from atmosphere-reliant power such as solar and wind; it tends to challenge renewable power. Nnergix is called predicting application based on web and data mining energy. Nnergix uses data from satellites and trains that data using ML to analyze companies' data to be better accurate and predict [35].

4.1.2 Xcel

Xcel is applying AI that goals at lecturing the contests allied with the undependability of climate-reliant power sources such as wind and solar. Xcel can express energy sources will vary in métier. Xcel is applied in retrieving climate intelligence having sophisticated correctness and glowing comprehensive [36].

4.2 Energy Efficiency

4.2.1 Google DeepMind

Google brings this AI application in 2014 to improve energy usage. It aims to decrease power ingesting and releases

subsequent when power is recycled. The application is utilized for cooling Google's information servers by 40 %, minimizing power utilization and bills [37].

4.2.2 Verdigris Technologies

This scheme deals an s/w technology that influences AI to enhance power ingesting. It is intended for big profitable constructions and executives of innovativeness amenities. The connection procedure instigates with IoT h/w connection. Shrewd sensors are devoted to the customer's electrical board to discover power ingesting [38].

4.3 Power Efficiency

4.3.1 Verv

It is a powered-based AI and is being used as a homegrown subordinate in power organizations. It delivers data on household utilization power utilization. Verv allows the employer to understand annals in what way every utilization at household customs power. It also assistances consumers in controlling their power expenditures [39].

4.3.2 PowerScout

Its goals at refining punter tutoring and admittance to the renewable power system. It customs AI-based model possible reserves on usefulness prices consuming manufacturing data. PowerScout influences analytics of data to recognize "shrewd homespun development" based on exclusive topographies and power customs at household customers [40].

5. CHALLENGES IN RENEWABLE ENERGY

Renewable energy is demarcated as power composed of natural assets. Newly, there has been rising attention to renewable power and converted one of the critical causes of energy cohort. Unmoving, there are numerous disputes and contests specified as [7]:

5.1 Availability of Power

The significant anxieties in renewable power are energy cohort contingent on usual assets that are irrepressible by persons. Such as, solar power current is produced only when sunlight is obtainable and goes off at nighttime; airstream energy also be influenced by the obtainability of the airstream. Therefore, if the airstream hustle is exactly little, turbines will not cause seizure, resulting in zippo energy movement to gridiron [9].

5.2 Issue in Power Quality

Dependably extraordinary energy excellence is desired to guarantee constancy and extraordinary competence of the system. The excellence of the energy source permits the scheme to effort glowing with super dependability, and inferior outlays [13].

5.3 Resource Location

Maximum renewable power florae that stake their power with the grid necessitate ample space. Maximum, renewable power generations are verbalized by the position which can be repellent to consumers. Certain renewable power generations are not obtainable in different areas [29].

5.4 Information Barrier

Improvement in this field is going on; deficiencies exist regarding information and consciousness, almost the welfare and requirement of renewable power. Speculation and assets stipends have been completed and are obtainable for enacting renewable vitalities. The vibrant necessity for administration assistance to support and guide contenders and possible receivers how smearing for renewable power inducements [33].

5.5 Cost Issue

The extraordinary preliminary connection fee is a key sprint in improving renewable power. Though the enlargement of petroleum herb necessitates around $6 apiece megawatt, it is clear that airstream and solar energy shrubberies are also great essential speculation. Therefore, stowing methods of the produced power is affluent and signify an actual contest in megawatt fabrication [38].

6. CONCLUSION

Current energy production and dissemination technology tendencies demonstrate that the smooth diffusion Grid has enlarged significantly. End manipulator utilizations are satisfying further slightly to energy excellence condition. This scenario offers a methodical evaluation of sources of energy excellence difficulties connected with renewable dispersal-produced systems (airstream and solar power). Power reduces with airstream infiltration and intensification with solar infiltration. In this paper, specific issues and challenges have been discussed in current renewable energy, paving the path for future investigations for researchers. The paper also discusses various types of renewable energies with its merit and demerits. Most recent applications in renewable energy, which show the basic idea for researchers for further improvements in energy generation, are also discussed.

7. REFERENCES

[1] Nawad, S.J, Sharma, S.K., Wyne, S. et al.: Quantum Machine Learning for 6G Communication Networks: State-of-the-Art and Vision for the Future. DCN. **7**,46317-46350 (2019).

[2] M.N. Akhter, S. Mekhilef, H. Mokhlis et al." Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques", IET Renewable Power Generation, vol. 13, pp. 1009-1023, 2019.

[3] N. Phuunagpornpitak and S. Tia" Opportunities and Challenges of Integrating Renewable Energy in Smart Grid System", EMSES, vol. 34, pp. 282-290, 2018.

[4] J. Heinermann and O. Kramer" Machine learning ensembles for wind power prediction", Renewable Energy, vol. 89, pp. 671-679, 2016.

[5] M.A.F.B. Lima, P.C.M. Carvalho, L.M.F. Ramirez et al." Improving Solar Forecasting Using Deep Learning and

Portfolio Theory Integration", energy, vol. 20, pp. 1-25, 2020.

[6] R.B. Ammar, M.B. Ammar and A. Oualha" Photovoltaic Power Forecast Using Empirical Models and Artificial 1 Intelligence Approaches For Water Pumping Systems", IJRE, vol. 20, pp. 1-45, 2020.

[7] T. Ahmad, H. Zhang and B. Yan" A review on renewable energy and electricity requirement forecasting models for smart grid and buildings", Sustainable Cities and Society, vol. 20, pp. 1-101, 2020.

[8] S.M. Dawaoud, X. Lin and M.I. Okba" Hybrid renewable microgrid optimization techniques: A review", RSER, vol. 82, pp. 2039-2052, 2018.

[9] E.M.Sandhu and T. Thakur" Issues, Challenges, Causes, Impacts and Utilization of Renewable Energy Sources - Grid Integration", JERA, vol. 4, pp. 636-643, 2014.

[10] A. Essl, A. Ortner and P. Hetteger" Machine Learning Analysis for a Flexibility Energy Approach towards Renewable Energy Integration with Dynamic Forecasting of Electricity Balancing Power", 978-1-5090-5499-2/17/$31.00 ©2017 IEEE, pp. 1-6, 2017.

[11] A. Gligor, C.D. Dumitru and H.S.Grif" Artificial Intelligence Solution for Managing a Photovoltaic Energy Production Unit", INTER-ENG, vol. 22, pp. 626-633, 2017.

[12] M. Gautam, S. Raviteja and R. Mahalakshmi" Household Energy Management Model to Maximize Solar Power Utilization Using Machine Learning", ICRTAC, vol. 165, pp. 90-96, 2019.

[13] P. Sharmila, J. Baskaran, C. Nayanatara et al." A hybrid technique of machine learning and data analytics for optimized distribution of renewable energy resources targeting smart energy management", ICRTAC, vol. 165, pp. 278-284, 2019.

[14] M. Alkandari and I. Ahmad" Solar Power Generation Forecasting Using Ensemble Approach Based on Deep Learning and Statistical Methods", Applied Computing and Informatics, vol. 19, pp. 1-26, 2019.

[15] M.Z. Susac, S. Mitrovic and Adela Has" Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities", International Journal of Information Management, vol. 20, pp. 1-12, 2020.

[16] T. Som, M. Dwivedi, C. Dubey et al." Parametric Studies on Artificial Intelligence Techniques for Battery SOC Management and Optimization of Renewable Power", ICCIDS, vol. 167, pp. 353-362, 2020.

[17] R. Ak, O. fink, E. Zio et al." Two Machine Learning Approaches for Short-Term Wind Speed Time-Series Prediction",NNLS, vol. 15, pp. 1-14, 2015.

[18] M.S.A.Musaylh, R.C. Deo, J.F. Adamowski et al." Short-term electricity demand forecasting using machine learning methods enriched with ground-based climate and ECMWF Reanalysis atmospheric predictors in southeast Queensland, Australia", Renewable and Sustainable Energy Reviews, vol. 113, pp. 3-22, 2019.

[19] K.N. Prakash and V.D. Prasanna" Machine learning based Residential Energy Management System", 978-1-5090-6621-6/17/$31.00 ©2017 IEEE, pp. 1-4, 2017.

[20] A.T.D. Perera, P.U. Wickramsinghe, V.M. Nik et al." Machine learning methods to assist energy system optimization", Applied Energy, vol. 243, pp. 191-205, 2019.

[21] S.S. Sanz, L.C. Bueno, L. Prieto et al." Feature selection in machine learning prediction systems for renewable energy applications", Renewable and Sustainable Energy Reviews, vol. 90, pp. 728-741, 2018.

[22] B. Zhang, W. Hu, Di Cao et al." Deep reinforcement learning–based approach for optimizing energy conversion in integrated electrical and heating system with renewable energy", Energy Conversion and Management, vol. 202, pp. 1-13, 2019.

[23] T. Ahmad, H. Chen, W.A. Shah " Effective bulk energy consumption control and management for power utilities using artificial intelligence techniques under conventional and renewable energy resources", EPES, vol. 109, pp. 242-258, 2019.

[24] E. Hossain, I. Khan, F.U. Noor et al." Application of Big Data and Machine Learning in Smart Grid, and Associated Security Concerns: A Review", 10.1109/ACCESS.2019.2894819, IEEE Access, pp. 1-40, 2017.

[25] S.K. Jha, J. Bilalovic, A. Jha et al." Renewable Energy: Present research and future scope of Artificial Intelligence", RSER, vol. 77, pp. 297-317, 2017.

[26] K.W. Kow, Y.W. Wong, R.K. Rajkumar et al." A review on performance of artificial intelligence and conventional method in mitigating PV grid-tied related power quality events", RSER, vol. 56, pp. 334-346, 2016.

[27] K.R. Kumar and M.S. Kalavathi" Artificial intelligence based forecast models for predicting solar power generation", PMME, vol. 5, pp. 796-802, 2018.

[28] M. Borunda, O.A. Jaramillo, A. Reyes et al." Bayesian network in renewable energy systems:A biblio graphical survey", RSER, vol. 62, pp. 32-45, 2016.

[29] M.A.M. Daut, M. Y. Hassan, H. Abdullah et al." Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review", RSER, vol. 16, pp. 1-11, 2016.

[30] M. Ramezanizadeh, M.H. Ahmadi, M.A. Nazari et al." A review on the utilized machine learning approaches for modeling the dynamic viscosity of nanofluids", RSER, vol. 114, pp. 1-15, 2019.

[31] M.Sharifzadeh, A.S. Lock and Nilay Shah" Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression", Renewable and Sustainable Energy Reviews, vol. 108, pp. 513-538, 2018.

[32] G.D. Sharma, A. Yadav and R. Chopra" Artificial Intelligence and Effective Governance: A Review, Critique and Research Agenda", AJTES, vol. 20, pp. 1-16, 2020.

[33] S. Sinha and S.S. Chandel" Review of recent trends in optimization techniques for solar photovoltaic–wind-based hybrid energy systems", RSER, vol. 50, pp. 755-769, 2015.

[34] A. Stetco, F. Dimohammadi, X. Zhao et al." Machine learning methods for wind turbine condition monitoring: A review", IJRE, vol. 18, pp. 1-23, 2018.

[35] C.Voyant, G. Notton, S. Kalogirou et al." Machine learning methods for solar radiation forecasting: A

review", Renewable Energy, vol. 105, pp. 569-582, 2017.

[36] S. Walker, W. Khan, K. Katic et al." Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings", Energy & Buildings, vol. 209, pp. 1-14, 2020.

[37] Z. Wang and R. S. Srinivasan" A review of artificial intelligence-based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models", Renewable and Sustainable Energy Reviews, vol. 16, pp. 1-13, 2016.

[38] H. Wang, Z. Lei and X. Zhang et al." A review of deep learning for renewable energy forecasting", Energy Conversion and Management, vol. 198, pp. 1-16, 2019.

[39] H. Wang, Y. Liu, B. Zhou et al." Taxonomy research of artificial intelligence for deterministic solar power forecasting", ECM, vol. 214, pp. 1-17, 2020.

[40] J.H. Yousif, H.A. Kazem, N.N.Alattar et al." A comparison study based on artificial neural network for assessing PV/T solar energy production", Case Studies in Thermal Engineering, vol. 13, pp. 1-13, 2019.

[40] S.M.Zahraee, M.K. Assadi and R. Saidur" Application of Artificial Intelligence Methods for Hybrid Energy System Optimization", RSER, vol. 66, pp. 617-630, 2016.

Design of a 14-bit Pipelined ADC using Ring Amplifier

Pan Luo
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Yu Guan
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Zhiyuan Tang
School of Communication
Engineering
Chengdu University of
Information Technology
Chengdu, China

Abstract: With the progress of integrated circuit technology, the intrinsic gain of transistors has become increasingly low, and the power consumption and complexity of OTA operational amplifiers have become higher, increasing the overall design difficulty of pipeline ADC. In order to improve the gain of the operational amplifier, improve the overall accuracy of the ADC, and reduce circuit power consumption, a 14-bit 20MSPS analog-to-digital converter with a pipeline structure of 2.5bit-2.5bit-2.5bit-2.5bit-2.5bit-2.5bit-2.5bit-2.5bit-2bit was designed using the SMIC 40nm process and analog drive digital technology under a 1.2V power supply voltage based on a new type of ring amplifier.The simulation results show that the SNDR of the input low-frequency signal pipeline ADC is 70.47dB, the SFDR is 85.5dB, and the ENOB is 11.45bit. When inputting high-frequency signals, the SNDR of the pipeline ADC is 68.35dB, the SFDR is 81.3dB, and the ENOB is 11.07bit.

1. INTRODUCTION

Currently, most pipeline ADC structures use traditional operational transconductance amplifiers (OTAs) or open loop amplifiers as interstage operational amplifiers, and the power consumption in pipeline ADC mainly comes from the operational amplifiers in each stage of MDAC[1]. Operational amplifiers in MDAC generally require very high gain to ensure transmission accuracy. With the progress of integrated circuit technology, the intrinsic gain of transistors has become increasingly low. In order to ensure the high gain of operational amplifiers, the use of gain bootstrapping, digital calibration, and other technologies have made the power consumption and complexity of circuits higher, and increased the overall design difficulty of ADCs. In order to improve the gain of operational amplifiers, improve the overall accuracy of ADC, and reduce circuit power consumption, the industry has been looking for a new type of operational amplifier that can replace traditional OTAs for pipeline ADC structures.

Table 1. Comparison of Common Operational Amplifiers

Amplifier type	Gain	Power	Linearity
Traditional OTA	limited	high	secondary
Open loop amplifier	secondary	secondary	low
Dynamic amplifier	high	very low	low
Ring amplifier	very high	low	high

With the power supply voltage is 1.2V, using traditional Cassode high-gain OTA as an operational amplifier will result in insufficient voltage redundancy. In order to ensure low power consumption while reducing the use of gain enhancement techniques, the gain of OTA will also be greatly limited. In order to avoid the problems of insufficient gain and high power consumption caused by traditional operational amplifiers, academic circles have also proposed solutions such as open-loop amplifiers[2], dynamic amplifiers[3], and ring amplifiers. A comparison of commonly used operational amplifiers is shown in Table 1. In practical design, open loop amplifier structures have PVT performance issues, and calibration techniques are often needed to ensure their gain. Dynamic amplifiers have fast amplification speed and low power consumption due to their switch controlled operation. However, due to their open loop structure, their accuracy and linearity are relatively low.

Ring Amplifier was first proposed and applied to pipeline ADC in 2012[4]. It has very high gain and rail to rail output swing, and its power consumption is very low. It meets the requirements of pipeline ADC for high gain and low power consumption of operational amplifiers, perfectly replacing traditional OTA operational amplifiers, and meets the overall design requirements of pipeline ADC.

In this paper, based on the SMIC 40nm process in SMIC International, a 14-bit 20MSPS high-precision analog-to-digital converter with a pipeline architecture of 2.5bit-2.5bit-2.5bit-2.5bit-2.5bit-2.5bit-2bit is designed using a ring amplifier at a 1.2V power supply voltage.The simulation results of the circuit built by Cadence show that the ADC meets the required indicators.

2. CIRCUIT IMPLEMENT

2.1 ADC Circuit Structure

The 14 bit 20 MSPS pipeline ADC designed in this article adopts a pipeline structure after considering redundancy bits. The sub level structure is shown in Figure 2.1. The pipeline ADC is mainly composed of six sub conversion stages and one level Flash-ADC, digital circuits, and clock circuits. Sub-ADC adopts Flash-ADC structure. MDAC (Gain Digital to Analog Conversion Unit): A new ring amplifier structure is used to perform digital to analog conversion of sub ADC outputs and amplify residual signals.
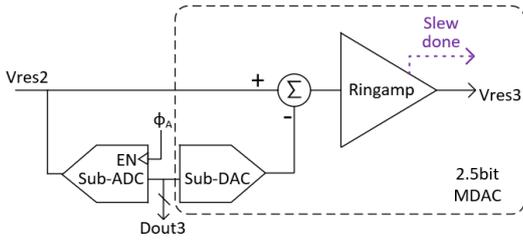
Figure. 2.1 The sub level structure of pipeline ADC

Each stage of the pipeline ADC consists of Sub-ADC, Sub-DAC, and subtraction operational amplifier circuits, wherein the MDAC consists of a switched capacitor circuit and an operational amplifier. The slew done signal generated by the dead band voltage of a fully differential ring amplifier is used to convert the input analog signal into an output that is close to a digital signal. The generated signal is used to start the next level ADC circuit, reducing the impact of clock jitter signals on each level of the waterline circuit, and improving the signal to noise ratio of the ADC.

2.2 Ring amplifier

2.2.1 Ring oscillator structure

The structure of a ring amplifier is derived from a ring oscillator, which has a similar circuit structure to that of a ring oscillator. A ring oscillator composed of a three-level CMOS inverter is shown in Figure 2.2.
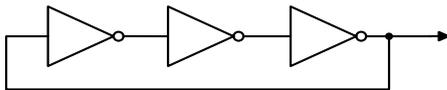

Figure. 2.2 Ring oscillator structure

The amplification function of a ring amplifier is modified based on a ring oscillator, which is created by dividing the oscillator into two signal paths and embedding different offsets in each path[5].When placed in switched capacitor feedback, a set of internal mechanisms generates stability and allows the oscillator to act as an amplifier.

2.2.2 Switched capacitor ring amplifier

Although the basic structure of a ring amplifier is simple, its operating principle is relatively complex. The traditional switched capacitor ring amplifier model is shown in Figure 2.3, when the loop amplifier is placed in a switched capacitor feedback structure and the dead band voltage V_{DZ} is equal to 0V, the loop amplifier is functionally identical to a three-level inverter loop oscillator.
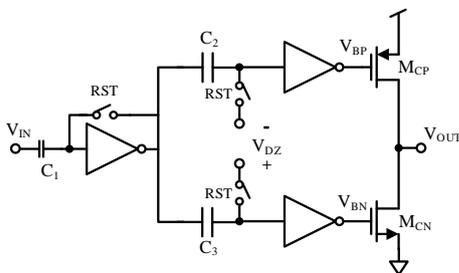

Figure. 2.3 Switched capacitor ring amplifier structure

When the deadzone voltage increases, the output waveform is stably formed at the intersection of the deadzone, and both the tubes M_{CN} and M_{CP} are closed. At this point, the effective gain through the oscillator is zero, so the average gain and oscillation frequency decrease. Under certain critical dead band

sizes, this effect reduces the unit gain bandwidth of the oscillator below the frequency corresponding to a 180 ° phase delay. When the deadzone is set large enough to produce stable settling, a ring amplifier is used as an amplifier, which is particularly suitable for working in small size, low voltage environments, is not affected by power supply voltage drops, and can be amplified with rail to rail output swing.

2.2.3 Fully differential ring amplifier

In order to fully utilize the high gain characteristics of a ring amplifier while addressing the limitations of a single ended structure, Figure 2.4 presents a fully differential ring amplifier.In order to make the ring amplifier fully differential, we replaced the first stage of two self biased ring amplifiers with a single differential pair[6].Based on an ordinary fully differential ring amplifier, a novel fully differential ring amplifier using a reverse parallel structure and adding a clock generation and control circuit is proposed.
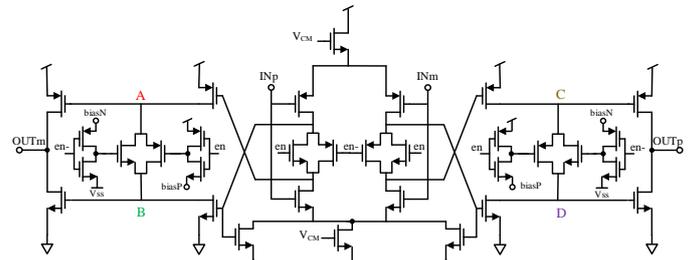

Figure. 2.4 Fully differential ring amplifier structure

The improved fully differential ring amplifier is based on a resistor self biasing differential structure, with the deadzone voltage control circuit changed from a resistor to an AP-CMOS structure[7].An en control terminal is added to the gate of the AP-CMOS, making it an adjustable CMOS control terminal between nodes A, B, C, and D. In the off state (en=0), these CMOS structures act as open switches, and the circuit is closed. In the open state (en=1), the circuit is of a working AP-CMOS ring amplifier structure. The CMOS structure operating in a floating battery state increases the maximum overdrive voltage of the device, reducing the size required for the circuit, thereby reducing parasitic effects within the ring amplifier, and improving the speed and efficiency of the overall ring amplifier.When using a smaller process size of 40 nm, the parasitic effects inside the ring amplifier will be significantly reduced. The size of each stage of the ring amplifier transistor is different, and the size of the first stage inverter A1 needs to be set large enough to increase the non dominant pole frequency. The transistor sizes for each stage from A1 to A3 are sequentially reduced to ensure sufficient phase margin for the ring amplifier.The transistor sizes of each amplifier stage of a fully differential ring amplifier are shown in Table 1.

Table 2. Device parameters of ring amplifier

RAMP	NLength	NWidth	PLength	PWidth
A1	40nm	1um	40nm	3.5um
A2	40nm	1um	40nm	3um
A3	500nm	200nm	500nm	900nm

2.3 MDAC circuit

2.3.1 Capacitance flip MDAC

The function of MDAC is to subtract the input signal from the result obtained by Sub-ADC, then multiply the difference to

obtain the residual signal of the ADC at the current level, and then transmit the signal to the input terminal of the next level.Figure 2.5 shows the structure diagram of a 1.5bit capacitor flip type MDAC[8].
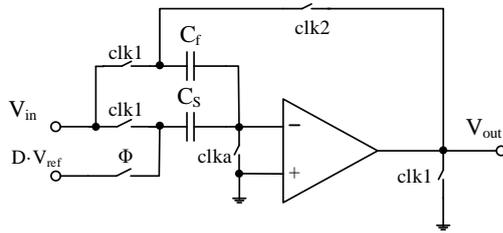

Figure. 2.5 1.5bit capacitor flip type MDAC

For capacitive flip over type MDAC, the capacitance C_s serves as both a sampling capacitor and a feedback capacitor, which is a loop feedback capacitor. The total sampling capacitance is $C_s + C_f$, and its feedback coefficient is:

$$\beta = \frac{C_f}{C_s + C_f} \qquad (1.1)$$

During the operation of the MDAC circuit, both the sampling phase and the holding amplification phase capacitances satisfy the charge conservation principle. At this time, the operational amplifier gain is and ignores the finite gain and parasitic capacitance effects. The transfer function of the capacitance flip type MDAC is:

$$V_{out} = \frac{A}{1 + A\beta}\left(V_{in} - D_i \cdot V_{ref}\frac{C_s}{C_s + C_f}\right) \qquad (1.2)$$

2.3.2 Overall structure of MDAC

The 14-bit pipelined ADC in this article adopts a capacitance flip type 2.5bit per stage MDAC structure, which reduces the impact of non ideal factors and reduces the difficulty of amplifier design. The overall circuit structure of the MDAC is shown in Figure 2.5.
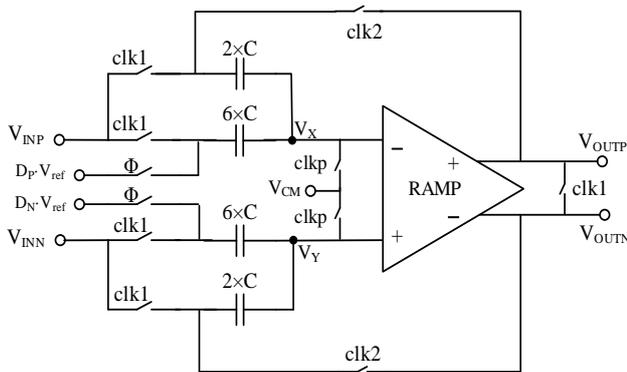

Figure. 2.6 Overall structure of MDAC circuit

Due to the 2.5bit redundant bit design of the MDAC, its interstage gain is 4.The MDAC composed of a fully differential structure has six sampling capacitors and two feedback capacitors on each side. It operates in the sampling and holding amplification phases using a bidirectional non overlapping

FFT analysis of ADC output data can obtain the spectrum characteristics of ADC, and then calculate various dynamic indicators.The simulation results of ADC dynamic performance are shown in Figure 3.2 and Figure 3.3.

clock, and its operating principle is similar to that of a 1.5bit capacitor flip type MDAC.The entire capacitance of the MDAC circuit during operation satisfies the law of charge conservation.The charge transfer equations for the input nodes X and Y of the operational amplifier during the sampling phase are as follows:

$$Q_X = (V_{CM} - V_{INP}) \cdot 8C \qquad (1.3)$$

$$Q_Y = (V_{CM} - V_{INN}) \cdot 8C \qquad (1.4)$$

Assuming that the number of 1 in each input digital code is n while maintaining the amplification phase, the charge transfer equations for nodes X and Y are as follows:

$$Q_X = (V_X - V_{REFP}) \cdot nC + (V_X - V_{REFN}) \cdot (6 - n)C + (V_A - V_{OUTP}) \cdot 2C \qquad (1.5)$$

$$Q_Y = (V_Y - V_{REFN}) \cdot nC + (V_Y - V_{REFP}) \cdot (6 - n)C + (V_Y - V_{OUTN}) \cdot 2C \qquad (1.6)$$

Without considering the parasitic capacitance at the input of the operational amplifier, the charge transfer equation for the sampling phase and the holding amplification phase is derived from the charge conservation law, and the 2.5-bit capacitance flip type MDAC transfer function is:

$$V_{OUT} = \frac{A}{A + 4} \cdot 4\left(V_{IN} - \frac{n-3}{4}V_{REF}\right) \qquad (1.7)$$

In the actual MDAC design, it is necessary to consider the limited gain and input parasitic capacitance of the operational amplifier, and the actual feedback coefficient of the MDAC will become smaller. Using a high-performance ring amplifier can improve the overall accuracy of the ADC.

3. SIMULATION RESULTES

Perform stability simulation on the ring amplifier, and the simulation results are shown in Figure 3.1. From the figure, it can be concluded that the low frequency gain of the operational amplifier is about 86.79 dB, and the unit gain bandwidth is about 670 MHz. At this time, the phase margin is about 62 deg, which meets the design requirements of the operational amplifier in MDAC.
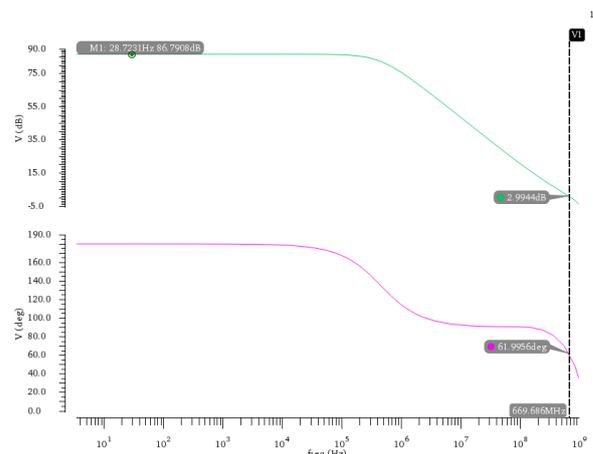

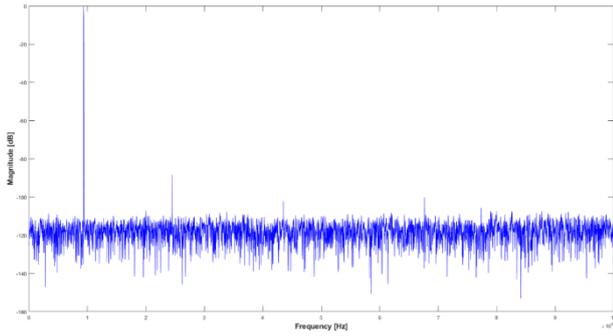Figure. 3.1 Stability Simulation of Ring Amplifier

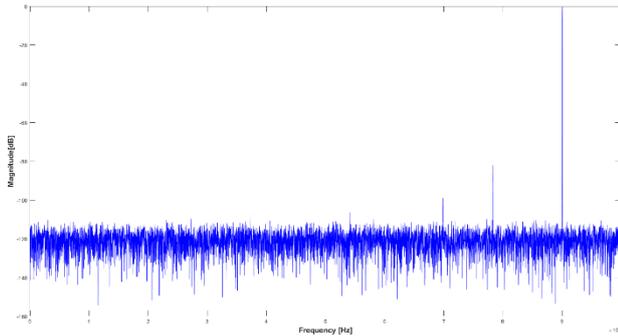Figure. 3.2 Low frequency signal FFT simulation results

Figure. 3.3 High frequency signal FFT simulation results

The input signal frequency is 0.902099 MHz, the SNDR of the pipeline ADC is 70.47 dB, the SFDR is 85.5 dB, and the effective bit ENOB reaches 11.45 bits. The input signal frequency is 9.024658 MHz, the SNDR of the pipeline ADC is 68.35 dB, the SFDR is 81.3 dB, and the effective bit ENOB is 11.07 bit.

4. CONCLUSION

Based on the SMIC 40nm process, a 14-bit pipelined ADC with a sampling rate of 20 MSPS is designed and implemented under a 1.2V power supply voltage. A new type of ring amplifier is used, which can generate a slow done signal through an AP-CMOS structure. This solves the problem of high power consumption in traditional OTAs, reduces clock jitter, and improves the gain and bandwidth of operational amplifiers in the ADC.The simulation results show that when the input

signal frequency is 0.902099 MHz, the SNDR of the pipeline ADC is 70.47 dB, the SFDR is 85.5 dB, and the effective bit ENOB reaches 11.45 bits. When the input signal frequency is 9.024658 MHz, the SNDR of the pipeline ADC is 68.35 dB, the SFDR is 81.3 dB, and the effective bit ENOB is 11.07 bit.

5. REFERENCES

[1] Keane J P, Hurst P J, Lewis S H. Background interstage gain calibration technique for pipelined ADCs[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2005, 52(1): 32-43.

[2] Yu L, Miyahara M, Matsuzawa A. A 9-bit 1.8 GS/s 44 mW pipelined ADC using linearized open-loop amplifiers[J]. IEEE Journal of Solid-State Circuits, 2016, 51(10): 2210-2221.

[3] Lin J, Paik D, Lee S, et al. An ultra-low-voltage 160 MS/s 7 bit interpolated pipeline ADC using dynamic amplifiers[J]. IEEE Journal of Solid-State Circuits, 2015, 50(6): 1399-1411.

[4] Lim Y, Flynn M P. A 100 MS/s, 10.5 bit, 2.46 mW comparator-less pipeline ADC using self-biased ring amplifiers[J]. IEEE Journal of Solid-State Circuits, 2015, 50(10): 2331-2341.

[5] Lim Y, Flynn M P. A 100 MS/s, 10.5 bit, 2.46 mW comparator-less pipeline ADC using self-biased ring amplifiers[J]. IEEE Journal of Solid-State Circuits, 2015, 50(10): 2331-2341.

[6] Lim Y, Flynn M P. A 1 mW 71.5 dB SNDR 50 MS/s 13 bit fully differential ring amplifier based SAR-assisted pipeline ADC[J]. IEEE Journal of Solid-State Circuits, 2015, 50(12): 2901-2911.

[7] Lagos J, Hershberg B, Martens E, et al. A single-channel, 600-MS/s, 12-b, ringamp-based pipelined ADC in 28-nm CMOS[J]. IEEE Journal of Solid-State Circuits, 2018, 54(2): 403-416.

[8] Oliveira J P, Goes J, Paulino N, et al. New low-power 1.5-bit time-interleaved MDAC based on MOS capacitor amplification[C]//2008 15th IEEE International Conference on Electronics, Circuits and Systems. IEEE, 2008: 251-254.

6G and Blockchain Communication: Applications, Challenges, and Predictions

Omkar Singh*
National Institute of Fashion
Technology
Patna, India

Navanendra Singh
National Institute of Fashion
Technology
Patna, India

Abhilasha Singh
National Institute of Fashion
Technology
Patna, India

Vinoth R
National Institute of Fashion Technology
Patna, India

Abstract: In today's era 5G network is being used globally, and research/technical officialdoms worldwide started work with 5G communication. But providing high-quality services (QoS) and power proficiency 6G communication is predictable in the network. This paper includes a comprehensive analysis of the development of 6G communication. In this paper, we have discussed diverse network generations with specifications and growths, detailed study of different technologies used in 6G communication has been elaborated. This chapter also deliberates on upcoming challenges in 6G communication for canvassers. Various 6G application areas are also presented in this survey. The critical focus is associated with 6G architecture which includes protocols stacks, coverage, and artificial intelligence. This paper aims to give informative direction for consequent 6G communication research.

Keywords: 6G, AI, Architecture, Wireless Network, Blockchain, Security, THz Communication.

1. INTRODUCTION

It is flattering ostensible that numerous facets of our atmosphere can be observed as a schmoozed ecosphere [1]. From broadcasting networks to the worldwide environment, highway transportation networks to stock marketplaces, and organic to communal systems, enormously interrelated and cooperating components create comparatively energetic systems in the ecosphere. These methods can be categorized as various systems [2]. Multifarious systems investigation can be reflected as a science that educates how the fundamentals of a system advance its cooperative performances, and system interrelates through its surroundings. To comprehend the comportment of a manifold system, it is necessary to initially understand its essential fundamentals and how they treat unruffled to prescribe the comportment of the whole system. Multifarious methods and their anticipated comportment often include allusions to adaptability, emergence, self-organization and development, robustness, resilience, regionalization, and speed. Just work emphasizes fundamental features of diverse methods in this perspective that can be classified as reorganized [3]. Multifarious methods, as linkages of interrelating objects, are premeditated empirically through the support of a rapid upsurge of available statistics of different areas. Concomitantly, these other areas stake numerous novel and important hypothetical queries. Portable communication systems and particularly 5G and upcoming 6G are characteristic illustrations of systems that increase quickly [4]. Grasping their difficulty in terms of heterogeneous fundamentals and significant level of independence develops a foremost obstacle, intimidating to interrupt the material rebellion. Conniving, monitoring, demonstrating, and observing the performance of such networks are the significant challenges that need to be tackled. We requisite novel models as we quickly touch from network based on fastened classified or semi-classified edifices to expose and dispersed schmoozed networks [5]. From a broadcasting network's viewpoint, the crucial issue is to acquire how to project these networks that can self-establish, adapt and improve their connections and tasks in an unremitting and healthy way to alleviate customer requests. The diverse systems detained can offer replicas, concepts, tools, and methods that permit a proper project technique to be developed to tackle this main issue [6]. Portable broadcasting systems and particularly 5G systems, and forthcoming 6G systems, are receiving extra intricate and diverse. The symbolic process of these systems through more solid placements, additional proper positions, uncountable users, and novel tools that are predictable to be announced in 6G systems such as Machine Learning (ML), Artificial Intelligence (AI), Terahertz (THz) band broadcasting and much more [7]. It is again impaired by a tendency to develop software of networking specifications and energetic adaptation of networked amenities. Multifarious schemes could create a valuable and operative contrivance proficient to prototypical the specific performance of networks [8].This paper presents diverse systems from a broadcasting network's viewpoint, enlightening issues and challenges towards 6G networks. Work focuses on 5G/6G communication systems, but the key focus is on 6G study by evolving 5G. The rest of the paper presents 6G architecture, technology, and applications.

1.1 Motivation

The world has perceived novel network generations each decade since the 1980s. Every novel generation provides better features than the previous one. In a newly evolving era, 5G has numerous innovative features. Though, it is suspected having multiple deficiencies of 5G when associated with other modern ICT substitutes. These deficiencies will be a crucial motivation for subsequent new network generation called the Sixth Generation of Mobile Communication (6G).

1.2 Research contribution

In this chapter, the following key research contributions are:

- Upcoming challenges in 6G communication for canvassers are deliberated.

- Various 6G application areas are also presented in this survey.
- The critical focus is associated with 6G architecture which includes protocols stacks, coverage, and artificial intelligence.
- A comprehensive analysis of the development of 6G communication is discussed.
- Diverse network generations with specifications and growths are elaborated.
- The main goal of this chapter is to give informative direction for consequent 6G communication research.

2. ADVANCEMENT OF MOBILE COMMUNICATION NETWORK

There has been a phenomenal advancement in mobile communication networks since the emergence of an analog communications network in the 1980s. This advancement is not a one-step process but consists of several generations with different standards, capacities, and techniques. The new generation has been introduced nearly every ten years [9].

2.1 1G-3G Network

In the 1980s 1G communication network was presented, which was premeditated for voice amenities, with a 2.4 kbps data rate. It implemented analog signal to communicate data without presentable wireless customary, prominent to numerous shortcomings including challenging hand-off, squat broadcast competence, and no safety. Associated to 1G system, two generations were built on digital intonation tools including TDMA and CDMA having data rate up to 64kbps, subsidiary voice, and SMS services. All network standards in the 2G epoch were GSM. 3G was projected in 2000 to have high data transferring speed and provide a 2 Mbps data transmission rate accessing the internet [10].

2.2 4G Network

4G was introduced in the 2000s is IP based, which is proficiently providing high-speed data rates up to 1Gbits/s. It advances phantom competence and decreases expectancy, cooperative necessities set by advanced applications such as DV), TV contented, and video conversation. Furthermore, 4G allows incurable flexibility to offer wireless services anywhere and anytime, over instinctive roaming transversely topographical restrictions of networks. LTE-A and WiMAX are deliberated as 4G ethics. LTE assimilates prevailing and novel tools such as CoMP, MIMO, and OFDM [11].

2.3 5G Network

5G communication system has just about accomplished preliminary rudimentary tests, h/w amenities erection, and calibration procedure and will rapidly be placed into commercial use. The objective of 5G is to create radical progresses in connectivity, network reliability, latency, data, and energy proficiency. Not solitary customs, the novel range of microwaves also creatively customs millimeter-wave the first period and enhances up to 10 Gb data rate. 5G smears advanced admittance tools with BDMA FBMC. Numerous developing tools are united into 5G to recover system concert:

MIMO for capability upsurge, SDN for tractability in the system, D2D for phantom proficiency, ICN for decreased traffic in the network, and system slicing for rapid utilization of several services [12].

2.4 6G Network

As we know, 5G is in the profitable distribution stage; research organizations worldwide have instigated to pay courtesy to 6G network, which is deliberated to be installed in around the 2030s. Green six generation is anticipated to improve the concert of data broadcast up to 1 Tbps and ultra-least dormancy in microseconds. It topographies THz broadcasting and latitudinal multiplexing, offering 1000 times more sophisticated capability than 5G systems. The solitary aim of 6G is to attain universal connectivity by assimilating satellite broadcasting systems and underwater broadcasting to provide global exposure. Power reaping tools and custom of innovative resources will significantly recover the network power effectiveness and understand supportable green systems [13].

3. RELATED WORK

Z. Lv et al.[14] proposed a practical elucidation to decrease signal intrusion for better transmitting associated signals. Furthermore, node info in sensors is handled by edge and fog computing, and network broadcasting excellence is also arbitrated by transmission power utilization and packet failure rate. Dual-channel design can distinctly transfer control communications, which minimizes single-channel traffic load and evades crashes amid control and sensor messages. It recovers the performance of transmission information of 6G/IoE. The utility representation of the proposed model is given as:

$$I_{utility} = \frac{I_{times}}{T_{current} - T_{first}} \tag{1}$$

The author in [14] has summarized notations and symbols used in Eq. (1).

The broadcast strategy has been professed as an arbitrarily performing object amid receiver and transmitter by E. Basar et al. [15], which reduces the excellence of acknowledged signal because of irrepressible connections of communicated radios by nearby substances. Current initiation of recon intelligent outsides in broadcast enables. Instead, system machinists control radio waves' sprinkling, response, and bending features by incapacitating undesirable properties of expected propagation in wireless. Baseband signal strength is used as follows:

$$r(t) = \frac{\lambda}{4\pi} \left(\frac{e^{\frac{-j2\pi l}{\lambda}}}{l} + \frac{R \times e^{\frac{-j2\pi(r_1+r_2)}{\lambda}}}{r_1 + r_2} \right) x(t) \tag{2}$$

The author in [15] has described specific symbols and notations mentioned in Eq. (2).

M.S. Sim et al. [16] presented DNN construction and clarified how to evaluate PDP of sub-6 GHz station, which is implemented as input of DNN, then authenticate its recital by actual surroundings-built 3D ray-outlining imitations over midair experimentations using mmWave archetype. The m-th ray is used to design this prototype is as:

$$h_m^{sub} = \sqrt{P_m} e^{j\phi_m^{sub}} F_{TX}^{sub}(\theta_m^{ZoD}, \phi_m^{AoD})$$
(3)

The author in [16] has summarized symbolizations used in Eq. (3).

An IAP-SP reduces the operational difficulty in preserving precise channel retrieval. Using the assessed station, data rate expansion delinquent is expressed and is transformed into a different stage modification exploration problem. X. Ma et al. [17] proposed a comprehensive exploration scheme to acquire an optimum broadcast rate and tolerate a tremendously high operational load. Formerly, a local exploration scheme is developed to minimize separate-stage IRS candidates who experience palpable performance defeat. The transition contribution is implemented to develop the schemes.

$$\sigma = \frac{2e^2}{\pi h^2} k_B T . \ln\left[2\cosh\left(\frac{E_F}{2k_B T}\right)\right] \frac{i}{\omega + i\tau^{-1}}$$
(4)

The author in [17] has elaborated on notations and symbols used in Eq. (4).

R. Ahmad et al. [18] proposed TPCSS technique for 6G solution. A complexity exploration is conceded to evaluate the effect of medium and laser constraints on TPCS and comparative TPA drifts. Reconnoitered laser constraints are minimum power and beat size. Variable medium constraints are peroxide elucidation attention and illustration length. TPCS is originated to be sovereign of variation in minimum power. The best-fitted data can be calculated using TPA as follows:

$$T(z) = \sum_{n=0}^{\infty} \frac{(-q_0)^n}{(n+1)^{3/2}(1+x^2)^n}$$
(5)

The author in [18] has described symbolizations used in Eq.(5).

The SERS technique is proposed by N. Sykam et al. [19] to uncover the squat attentiveness of peroxide molecules to perceive water prevention. Here, there is little cost, a quick and productive method for the manufacture of EG below microwave radioactivity in 1 min at 800W, and outstanding adsorption substantial for R6G. The consequence of adsorption procedure constraints, including contact time, pH, isotherm replicas, and kinetic reproductions on rinse

elimination underneath aqueous resolutions, were explored. Adsorption capability can be intended as follows:

$$q_e = (C_0 - C_e)\frac{V}{m}$$
(6)

The author in [19] has described notations used in Eq. (6).

V.A. Chhabra et al. [20] stated the amalgamation of merged polyaniline using PbS QDs, which was successively engaged for snap catalysis of peroxide and Rh-6G. This PbS/PANI amalgamation was organized by commissioning biochemical oxidative polymerization in the occurrence of PbS QDs. Amalgamated has been considered by X-ray precipitate deflection, FTIS, broadcast microscopy electron, and UV spectroscopy. NOMA and mmWave NOMA for forthcoming B5G and 6G network explored by L. Zhu et al.[21]. A unique characteristic of mmWave NOMA is to receive/transmit beam-founding using big staged arrays. A. Yazhar et al. [22] projected 5G NR with possible waveform edifices to estimate waveform constraints in 6G. There will be various waveform constraint selections in the forthcoming. TPs will implement these waveform constraint possibilities even though transmission to dissimilar users using optimum resource distribution pronouncements. R. Goul et al. [23] measured CVD on tasters of dissimilar attentions of R6G on Grapheme/AuNPs substrates by implementing minimum energy. Limited component simulations were done for a network using hemi ellipsoids over numerous circumstances including R6G investigative wrapper the superficial with graphene nanoparticles. Graphene exists amid nanoparticles produced by redshift in plasmatic timbre rate and grapheme reduced electric arena of external.

4. TECHNOLOGIES USED IN 6G
4.1 Spectrum Communication Technique
Spectrum is the underpinning of broadcasting. Since the increase of communication in the 1980s, we have perceived the marvelous development of spectrum possessions in each novel generation because of endless recreation for data rates. The maximum marks of 6G provide Tbps accumulated bit. Terahertz (THz) and visible light are two striking spectrums [24].

4.1.1 THz Communication
The THz is a spectral ensemble amid warm and visual costumes with frequencies from 0.1 THz to 10 THz. Excluding great immature spectrum assets, numerous sole features inspire to implement of THz for forthcoming networks [25].

4.1.2 Perceptible Light Communication
OWC are reflected as harmonizing tools for RF-based broadcasting, and frequency limit contains ultraviolet, perceptible light, and infrared spectrum. The detectable nimble range is the most auspicious continuum of OWC because of scientific improvements and the extensive implementation of LED. LED differentiates from grownup

lighting expertise because of switching to dissimilar light strength echelons rapidly, which allows encoded data in produced light in diverse ways. Perceptible light broadcasting receipts the full benefit of LED to attain double goals of the whirlwind and large rapidity data transmission [26].

4.2 Fundamental Techniques

4.2.1 Blockchain for Decentralized Security

Blockchains are distributed record-based catalogs, and transactions can be steadily recorded and restructured, deprived of central mediators. The intrinsic geographies of blockchain include dispersed interference, confrontation, and secrecy, create it perfect for numerous applications. Blockchain is reflected as a subsequent rebellion for forthcoming broadcasting tools. It assurances tougher safety topographies during broadcast since it allows multiple system objects to strongly admittance precarious. Blockchain also delivers numerous assistances in resource instrumentation and system access [27].

4.2.2 Flexible and Intelligent Material

Notwithstanding marvelous achievements in the broadcast system in previous years, a concert of customary semiconductor resources such as silicon appears to grasp its restrictions and resources by improved large-frequency and large-temperature features are imperative for ultrahigh broadcasting. Silicon and Grapheme used to project succeeding-generation broadcasting devices [28].

4.2.3 Energy Harvesting and Management

Reliable calculation strains for AI dispensation and enhancing the explosion of IoT strategies are posturing critical challenges to the power efficacy of broadcast equipment. Consequently, power-competent broadcast tools will sparkle in 6G, having shorter broadcast distances. In previous years, much exertion has been expended on power reaping and organization investigations. SR tool provides a conceivable elucidation to power delinquent, which assimilates unreceptive backscatter strategies with vigorous broadcast network [29]

4.3 New Communication Paradigm

In the 1980s 1G communication network was presented, which was premeditated for voice amenities, with a 2.4 kbps data rate.

4.3.1 Molecular Broadcast

A novel broadcast model stirred by the environment is a conceivable elucidation that implements biological signals to transmit data, mentioned to molecular broadcast (MB). In MB, biological signals are characteristically minor elements of certain nanometers to some micrometers in extent, including phospholipid vesicles and atoms, which are typically broadcast in gaseous. Associated with the radio broadcast, MB has definite benefits in micro and macro gage [30].

4.3.2 Quantum Broadcast

Quantum broadcast (QB) is a different auspicious broadcast prototype using public safety. The ultimate variance between quantum broadcast and traditional binary-based broadcast is snooping can be sensed on-site. The data is encrypted in the quantum stage with photons/quantum atoms and cannot be retrieved or replicated, deprived of interference it because of quantum ideologies, including the association of tangled atoms and unchallengeable law. Moreover, QB can recover data rates because of the superposition environment of cubits. Another striking fact of QB is its enormous perspective in the large-distance broadcast. QR is a serious strategy for a large-distance universal quantum system. It is proficient in separating QB's detachment into smaller intermediary sections and modifying photon damage and process faults [31].1G-3G Network.

5. CHALLENGES IN 6G COMMUNICATION

5.1 Dynamic Topology

The network topology in the sixth generation is predictable to be dynamic. Because every operator over its system/smart devices forms IoT systems will be associated vigorously to network and offers preeminent QoS at a contemporary instant, will severely variation dynamic networks. Drones, UAVs, drones, radar, and satellite broadcasts will be rapid stirring nodes that should support the complexity. The essential appropriately prototypical meddling subtleties so that devices can become conferral to sub-network rapidly and maintain network reliability [33].

5.2 THz Frequencies

The necessities for sophisticated data rates and sizeable spectral power proficiency execute mistreatment frequencies outside mmWave at terahertz (THz). It provides slight cell enlargement, whose range is limited up to meters. Minute cells will initiate much thicker positioning. Thicker positioning will unavoidably dynamism the investigators to reflect on novel traffic flow supervision methods, novel movement controlling, mobbing mechanism algorithms, and much more [34].

5.3 Access Network for Backhaul Traffic

Network tools will necessitate an enormous escalation in data evolution and reduce the network's access for Backhaul unable to manage it. Research enhancement at progressive bands such as D-Band, including 60 GHz ranges, is accessible and will be personified. FSO and quantum broadcast also reflect 6G backhaul to come across necessities. Drones and telluric stations may involve satellite connection with squat trajectory satellites and CubeSats for providing backhaul sustenance and upsurging ample area attention [35].

5.4 Artificial Intelligence and Machine Learning

Because of the difficulty of 6G systems, it is predictable to include AI for the efficacious and effectual process of such types of systems. AI has previously been implemented in

broadcasting in each OSI layer. Concerning 6G systems, AI is predictable to simplify their process and probable to influence the difficulty. An inescapable outline of AI as the system's superiority is predictable to create important roles in such as semantic broadcast, ML, and NN to complete supervision of broadcasts [36].

5.5 Network Functions Virtualization & SDN

SDN & NFV are the tools that depend on virtualization. The motive of these tools is to permit system projects and structure in s/w execution by fundamental s/w crosswise generic h/w platforms and strategies. SDN emphasizes unraveling system mechanism tools from network advancing tools, whereas NFV eliminates system going and former schmoozing means from the h/w for running, foremost to solarisation of system tools. System amenities adaptation, execution of functioning processes convoluted in manipulative, making, and distributing E2E services. NFV and SDN will improve adaptively in 6G networks [37].

5.6 Blockchain

A blockchain is a predictable tool to sustain in 6G systems. Subsequently, it is reflected a tool that can expressively underwrite the organization of enormous statistics that are predictable for creating and touching in 6G broadcast networks. Blockchain is attained by peer-to-peer systems and exists deprived of being accomplished by integrated authority. Blockchain can probably deliver numerous amenities, including interoperability transversely devices, big traceability data, and communications of different IoT schemes automatically [38].

5.7 Moving Networks

As users will request extensive excellence Internet facilities for moving trains/vehicles/planes, much more is enormously enhancing. Users request similar services as stationary substructure operators, and 6G systems should be capable of offering it. To tackle these apprehensions idea of dynamic networks has been presented. Active networks are a distinct group of adhoc networks. Dynamic networks, because of their volatile environment, knowledge noteworthy quality disputes can be perceived because of the rapidity of vehicles [39]

5.8 Intelligent Surfaces

Intelligent Surfaces are encouraging the developing h/w tool to recover the range and power efficacy of WSN. MIMOs implement antennae composed of large number arrays to amend their radioactivity designs over periods and occurrence for both broadcast and receiving. HSFs know the broadcast surroundings of electromagnetic surfs through programmatically organized meta-exteriors to outfit specified purposes. The RIS-associated idea encompasses RIS array components. Meta-exteriors are thin planar, artificial edifices that have newly permitted the comprehension of original

electromagnetic and visual mechanisms by engineerable amenities [40].

6. 6G APPLICATION AREAS

To enhance Internet implementation in rural regions, users must get assistance offered for everyday existence, strengthening performance and ultimately making the commercial situation feasible for machinists. Significance to rustic operators can be revealed through. Moreover, the operator's necessity to espouse the available facilities, which will enhance the petition and inspire users to increase the connectivity delivered. There are various 6G applications areas as follows [41]:

6.1 Health

In rustic zones, patients are the absence of excellent healthcare. Tele prescription signifies an appropriate elucidation to tackle this issue. Main fitness hubs can be exploited with regular stopovers; however, transfer to subordinate and tertiary fitness hubs does not precisely include custom machinery to hand over patient info. Dissimilar methods are projected where community employees assist patients in the rustic zone to interconnect with medics, implementing hypermedia equipment trusting on strengthened broadband reinforced by broadcasting [42].

6.2 Education

Education is a significant facility for delivering distant rustic zones. In-distant teaching is delivered to rural zones via virtual programs based on multimedia. Flipped-based classroom archetypal was implemented to support online teaching. Candidates were incapacitated by poor connectivity in rustic zones that delayed the enactment of classroom setup for video conferencing. A communication teaching method implementing satellite TV networks is verified for rural regions. Because of the absence of dependable Internet broadcasting, learners can mostly use custom tablets to interpret eBooks [43].

6.3 Farming

Agriculture is a distinctive application region for rural zones that can assist in Internet broadcasting or LAN. For instance, IoT-based devices are used for achieving exact irrigation in fields through imperfect obtainability of irrigation marine. Trickle irrigation uses the accurate volume of aquatic and improves sprayed fields nearby herbal ancestries [61]. Sensor nodes in IoT are projected for agriculture applications in rustic areas deprived of broadcasting or cloud communication. UAVs are implemented to gather capacities from sensors connected with IoT and communicate them to adjacent 5G base stations [44].

6.4 Financial services

ATMs development and using POS tools posture many issues in rural regions. A key challenge is the need for dependable broadcasting for performing transactions through cards by users in real time. Improved safety can be attained with

smartphones accompanied by smart card rendering to the broadcasting approach. Though convinced disputes endure being tackled, the utmost significant possibility of overdrawing with definite bank A/C is to collect the amount on a card and permit reliable ATMs to modify it while withdrawing [45].

6.5 E-Commerce

A system for associating support assemblies in rustic regions was developed. The main goal is to provision micro businesspersons in rural zones to enlarge their productions and upkeep native frugality. The system implements IVR and mVAS to permit transactions amid rustic shareholders to upkeep micro businesspersons in rural regions and permit them to increase commercial events. In the mobile-based system, it is assumed that inhabitants lack good teaching and monetary revenues to individual workstations and access to web services [46].

6.6 eGovernment

A bottom-up method is projected before founding e-supervision in rural regions, where amenities connected to agriculture, health, and education in the rural areas should spread a certain standard level earlier boarding on the e-supervision plan. Or else, an e-supervision plan would not be attained envisioned assistance as the rural region is nevertheless prepared for acceptance [47].

6.7 Other Services

Kiosks offer to employ prospects in rural regions to upkeep native frugality. Career searchers can upload their info at the kiosk. However, workers upload available openings. Similar kiosks can be implemented for selling/buying objectives. Wherever vendors can promote their harvests at kiosks. Every kiosk is connected to the server through Internet. Bus ticket

service is also considered in different areas, and ticket retailing positions in rural regions agonized from unhurried connectivity to dominant servers. Consequently, the ML approach is implemented on dominant servers for handling services [48].

7. 6G PROPOSED ARCHITECTURE

6G networks are anticipated to attain power-proficient and communally unified wireless links worldwide. However, prevailing network design cannot assure forthcoming application provisions restraints such as high throughput, low latency, and consistency. Consequently, onward observing investigation on forthcoming network outlines is essential. In this paper architecture of the 6G network is proposed and formulated till 2030 [2]. Though it is impractical to precisely demonstrate the open 6G network architecture, we presented 6 G-associated architecture. This architecture has three basic categories: enhanced stratification, intelligent connection, and universal coverage. The Stratification network view includes content-driven routing, management plane, dynamic spectrum access, fluid antenna, blockchain, new spectrum THz, and VLC. Intelligent connection is a significant view of this architecture introducing novel technologies in distributed artificial intelligence, intelligent radio, and real-time intelligent edge [49]. 6G network coverage will include space, aerial, terrestrial, undersea, and much more. HTS methods are proficient in broadband network service similar to telluric facilities, including evaluating and bandwidth. NGSO system provides minimum-latency, large-bitrate worldwide network access, and numerous satellite collections to instigate commercialization. LEO networks have been definitude in concept and simulation surroundings with laser and RF mechanisms to deliver low latency broadcast than terrestrial [50].
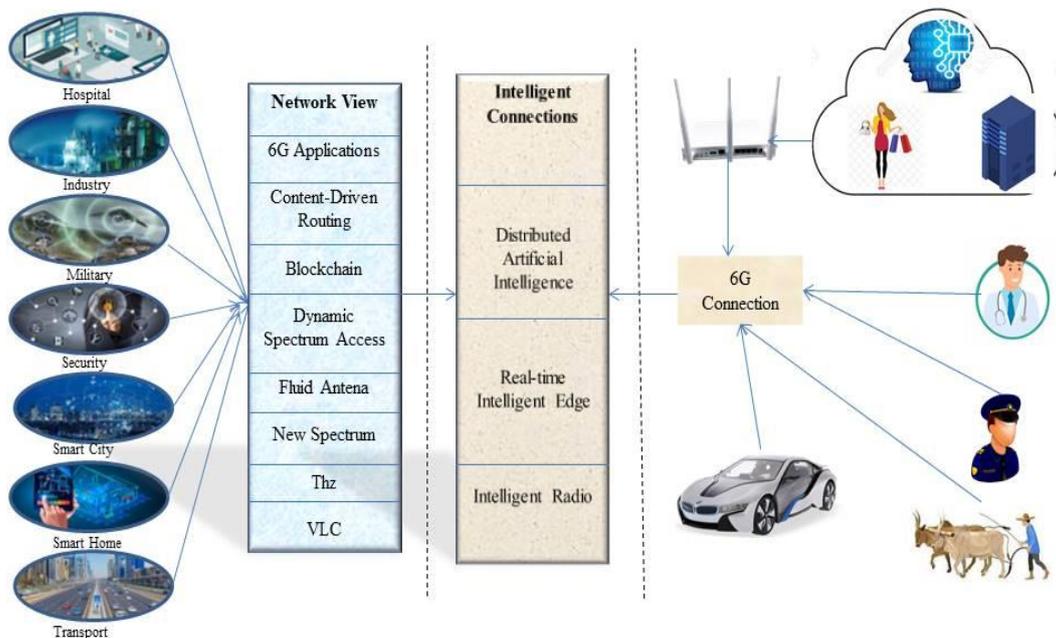

Fig. 1. Network Architecture for 6G

8. CONCLUSION

Nearly exponential upsurge in broadcast data, particularly multimedia statistics and quick propagation of all intelligent devices for subsequent broadcast development towards 6G. 6G communications are encouraging, a significant upsurge in QoS, and a supportable future. In this chapter, we have discussed a detailed survey of 6G communications. We instigate with development of network generations from 1G to 5G, which designates the expansion drift of 6G to a certain level. Detailed studies of different technologies used in 6G communication have been elaborated. This chapter also deliberates on upcoming challenges in 6G communication for canvassers. Various 6G application areas are also presented in this survey. The critical focus is associated with 6G architecture which includes protocols stacks, coverage, and artificial intelligence. The goal of this chapter is to give informative direction for consequent 6G communication research.

9. REFERENCES

[1] Nawad, S.J, Sharma, S.K., Wyne, S. et al.: Quantum Machine Learning for 6G Communication Networks: State-of-the-Art and Vision for the Future. DCN. **7**,46317-46350 (2019)

[2] Huang, T., Yang, Wu. and Wu, J.: A Survey on Green 6G Network: Architecture and Technologies. **7**, 175758-175768 (2019)

[3] Vishwanathan, H. and Mogensen P.E.: Communications in the 6G Era. GCA. **141**, 57063-57074 (2020)

[4] Sergiou, C., Lestas, M., Antoniou, P. et al.: Complex Systems: A Communication Networks Perspective Towards 6G. Complex Networks Analysis and Engineering in 5G and Beyong Towards 6G. **8**, 89007-89030 (2020)

[5] Yaacoub, E. and Aloumini, M.S.A.: A A Key 6G Challenge and Opportunity—Connecting the Base of the Pyramid: A Survey on Rural Connectivity. CNA. **108**, 533-582 (2020)

[6] Wang, J., Qiu, C., Mu, X. et al.: Ultrasensitive SERS detection of rhodamine 6G and p-nitrophenol based on electrochemically roughened nano-Au film. Talanta. **19**, 1-40 (2019)

[7] Rappaport, T.S., Kanhere, O., Mandal, S. et al.: Wireless Communications and Applications Above 100 GHz: Opportunities and Challenges for 6G and Beyond. millimeter-wave and terahertz propagation, channel modeling and applications. **141**, 78729-78757 (2019)

[8] Haas, H., Yin, L., Chen, C. et al.: Introduction to indoor networking concepts and challenges in LiFi. Journal of Optical Communications and Networking. **12**, 190-203 (2020)

[9] Lin, L. and Meng, W.: Convolutional-Neural-Network-Based Detection Algorithm for Uplink Multiuser Massive MIMO Systems. EUSPN.**8**,64250-64265 (2020)

[10] Balasubramanian, K. and Swaminathan, H.: Highly sensitive sensing of glutathione based on Fӧrster resonance energy transfer between MoS2 donors and Rhodamine 6G acceptors and its insight. Sensors and Acuators. **17**,1-41 (2017)

[11] Barzan, M. and Hajiesmaeilbaigi, F.: Investigation the concentration effect on the absorption and fluorescence properties of Rhodamine 6G dye. Optik. **159**,157-161 (2019)

[12] Dai,Y., Fei, Q., Shan, H. Y. et al. :Determination of Er3+ using a highly selective and easy-to-synthesize fluorescent probe based on Rhodamine 6G. Arabian Journal of Chemistry. **19**, 1-27 (2019)

[13] Manuel, M.L.R., Cantu, S., Lopez, E.P. et al.: Evaluation of calcium oxide in Rhodamine 6G photodegradation. Catalysis Today. **17**,1-17 (2017)

[14] J.Lv and Kumar, N.: Software defined solutions for sensors in 6G/IoE. Computer Communications. **20**, 1-11 (2020)

[15] Basar, E., Rosny, J.D., Aloumini, M.S. et al.: Wireless Communications Through Reconfigurable Intelligent Surfaces. Computer Communications. **7**, 116753-116773(2019)

[16] Sim, M.S., Park, S.H. and Chae, C.B.: Deep Learning-Based mmWave Beam Selection for 5G NR/6G With Sub-6 GHz Channel Information: Algorithms and Prototype Validation. Artificial Intelligence for Physical-Layer Wireless Communications. **8**, 51634-51646 (2020)

[17] X.Ma, Chen, Z., Chi, Y. et al. :Joint Channel Estimation and Data Rate Maximization for Intelligent Reflecting Surface Assisted Terahertz MIMO Communication Systems. Computer Networks. **8**,99565-99581 (2020)

[18] Ahmad, R., Rafique, M.S., Ajami, A. et al.: Influence of laser and material parameters on two photon absorption in Rhodamine B and Rhodamine 6G solutions in MeOH. Optik - International Journal for Light and Electron.**183**, 835-841 (2019)

[19] Sykam, N., Jayram, N.D. and Rao, G.M.: Exfoliation of graphite as flexible SERS substrate with high dye adsorption capacity for Rhodamine 6G. Applied Surface Science. **471**, 375-386 (2019)

[20] Chhabra, V.A., Kaur, R, Walia, M.S. et al.: PANI/PbS QD nanocomposite structure for visible light driven photocatalytic degradation of rhodamine 6G. Environmental Research. **20**, 1-32 (2020).

[21] Zhu, L., Xia, X. and Wu, D.O.: Millimeter-Wave Communications With Non-Orthogonal Multiple Access for B5G/6G. Millimeter-Wave Communications: New Research Trends And Challenges. **7**, 1161223-116132 (2019)

[22] Yazar, A. and Arslan, H.: A Waveform Parameter Assignment Framework for 6G With the Role of Machine Learning. VTS. **1**, 156-172 (2020)

[23] Goul, R, Das, S., Liu, Q. et al.: Quantitative analysis of surface enhanced Raman spectroscopy of Rhodamine 6G using a composite graphene and plasmonic Au nanoparticle substrate. Carbon. **111**, 386-392 (2017)

[24] Ardakni, A.G. and Rafiepour, P.: Random lasing emission from WO3 particles dispersed in Rhodamine 6G solution. Physica B: Physics of Condensed Matter.**18**, 1-16 (2018)

[25] Haroon, M., Wang,L., Yu, H. et al. :Synthesis of carboxymethyl starch-g-polyvinylpyrolidones and their properties for the adsorption of Rhodamine 6G and ammonia. Carbohydrate Polymers. **186**, 50-158 (2018)

[26] Kang, H., Fan, C., Xu, H. et al.: A highly selective fluorescence switch for Cu2þ and Fe3þ based on a new diarylethene with a triazole-linked rhodamine 6G unit. Tetrahedron.**74**, 4390-4399 (2018)

[27] Malfatti, L.,,Suzuki, K.,Erker, A. et al. :Photoluminescence of zinc oxide mesostructured films doped with Rhodamine 6G. Journal of Photochemistry and Photobiology A: Chemistry. **18**,1-14 (2018)

[28] Tang, F., Kawamoto, Y., and Kato,N. et al.: Future Intelligent and Secure Vehicular Network Toward 6G: Machine-Learning Approaches. CAN.**19**, 1-16 (2019)

[29] Viboonratansari, D., Pabchanda, S., and Prompinit, P.: Rapid and simple preparation of rhodamine 6G loaded HY zeolite for highly selective nitrite detection. Applied Surface Science. **18**, 1-33 (2018)

[30] Wang, Z., Zhang, Q., Liu, J. et al.: A twist six-membered rhodamine-based fluorescent probe for hypochlorite detection in water and lysosomes of living cells. Analytica Chimica Acta. **1082**,116-125 (2019)

[31] Wu, W.N., Wu, H., Zhong, ,R.B. et al. :Radiometric fluorescent probe based on pyrrole-modified rhodamine 6G hydrazone for the imaging of Cu2+ in lysosomes. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. **212**, 121-127 (2019)

[32] Lu, Y. and Zheng, X. :6G: A survey on technologies, scenarios, challenges, and the related issues. Journal of Industrial Information Integration. **19**, 1-52 (2020)

[33] Wei, Y., Peng, M. and Liu, Y.: Intent-based networks for 6G: Insights and challenges. Digital Communications and Networks. **5**, 1-11 (2020)

[34] Sen, P., Pados, D.A., Batalama, S.N. et al. :The TeraNova platform: An integrated testbed for ultra-broadband wireless communications at true Terahertz frequencies: Computer Networks.**179**,1-11 (2020)

[35] Fu, Y., Doan, K.N. and Quek, T.Q.S.: On recommendation-aware content caching for 6G: An artificial intelligence and optimization empowered paradigm. Digital Communication and Networks. **10**, 1-11 (2020)

[36] Zhao, Y., Zhao, J., Zhai, W. et al. :A Survey of 6G Wireless Communications: Emerging Technologies. Computer Networks. **20**,1-10 (2020)

[37] Zhou, Y., Liu, L., Wang, L., et al.: Service aware 6G: An intelligent and open network based on convergence of communication, computing and caching. Digital Computer Networks**. 11**, 1-11 (2020)

[38] Xu, H., Klaine, P.V., Onireti, O. et al.: Blockchain-enabled resource management and sharing for 6G communications. DCN**. 15**,1-13 (2020)

[39] Long,O., Chen, Y., Zhang, H. et al. :Software Defined 5G and 6G Networks: a Survey. Mobile Networks and Applications. **5**, 1-21 (2019)

[40] Akyyildiz, J.F. and Nie, S. :6G and Beyond: The Future of Wireless Communications Systems. WCS.**141**,1-36 (2020)

[41] Huang, T.,Yang, W., Wu, J., et al. :A Survey on Green 6G Network: Architecture and Technologies., GIT.**7**, 175758-175768 (2019)

[42] Shah, S.H.A., Balkrishan, S., Xin, L. et al.: Beamformed mmwave system propagation at 60GHz in an office environment. ICC**. 11**,1-7 (2020)

[43] Rahman, T.A., Aziz,O.A., Hinda, M.N. et al :Channel Characterization and Path Loss Modeling in Indoor Environment at 4.5, 28, and 38 GHz for 5G Cellular Networks. IJAP. **2018**, 1-14 (2018)

[44] Mitra, R.N. and. Agarwal, D.P. :G mobile technology: A survey. ICT Express. **1**, 132-137 (2015)

[45] Alsharif, M.H., Kelechi, A.H., Albreem, M.A., et al.: Sixth Generation (6G) Wireless Networks: Vision, Research Activities, Challenges and Potential Solutions. MDPI. **12**,1-21 (2020)

[46] Elmeadawy, S. and Shubair, R.M., :6G Wireless Communications: Future Technologies and Research Challenges. ICECTA. 1-5 (2019)

[47] Saad,W.,.Bennis, M. and Chen, M. :A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems. Computer Science. 21,1-26 (2019)

[48] Zhang,Z., Xiao, Y., Ma, Z., et al. :6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies. EUSPN.**14**, 1-14 (2019)

[49] Huang, T., Yang, Wu., Wu, J., et al.: A Survey on Green 6G Network: Architecture and Technologies. GREEN INTERNET OF THINGS.**7**,1-14 (2019)

[50] Yang, P., Xiao, Y., Xiao, M., et al :6G Wireless Communications: Vision and Potential Techniques. Licensed and Unlicensed Spectrum For Future 5g/B5g Wireless Systems.**33**, 70-75 (2019)

A Survey on Blockchain-Based Supply Chain Management in Agriculture

Manikantha K
Assistant Professor, Department of Computer Science and Engineering
BNMIT
Bengaluru, India

Abhay Nataraj
Department of Computer Science and Engineering
BNMIT
Bengaluru, India

Mohammed Affan
Department of Computer Science and Engineering
BNMIT
Bengaluru, India

Calden Michael D'Souza
Department of Computer Science and Engineering
BNMIT
Bengaluru, India

Charan Kannati
Department of Computer Science and Engineering
BNMIT
Bengaluru, India

Abstract: India is one of the fifth largest economies in the world by nominal GDP with 70 per cent of the rural households still depending on agriculture as their primary source of livelihood. India is responsible for 7.39 per cent of total world agricultural production. But the agricultural supply chain used in India is inefficient and caters less to the producer. Farmers who are the producers get very less profit from selling their produce. It is estimated that about 80 percent of the profits are secured by the middlemen who have significant control over the agriculture trade in the country. The lack of communication between different stages in the chain and data transparency is also one of the major reasons why the current supply chain in India is flawed. Proper communication is essential in any system to avoid the spread of any misinformation and at each stage in the supply chain sensors are required to share data with the users so that information is collected and made available to everyone concerned with the particular transaction of goods. The combination of blockchain and IoT in the agricultural sector can provide a better supply chain system from the ground up which addresses communication and transparency issues suffered by the latter systems.

Keywords: blockchain, IoT, agriculture, India, supply chain

1. INTRODUCTION

Agriculture is a primary contributor to India's gross domestic product (GDP). According to the yearly stats provided by the government for Financial Years 2017-2021, we can notice that there is an increase in employment for agricultural jobs. In 2017 the employees in the agricultural sector counted 145.66 million and despite the outbreak of the coronavirus pandemic, the number of employees in 2020 increased to 145.88 million, and the next year India witnessed an even steeper increase with 151.79 million employees joining the agriculture sector. There was about a 4.21% increase in the number of employees from the year 2017-2021. Since it is one of the most active sectors in India it is vital that they have an efficient supply chain that is efficient, fair and rewarding.

A supply chain is a network of facilities helpful in getting the product from production to its eventual distribution to consumers. The supply chain plays a vital role in any industry as it helps the company/organization in managing inventory so as to stock the optimal amount of supplies needed to hit their sales target, better forecasting to predict their future sales and profit margins and improve relationships with suppliers. The agricultural supply chain in India has four different entities, producers(farmers), processors, distributors and retailers.

The supply chain currently in use has a lot of issues. Some of these issues are mentioned below:

-

- Middlemen in the system pocket most of the profits which makes it less rewarding for the producers and ultimately demotivates them, resorting to growing low quality crops as they could get the same amount of profits from producing lower quality crops rather than spending more money on growing and maintaining higher quality crops.
- Lack of communication between different levels in the supply chain leads to poor information exchange and knowledge transfer. The middlemen in the chain have the most knowledge. The market information known by the distributor is not shared with the producer. This allows the distributor to make more profits as he sells the product directly to the consumer and this information would not be revealed to the producer.
- Lack of proper storage facilities among the producers leads to wastage of harvested crops. Certain crops require constant controlled temperatures if such storage facilities aren't made available then the producer will incur losses.

A Blockchain is a decentralized database with a special digital file called the ledger which is available to every node in a network. In a Blockchain The transactions are recorded and distributed, The nodes in the Network do not hold the privilege to edit this special file. Thus we can say that the ledger is immutable, thereby securing the transactions that are recorded,

they cannot be altered, deleted or destroyed. This is more of an efficient way to allow companies to complete their transactions without the requirement of third parties.

The foundation set by blockchain makes the relationship between vendors and suppliers more transparent and trustworthy. As mentioned in [8] trust and decentralization are the two most important requirements in SCM which can be delivered by implementing blockchain in it. As mentioned in the discussion earlier issues related to SCM are related to distrust, information exchange, commitment, and transparency. Blockchain addresses these shortcomings using the following features:

1. Immutable data documentation
2. Distributed Storage
3. Authorized user permissions

2. PURPOSE

In all of the subjects and motives put forth by the articles Referred, The system proposed by the authors is inadequate to solve all the problems faced in Agricultural Supply chain management, To address problems like Management, Traceability and Storage of data, various solutions are presented and putting it all together and deriving a system capable of eliminating any more shortcomings which is our goal.

3. LITERATURE REVIEW

In this section, we analyze the previous systems proposed in order to understand the current situation of blockchain in the field of agriculture. The authors of [1] portray a system to ensure the traceability of Agri-food items from producers to consumers coupled with a reputation system, the trust of an entity in the blockchain is based on the trust values stored in it. The higher the trust value of the entity more is its reliability. This system provides a trading and delivery mechanism that ensures product authenticity by avoiding the manipulation of information by the Logistic Company(LC) during the transaction. Efficient tracking of the product is explained in [2] by using a unique digital identification tag, location and navigation sensors with every item that helps keep track of them. Sensors can send the real-time data collected continuously to the application by making use of IoT architecture to provide a clear path for data to travel from the sensor to the application for processing. Digital code scanners installed in the entry and exit points of the warehouses inform the application about the product's status, this gives the entities in the supply chain more clarity about the product being shipped and its location in real-time. A more theoretical understanding of implementing blockchain-based agriculture systems is given in the paper [3]. The paper mainly talks about the communication problems between different levels in the supply chain.Market research done by the higher tier entities(retailers, distributors, etc) is not shared with the lower tier entities(producers) and minimal knowledge transfer between the higher and lower level entities in the chain may lead to the disproportionate supply of products which causes unnecessary wastage or shortage of goods. These communication issues in the supply chain are to be solved by using a blockchain architecture that stores each transaction in it with a timestamp so as to avoid the manipulation of data by third parties. The transparency of the system is ensured by the miners and forgers who gather all the transactions taking place in the system and add it to the chain so that the data is made publicly available.There are many individuals in the network working towards their personal goals and these can be achieved by using smart contracts. Smart Contracts are implemented in

such a way that the system isn't biased to any one tier in the supply chain but rather a solution to achieve a level playing field. With the decreasing quality of crops being grown on farms due to the use of pesticides, the demand for organic crops has risen in recent times. [4] talks about introducing organic crops with valid and authentic certifications, into a public blockchain. This will ensure better food quality for the consumer and thereby an increased demand for the product. The tracking mechanism of the product in the chain involved the usage of trackers on the container as a whole rather than each product present in it. Thereby the container was tracked rather than the product. The methodology talks about the development of a DApp, wherein smart transaction models were developed on solidity and a QR code model was implemented to validate the products. QR code acted as the point of contact between the physical and digital worlds. The product contract could be deployed only by the farmers in order to indicate a physical product. They need to maintain a minimum reputation level in order to issue a product contract. Each product contract deployed, an address will be returned which will be used to generate a QR code that will help identify the product physically. The history of the product would be added to each smart contract within each transaction. They were then updated with new owners irrespective of whether it was sold. Smart contracts ensure trust between users as it regulates the rules. A major highlight of this paper is its reputation system. At the start of the certification process, persons are identified to take part in the validation process. In order to gain a reputation, the farmer needs to place a request which needs to be validated by agriculture officers and other related people. The motivation for others to take part in validating the farmer's transaction is that they are rewarded with reputation points. All the events would be stored in the blockchain as soon as the validation process would return a success. The paper [5] summarizes the concept of using NFC(Near Field Communication), RF ID (Radio Frequency Identification) and QR codes in product delivery which is to ensure the safety of the product and also that either wholesaler, consumer or retailer knows the complete information and details that they must know about the product. By knowing the characteristics of each of the enhancing features like that mentioned, we can enhance safety in the system. The NFC tag is used to track and trace the product in its current path from the manufacturer till it reaches the consumer. Similarly, in the case of RFID tags, they are used to track the product but also the consumer can detect counterfeit. The next module is the QR code, this is another similar implementation of NFC and RFID. QR Codes also ensure that the product is not duplicated hence maintaining consistency in product delivery. [6] conveys the entirety of the management module of medical equipment, and the difficulties faced by the manufacturers, distributors, suppliers and consumers. This article also delivers the standard requirement to price highly consumable stock and the stock that is stored for a long time unused. Comparing this to the various agricultural products, their management and pricing is also an important aspect that needs to be taken into consideration. Various distributors have facilities not large enough to store the incoming stock which may be returned or sold, this paper gives us a solution to deal with this problem. To effectively cost the product is to the consumer so as to not transmit the loss in the chain of Retailers, Suppliers, Dealers, Distributors, manufacturers and producers, Effective management of these goods also plays a vital role in the pricing of the product along with the charges of its storage, production and processing. In the paper [7] blockchain can be coupled with IoT, to monitor the process and the condition of

goods. Whenever goods are transported by farmers, the qualitative and quantitative details are measured and collected via a form and stored in the blockchain. The obtained parameters are verified by the smart contract. Farmers and consumers are informed about the condition of the goods measured using IoT. Sensors placed at each stage of SCM(in Storage and Transportation vehicles). Changes in market price are also incorporated in the blockchain, when the products are sold, information about the quantity and the profit is added to the blockchain for farmers' understanding of the market scenario.

4. CHALLENGES

Considering the technologies being used in various systems proposed, this survey requires a basic understanding of several domains such as Blockchain, IoT, Computer vision, Agricultural Science, Operations management, logistics, procurement, and information technology. Having such knowledge beforehand helps navigate through all the papers referred to. There are many choices for papers with similar discussions on the chosen topic which led to selecting only those papers which helped give a clearer picture of the said matter. All of the papers that we chose had relevant points to understanding the flaws of the system and how to put right every wrong turn, some papers talk about correcting the system and a few list improvements to the suggested system. we could not pick one way to perfect the system which led us to choose one of the many paths that were laid out before us. Thus we chose to eliminate only the difficulties that might cause the producer i.e the farmer, to face the main issue to address as mentioned in many of the papers we reviewed.

5. FIGURES/CAPTIONS

a) Current System: With reference to the earlier discussion current system of SCM is referred to in "Fig. 1".
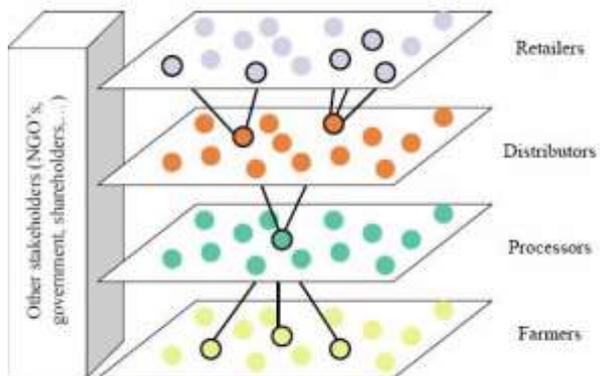

Fig. 1. Note: Systematic flow of Supply Chain in Agriculture from a processor point of view. Reprinted from [9] the issue, pp. 4.

b) Proposed System: As mentioned in [2] the proposed framework of the system with implementing blockchain in Supply Chain Management.
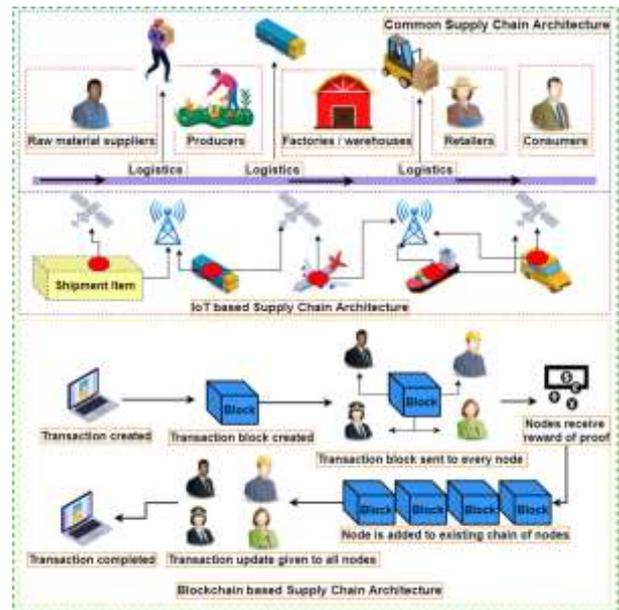

Fig. 2. Note: Blockchain and IoT-based Supply Chain Management Framework Reprinted from [2].

6. CONCLUSION

This survey includes referencing various papers which bring about the detailed architecture of Supply Chain Management and also understand its shortcomings. Some of the research papers like [1], [3], [4], [7] propose a solution to various issues in Supply Chain Management(SCM) like distrust, communication issues, commitment issues and consistency. Addressing all these issues we have concluded that blockchain along with IoT could be a possible solution for the inadequate existing SCM. This proposal is efficient and reliable. Throughout the discussion, in the paper, we talk about various blockchain properties like transparency, decentralization and security addressing those issues mentioned above about the SCM. This survey also follows up with the implementation of the solution with detailed research.

7. REFERENCES

[1] A. Shahid, A. Almogren, N. Javaid, F. A. Al-Zahrani, M. Zuair, and M. Alam, "Blockchain-Based Agri-Food Supply Chain: A Complete Solution," IEEE Access, vol. 8, pp. 69230–69243, 2020,

[2] M. N. M. Bhutta and M. Ahmad, "Secure Identification, Traceability and Real-Time Tracking of Agricultural Food Supply During Transportation Using Internet of Things," IEEE Access, vol. 9, pp. 65660–65675, 2021,

[3] B. Hegde, B. Ravishankar, and M. Appaiah, "Agricultural Supply Chain Management Using Blockchain Technology," 2020 International Conference on Mainstreaming Block Chain Implementation (ICOMBI), Feb. 2020,

[4] B. M. A. L. Basnayake and C. Rajapakse, "A Blockchain-based decentralized system to ensure the transparency of organic food supply chain," 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE), Mar. 2019,

[5] N. N. Ahamed, P. Karthikeyan, S. P. Anandaraj, and R. Vignesh, "Sea Food Supply Chain Management Using Blockchain," 2020 6th International Conference on Advanced

Computing and Communication Systems (ICACCS), Mar. 2020,

[6] Y. Yue and X. Fu, "Research on Medical Equipment Supply Chain Management Method Based on Blockchain Technology," 2020 International Conference on Service Science (ICSS), Aug. 2020,

[7] V. Sudha, R. Kalaiselvi, and P. Shanmughasundaram, "Blockchain based solution to improve the Supply Chain Management in Indian agriculture," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Mar. 2021,

[8] S. Yousuf and D. Svetinovic, "Blockchain Technology in Supply Chain Management: Preliminary Study," 2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS), Oct. 2019,

[9] National Institute Of Agricultural Extension Management, "Training programme on Supply Chain Management in Agriculture," unpublished.

An Algorithm for Finding Equivalence in Referential Relations

Kalman Gulzhamal
doctoral student of specialty
"Information System" at
L.N.Gumilyov Eurasian
National University,
Nur-Sultan, Kazakhstan

Ilyubayev Adelzhan
Teacher,Master of Technical
Science, Abay Myrzakhmetov
Kokshetau University,
Kokshetau, Kazakhstan

Aktaeva Dilara Aidosovna
Teacher,Master of Technical
Science, Abay Myrzakhmetov
Kokshetau University,
Kokshetau, Kazakhstan,

Kasym Karlygash
Kydyrbekkyzy

Teacher, Master of Technical
Science, Abay Myrzakhmetov
Kokshetau University,
Kokshetau, Kazakhstan

Likerova Zinaida Valentinovna
Teacher, Master of Technical
Science, Abay Myrzakhmetov
Kokshetau University,
Kokshetau, Kazakhstan

Dauletbek Aigerim
Graduate Student of Specialty
"Information System" at
Abay Myrzakhmetov
Kokshetau University,
Kokshetau, Kazakhstan,

Abstract: Establishing referential relations in discourse is one of the most relevant but difficult to model problems of automatic text analysis. Reference means attributing a textual unit (linguistic expression) to a non-linguistic object (referent). The correct interpretation of the statement in the analyzed text requires solving the referent of the textual reference about the object, that is, the reference of the textual expression. In this work, a model for resolving referential relations is proposed.

The method used to solve the referential relationship is finding equivalence, that is, by calculating the degree of similarity of possible elements in the word, the similarity in the referential relationship is calculated, and as a result, the relationship value of the referent in the sequence of sentences is obtained.

These opportunities will be a new beginning for solving referential relationships in the Kazakh language. Based on the possibilities of computer technology, this direction in developing language education creates a great opportunity for extensive research and analysis at various levels of the language treasure as a bright mirror of the spiritual and material culture of the people.

Keywords: Anaphora, coreference, discourse, information object, ontology.

1. INTRODUCTION

Anaphoric, cataphoric, and coreference are the main elements of resolving referential relations.

Anaphora is a relation between sentences in which the meaning of one word or phrase is defined by another word or phrase. The first member of an anaphoric relation is called an antecedent, and the second member is called an anaphor or anaphor.

A cataphoric relation is contrasted with an anaphora where the anaphora occurs in the first clause and the antecedent occurs in the second clause. Anaphoric and cataphoric relations are formed from nouns.

The most common form of anaphoric relation is pronoun anaphora. This type of anaphora includes the third person form of the pronoun, among the pronouns, the most anaphoric functions are the classification pronoun and the reference pronoun. We can see the anaphoric function of the relative pronoun in the following examples

For example: *Труба түбіндегі жапырық тас үй – мехцех. Бұл - әншейін келешегіне қарай қойылған ат, әйтпесе нобайы түзу бір механизм жоқ.(The leafy stone house at the bottom of the pipe is a mehtzeh.*

This is a name given to the future, otherwise there is no straight mechanism.)

This syntactically complex unit that connects the first sentence and the second sentence is an anaphoric relationship, where the word **мехцех** in the first sentence is repeated with pronoun **бұл** in the second sentence.

Елжас өткен айда Бразилияда болды. Ол сол елден саған сыйлық алып келіпті. (Eljas was in Brazil last month. He brought you a gift from that country)

In the example of the first sentence, it is anaphoric to have the classificatory pronoun **Елжас** and in the second sentence the word **ол** is repeated through it in third person form and repeated with the same indicative pronoun.

Solving this problem faced by natural language is the most difficult for other languages and has not yet been fully resolved. For example, the scheme described in [6] includes mentions of people and organizations and the task of solving anaphora, while the task of solving anaphora in [7] and [8] is limited to pronoun anaphora.

Solving anaphora is a very important task, many researchers take different approaches to this problem and use different methods: traditional (syntactic and semantic) and alternative (statistical) methods, recently corpus and ontological methods are also used.

The corpus method is a collection of texts within the discipline and thus research.

Most of the similar resources available today are English-language resources. A research group from Stanford University recommends using Wikipedia to solve co-references [4]. The approach itself is based on the use of several simple filters together. The system based on this approach has now been expanded with new filters [2]. Two of the five proposed new filters use external resources such as WordNet [8], Wikipedia, and Freebase [5]. Projects such as WordNet and Freebase are suitable for the English language, which has a significant impact on research in the field of English text processing.

Solving such relations for the Kazakh language will be very complicated, however, in solving the referential relations for the Kazakh language, considering the semantic analysis model and the peculiarities of the Kazakh language, combining similar elements in the sentence, searching for corresponding elements, etc. a model for solving referential relations in the Kazakh language is created by using methods.

2. METHOD

2.1 The Architecture of Anaphora Resolution

In this method, we used the primary analysis component and the semantic analysis model, identification component, and data validation component.

This section presents the architecture of the anaphora resolution system in the Kazakh language (figure 1).
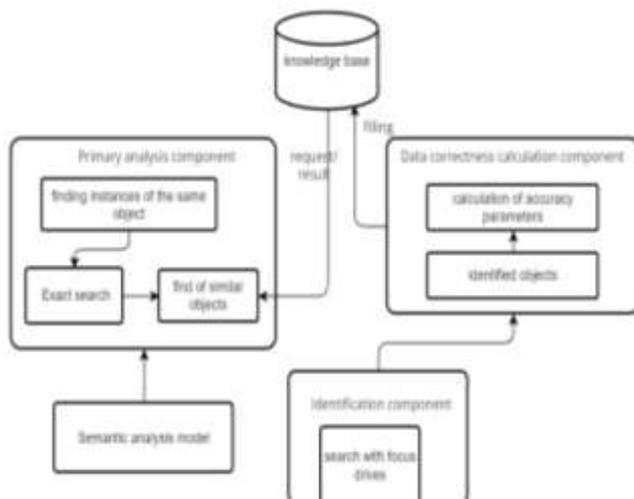

Fig.1. Reference resolution model.

The general operation of the model is as follows:
- Primary analysis. The information objects obtained from the document are sent to the primary analysis component, where checks are made for the presence of correlation and the matching of the main attributes with the DB objects according to the tuple. Information objects that have achieved correspondence with a single object of the database or a set of basic attributes are fully defined are identified.
- Identification component. The rest of the information objects fall into the identification component, where, if necessary, the collections of the closest objects of the database are filtered, which are expanded in the hierarchy of ontology classes or in other relationships of the ontology.
- Semantic analysis model. It helps to fully reveal the semantic relationship of the elements of the word, the meaning of the word, the semantic analysis model allows to speed up the work of the remaining components.

In the Kazakh language, the order of the words in a sentence is, in general, stable, not as free as in the Russian language, as a rule, the narrative is at the end of the sentence, the initial is before it, the determiner, the modifier, the complement is before the related words.

Pronouns often have the primary, complementary function in a sentence.

In this study, we take data sets from tengrinews news collection and stories of G. Mustafin: personal pronouns, demonstrative pronouns, reflexive pronouns are resolved in this dataset. A proper anaphora resolution system requires subject object matching. The proposed training dataset is defined as a subject, object, number, animate or inanimate object. This POS tagging system for Part-of-Speech Tagger for Kazakh language we use annotated corpus, Noun and Noun Phrase are selected using rules for Kazakh text.

2.2 Equivalence calculation algorithm

When searching for similar objects, candidate attributes are searched by pairwise comparison of the set of their values. To calculate the D-similarity measure, we obtained the following values.

$$a, b \in A, a \approx b, and\ e\ \alpha_a \in Dat_a\ ,\beta_b \in Dat_b, \rho_a \in Rel_a, \vartheta_b \in Rel_b$$

The similarity measure for data attributes is determined by the number of equal values of the attributes:

α_a similarity across data

$$\beta_b(\alpha_a \sim \beta_b) \Leftrightarrow (\alpha = \beta \lor \alpha \ll \beta \lor \alpha \gg \beta) \lor N^d\ and\ S^d = V_{a_a} \cap V_{\beta_a} \neq \emptyset$$

D is similar in terms of data:

$$D(\alpha_a, \beta_b) = \frac{|S^d|}{2}\left(\frac{1}{|V_{a_a}|} + \frac{1}{|V_{\beta_a}|}\right) \qquad (1)$$

Where Dat_a set of data attributes, V_{a_a} a set of information values, Rel_a a set of relationship attributes, S^d structured information (input data), α_a, β_b —co-referent-candidate.

The algorithm for finding coreference is as follows:

a and b are candidate coreferents

$$a \approx b \Leftrightarrow c_a = c_b \lor c_a < c_b \lor c_a > c_b\ and\ Atr_a^k \subseteq Atr_b^k \lor Atr_b^k \subseteq Atr_a^k$$

$$coR(x) = \bigcup_{x \in X} coR(x) \qquad (2)$$

The degree of similarity of possible candidates is calculated from the formula (1) given above, and the algorithms for calculating the coreference using the co-referent-candidate a and b given by the formula (2) are given.

The model and calculation algorithms we used in this work are not designed to solve all types of referential relations, these calculation algorithms may not work properly due to the influence of factors affecting anaphora and cataphora and their types in sentences.

2.3 Combining information objects

D is a binary relation in the set of information objects, $s^1 \text{ and } s^2$ objects, and $s^1 D s^2 \iff s^1 \text{ and } s^2$ recognized referentially identical. Relationship D is the ratio of equivalence and, therefore, it divides many O_Q into non-overstepping subset of the clusters. Equivalence classes in relation to D coincide with components of the coherence of the column in Fig. 2. All information contained in the elements of the cluster is combined in one object called a nodal object or not. The g-equivalent of the chain of objects is considered nodular. Obviously, there will always be such (Fig. 3).
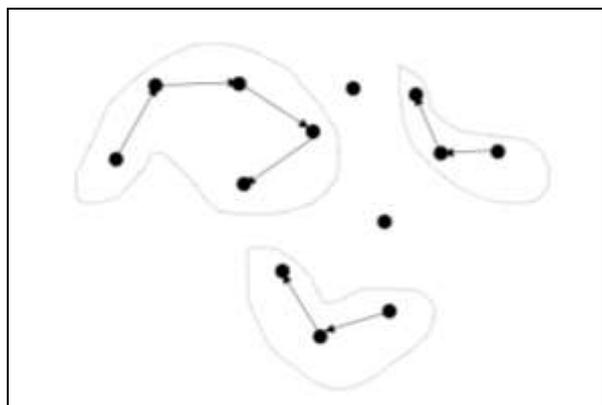

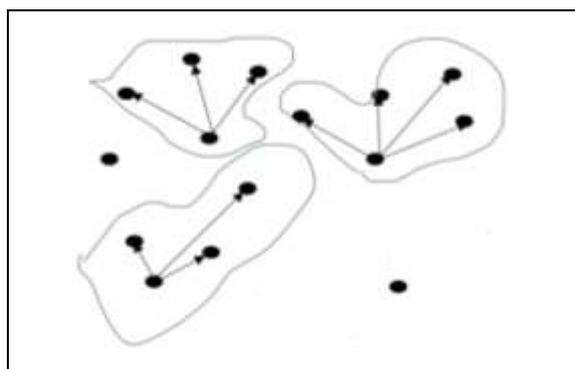Fig.2. Marking up multiple objects

Fig. 3. Node objects

Combining objects in accordance with relationships, shown in fig. 3, we achieve that the cardinality of the set of objects becomes equal to cardinalities of the set of equivalence classes with respect to relation D. It's clear, what.

$$|O_Q/D| \leq |O_Q|$$

3. EXPERIMENTAL PROCEDURE AND RESEARCH RESULT

To realize the algorithm, we used the ontology created in the subject area of computer sciences.

For practical tests of our method and obtaining the first approximate values of $f \text{ and } k$, an editor of objects with built -in mechanisms for calculating the similarity of any two selected $s^1 \text{ and } s^2$ objects and resolving reference connections in each many information objects was developed.

The selected fragment, 300 words, contain all referentially identical objects. Volume the total text is 420 words, the number of extracted objects is 30. Objects that are not included in this fragment do not fundamentally affect the result of resolving referential links. Ontological class instances have been retrieved: The result of participate (қатысу), Internet resource (интернет ресурстар), Scientific activity (Қызмет), Organization (ұйым) and Person(тұлға).

The result of scientific activity (Қызмет), class has 4 attributes: title, number, start date, completion date.

of Organization (ұйым) class has 4 attributes: e mail, abbreviation, address, founding date, organization name, organization description, phone number.

of participate (қатысу) class has 5 attributes: involved end, involved start, involved role, person involved, involved activities.

Fig. 4. Main window of the object editor

as you can see from the picture, no attributes of q² have been determined, we look for a candidate close to q² based on the above equivalence search algorithm, Obviously, since both objects are instances of the Scientific Event class, therefore the object q¹ is the equivalent of the object q².

$$\textbf{Candidate}(Q^1 Q^2)=1$$

As a result of comparing the class attributes of the ontology, it is possible to find the equivalence of the attributes, for the purposes of our research, this is called coreference, that is, the coreference of the attributes with each other.

The general results of the research work can be seen in the table below.

Table 1. the result of finding the referent

Amount of text	Number of ontological classes	Number of attributes	Referent number of attributes
300	87	156	64
245	75	123	48
345	57	103	56
Total	219	382	168

4. CONCLUSIONS

In this article, we proposed a method for solving refractive relations, considering as an example the main elements of referential relations and their differences from each other, this method was obtained based on the algorithm for calculating the degree of similarity of attributes and semantic analysis, the algorithm for solving the referential relationship based on determining the mutual semantic relationship of elements in a word and finding the degree of similarity of a potential candidate, this work plays a key role for us in the development of work on solving the refractive relationship in the kazakh language.

REFERENCES

1. Caroline V. Gasperin Statistical anaphora resolution in biomedical texts. Technical report, University of Cambridge Computer Laboratory. 2009. ISSN 1476–2986

2. Heeyoung Lee, Yves Peirsman, Angel Chang, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. 2011. Stanford's Multi-Pass Sieve Coreference Resolution System at the CoNLL2011 Shared Task. In Proceedings of the CoNLL2011 Shared Task.

3. Mitkov, R. Anaphora resolution: the state of the art, Working paper, (Based on the COLING'98/ACL'98 tutorial on anaphora resolution), University of Wolverhampton, Wolver hampton, 1999.

4. Karthik Raghunathan, Heeyoung Lee, Sudarshan Rangarajan, Nathanael Chambers, Mihai Sur deanu, Dan Jurafsky, and Christopher Manning. 2010. A Multi-Pass Sieve for Coreference Resolution. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP 2010)

5. The main page of the Freebase project. [Electronic resource] - Access mode http://www.freebase.com/

6. Ermakov A.E. Reference designations of persons and organizations in Russian-language media texts: empirical patterns for computer analysis. // Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2005". M.: Nauka, 2005.

7. Zagorulko Yu.A., Borovikova O.I., KononenkoI.S., Sidorova E.A. An approach to building a subject ontology for a knowledge portal in computational linguistics. // Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2006". M.: RGGU, 2006.

8. Official website of Princeton University. [Electronic resource]. Main page of the WordNet project. – Access mode http://wordnet.princeton.edu

9. Potepnya V.N. Resolution of pronominal anaphora in multilingual information systems. // Artificial intelligence-2006 №4 P.619-626.

10. [10] Gray A.S., Sidorova E.A. Identification of objects in the task of automatic document processing. // Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2011". M.: RGGU, 2011. S. 580-591.

11. Tolpegin P.V., Vetrov D.P., Kropotov D.A. Algorithm for automated resolution of third person pronoun anaphora based on machine learning methods. Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialogue 2006" // Ed. N.I. Laufer, A.S. Narinyani, V.P. Selegeya. - M.: RGGU, 2006

A Pose Estimation Method Combining Instance Segmentation and Point Pair Features

Yu Xin
College of Communication Engineering
Chengdu University of Information Technology
Chengdu China

Hao Peng
College of Communication Engineering
Chengdu University of Information Technology
Chengdu China

Abstract: For the traditional pose estimation method based on point pair features, both the preprocessing of scene point cloud and the construction of point pair features of scene point cloud have serious time-consuming problems, which cannot meet the needs of actual industry. Therefore, this paper proposes a pose estimation method that combines instance segmentation and point pair features. Therefore, this paper proposes a pose estimation method that combines instance segmentation and point pair features. First, use the Mask R-CNN-based instance segmentation network to obtain the location of the target object in the two-dimensional image of the scene; then, obtain the local point cloud data of the space where the target object is located from this position and the depth information of the scene; finally, the local the point cloud data is used as the scene point cloud based on point pair feature pose estimation to perform feature matching with the point cloud of the target object.

Keywords: point cloud data; point pair features; Mask R-CNN; pose estimation

1. INTRODUCTION

3D feature descriptors are key factors in 3D pose estimation, they encode the relationship between point cloud data into low-dimensional feature vectors. Global feature descriptors are constructed by using the overall geometric information of the model, such as Shape distribution feature descriptors, Spherical harmonics feature descriptors, SPR feature descriptors, etc. However, global feature descriptors are less stable in the presence of occlusions and simple object shapes. The local feature descriptor describes the local structure and shape information of the target object through the domain relationship and normal vector of the 3D point cloud, such as the 3DSC feature descriptor, the FPFH fast point feature histogram, and the SHOT feature descriptor. However, in the subsequent feature matching, local feature descriptors need to perform seriously time-consuming hypothesis verification, which will cause the matching efficiency to become very low. In order to make full use of the respective advantages of the global feature descriptor and the local feature descriptor, Drost et al. proposed the point pair feature descriptor PPF (Point Pair Feature), whose main idea is to extract point pair features from the point cloud of the target object and the scene, The pose of the object in the scene point cloud is estimated according to the relationship between the point pair features and the pose combined with the Hough voting strategy. However, the pose estimation algorithm based on PPF also has certain defects, such as being susceptible to background and noise interference and a large amount of calculation when voting on the scene point cloud. Improved PPF algorithms for everyday objects in cluttered scenes by Birdal et al. and Hinterstoisser et al. Wu et al. used the PPF algorithm to perform robotic trash bin picking and achieved a recognition rate of 93.9%. Choi et al. introduced boundary points with directions and boundary line segments into an algorithm for estimating planar industrial objects, which performed a higher recognition rate and faster speed than traditional PPF. In this paper, a point-to-feature pose estimation algorithm combined with instance segmentation is proposed. This method only selects the local point cloud data containing the range of the target point cloud as the matching scene point cloud for pose estimation, which not only reduces the number of points also predicts in advance the spatial extent of the point cloud of the target object.

2. METHOD

2.1 Original PPF Algorithm

Drost et al. used "global description, local matching" to describe the algorithm, so we can know that the algorithm can be divided into two stages. The first stage is for the model to train the global feature descriptor of the target object offline, that is, the point pair feature descriptor. The second stage is to match the feature description of the target global point pair trained in the first stage with the scene point cloud, establish a local coordinate system from the relationship between the point pair features, and calculate the three-dimensional rigid body transformation of the target object in the scene point cloud. Then, by voting on all possible pose transformations, the pose of the target object is obtained, and finally the obtained pose is optimized to complete the pose estimation of the target object in the scene point cloud. Figure 1 shows the technical flow chart of pose estimation based on point pair features.
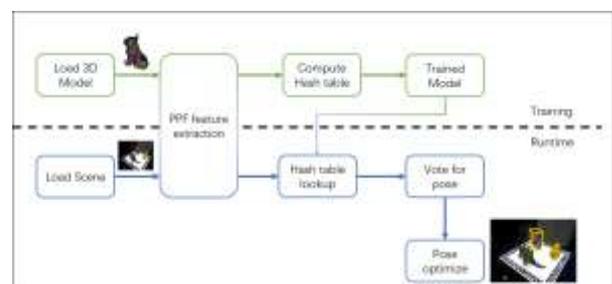

Figure. 1 Original PPF algorithm process

2.2 Our Algorithm

In this paper, we propose a method combining deep learning for instance segmentation and point-to-feature based voting. This method is divided into two stages. In the first stage, the advanced Mask R-CNN network is used to find the target

object in the scene RGB image, and the type ID and area of the target object are returned. Then, in the second stage, the detected target object is located. The region adopts point-to-feature voting method to obtain the 3D pose of the target object. This method in this chapter combines the advantages of the two methods. The instance segmentation network based on deep learning can quickly filter out the data in complex real scenes, so that the search space is reduced when feature matching, and the method of point-to-feature voting is retained. Recover the robustness of the target pose. Figure 2 shows the network flow diagram of the method proposed in this paper.
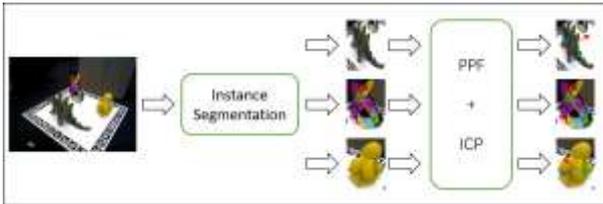

Figure. 2 Our algorithm process

2.2.1 Instance Segmentation Stage

Our data set comes from the HomebrewedDB data set of the BOP challenge competition. This data set is mainly used for 6D pose estimation of 3D objects. This data set contains a total of 33 model objects and 13 real scenes composed of these 33 model objects. As shown in Figure 3, the target sizes in the scene pictures in this dataset are almost the same, so it is necessary to enhance the data of scene target object instance segmentation before training the Mask R-CNN network.

Figure. 3 HomebrewedDB dataset scene images

Since the HomebrewedDB dataset has the target 3D model point cloud data, it is possible to obtain pictures from different angles of the target through the 3D point cloud data of the target, and then randomly paste these pictures from different angles into the VOC2017 dataset according to different proportions. On the picture, as shown in Figure 4.

Figure. 4 Synthesize objects containing model objects on the VOC2017 dataset

2.2.2 Pose Estimation

After segmenting the scene point cloud by instance, we obtain the local point cloud data of the space where the target object is located, and then calculate the matching between the point pair features of this local point cloud and the global feature description of the target object. In order to ensure that the error in calculating the point pair feature descriptor is small, we use the method of consistent normal vectors to process the target object and local point cloud, as shown in Figure 5.
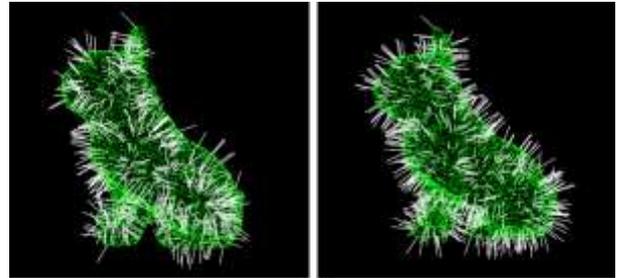

Figure. 5 Consistency processing of model point cloud normal vectors

3. EXPERIMENT

3.1 Experimental Environment

This experiment was implemented in the Pytharm development tool in the Windous 11 operating system, with Python version 3.8. The deep learning framework is Python, version 1.13.1, corresponding to CUDA version 11.6. The experimental GPU is NVIDIA GeForce RTX 3060, and point cloud data is processed using OpenCV library version 4.6 and PCL library version 1.12.1.

3.2 Instance Segmentation Results

After using the Mask R-CNN network to train the cat, dinosaur, and rabbit models in the HomebrewDB dataset, 340 scene images were tested in the HomebrewDB dataset. The recognition accuracy P and recall R of the three target objects were calculated, and the recognition effect was comprehensively evaluated using the harmonic mean F, as shown in Table 1.

Table 1. Mask R-CNN training results

Model	P	R	F
Cat	61.19%	72.21%	66.70%
Dinosaur	69.34%	78.52%	73.93%
abbit	67.14%	73.68%	70.41%
Total	65.89%	74.80%	70.34%

The target recognition results for some example images are shown in Figure 6, which intuitively shows that the target recognition position matches the actual position of the image, indicating that this training result can effectively recognize the position of the target.
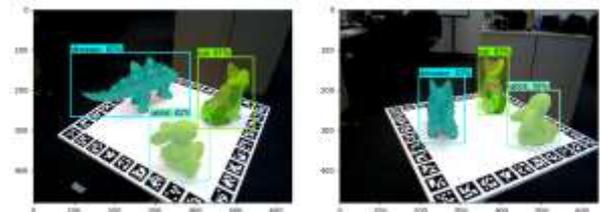

Figure.6 Target object recognition results

3.3 Pose Estimation Results

Compare the efficiency of target pose estimation of traditional point to feature and point to feature target pose estimation based on instance segmentation in the real environment. First, use the same Downsampling scale coefficient of 20% for model and scene point cloud to Downsampling them, respectively process the scenes in HomebrewdDB dataset in the form of traditional filtering and instance segmentation, and then input them into the improved point to feature pose estimation method in Chapter 3 Perform point to point feature calculation on three rabbit models, and verify their average error ADD after matching with the scene. If the average error is less than 0.05, it is determined that the matching is successful. Finally, the recognition rates and recognition time consumption of these three models in 340 real scene point clouds from the HomebrewedDB dataset were calculated, as shown in Table 2.

Table 2. Comparing the traditional PPF algorithm with our algorithm on three models

Model	Drost-PPF		Our	
	R	Time (s)	R	Time (s)
Cat	0.615	17.3	0.690	5.5
Dinosaur	0.560	34.5	0.727	7.7
Rabbit	0.673	13.2	0.672	4.1

At the same time, compared with other algorithms that also test the target model pose estimation on the HomebrewedDB dataset, these algorithms have different advantages in recognition accuracy and recognition efficiency. The algorithm in this paper has obvious advantages in target pose estimation efficiency and accuracy. The improvement, especially the improvement in efficiency is very obvious, as shown in Table 3.

Table 3. Comparing the traditional PPF algorithm with our algorithm on three models

Method	Data Type	R	Time (s)
Drost-PPF	D	0.615	20.7
CRT-6D	RGB	0.603	3.8
PointVoteNet2	RGB-D	0.556	21.9
Pix2Pose	RGB-D	0.711	10.5
Our	RGB-D	0.693	6.3

The following visually shows the 3D target recognition and pose estimation effect of the algorithm proposed in this paper. The red point cloud is the pose of the target in the scene, as shown in Figure 7.
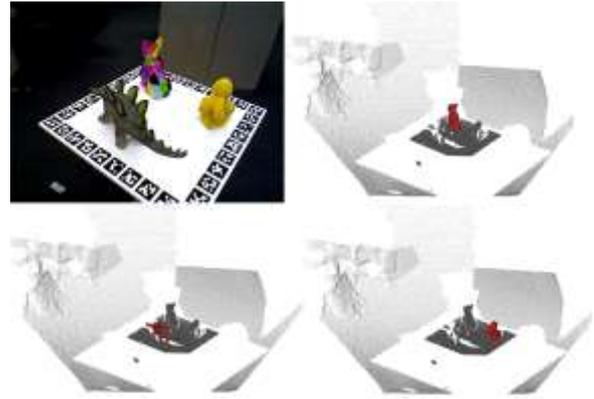

Figure. 7 Model pose estimation results

4. CONLUSION

This paper proposes a pose estimation algorithm based on Mask R-CNN network instance segmentation after processing the target object in the scene. Compared with the traditional pose estimation algorithm based on point pair features, this algorithm greatly reduces the input in real scenes. The number of scene point clouds; at the same time, when constructing the point pair feature between the target object and the scene point cloud, ensure that the normal vector direction of the point cloud is consistent. The method in this paper effectively solves the problem of low efficiency and accuracy of pose estimation due to the huge amount of scenic spot cloud data in the real scene.

5. REFERENCES

[1] Osada R, Funkhouser T, Chazelle B, et al. Shape distributions[J]. ACM Transactions on Graphics (TOG), 2002, 21(4): 807-832.

[2] Kazhdan M, Funkhouser T, Rusinkiewicz S. Rotation invariant spherical harmonic representation of 3 d shape descriptors[C]//Symposium on geometry processing. 2003, 6: 156-164.

[3] Wahl E, Hillenbrand U, Hirzinger G. Surflet-pair-relation histograms: a statistical 3D-shape representation for rapid classification[C]//Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. IEEE, 2003: 474-481.

[4] Frome A, Huber D, Kolluri R, et al. Recognizing objects in range data using regional point descriptors[C]//Computer Vision-ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part III 8. Springer Berlin Heidelberg, 2004: 224-237.

[5] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration[C]//2009 IEEE international conference on robotics and automation. IEEE, 2009: 3212-3217.

[6] Tombari F, Salti S, Di Stefano L. Unique signatures of histograms for local surface description[C]//Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part III 11. Springer Berlin Heidelberg, 2010: 356-369.

[7] Drost, B.; Ulrich M.; Navab N.; Ilic S. Model globally, match locally: Efficient and robust 3d object recognition. In Proceedings of the 2010 IEEE Conference on

Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 13-18 June 2010; pp. 998-1005.

[8] Birdal T.; Ilic S. Point pair features based object detection and pose estimation revisited. In Proceedings of the 2015 International Conference on 3D Vision (3DV), Lyon, France, 19-22 October 2015; pp. 527-535.

[9] Hinterstoisser, S.; Lepetit, V.; Rajkumar, N.; Konolige, K.; Going further with point pair features. In Proceedings of the European Confer ence on Computer Vision (ECCV 2016), Amsterdam, The Netherlands, 8-16 Octorber 2016; pp. 834-848.

[10] Wu, C.H.; Jiang, S.Y.; Song, K.T. CAD-based pose estimation for random bin-picking of multiple objects using a RGB-D camera. In Proceedings of the 2015 15th International Conference on Control, Automation and Systems (ICCAS), Busan, South Korea, 13-16 October 2015; pp. 1645-1649.

[11] Choi, C.; Taguchi, Y.; Tuzel, O.; Liu, M.Y. Voting-based pose estima tion for robotic assembly using a 3d sensor. In Proceedings of the 2012 IEEE International Conference on Robotics and Automation (ICRA), Saint Paul, MN, USA, 4-18 May 2012; pp. 1724-1731

[12] Castro P, Kim T K. CRT-6D: Fast 6D Object Pose Estimation with Cascaded Refinement Transformers[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2023: 5746-5755.

[13] Sundermeyer M, Hodan T, Labbe Y, et al. Bop challenge 2022 on detection, segmentation and pose estimation of specific rigid objects[J]. arXiv preprint arXiv:2302.13075, 2023.

[14] Park K, Patten T, Vincze M. Pix2pose: Pixel-wise coordinate regression of objects for 6d pose estimation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 7668-7677.

Augmented Reality Based on Discovery Learning: The Human Circulatory System to Improve Biological Literacy

Rogaya
Education Technology,
Postgraduate,
Universitas Negeri Medan,
Medan, West Sumatera,
Indonesia

R. Mursid
Education Technology,
Postgraduate, Lecture
Universitas Negeri Medan,
Medan, West Sumatera,
Indonesia

Baharuddin
Education Technology,
Postgraduate, Lecture,
Universitas Negeri Medan,
Medan, West Sumatera,
Indonesia

Abstract: To produce Augmented reality learning media products based on Discovery learning on circulatory system material, assessing the feasibility and effectiveness of Discovery learning based Augmented reality learning media products on circulatory system material in the learning process of class XI students at SMA Plus Sedayu Nusantara. This research is a Research and Development (R & D) development research with the ADDIE development model. Subjects included 2 material experts, 2 media experts, 3 students in the individual test, 9 students in the small group test, and 30 students in the limited field test. The results showed that: Augmented Reality learning media based on Discovery Learning on the circulatory system is very suitable to be used as a learning resource in learning material for the Circulatory System Class XI SMA; Discovery Learning-based Augmented Reality learning media on the circulatory system is very effective in use, testing the effectiveness of Discovery Learning-based Augmented Reality learning media products on the circulatory system fulfills the criteria of being effectively used which is shown from the results of product effectiveness testing using the t-test on different independent test tests. The average result of the assessment in the experimental class was that the average score of the learning outcomes test was higher than the average result of the control class students who did not use learning media.

1. INTRODUCTION

IPA is knowledge related to how to find out about nature systematically so IPA is not only mastering a collection of knowledge in the form of facts, concepts, or principles but also a process of discovery because IPA is a collection of systematic theories. , its application is generally limited to natural phenomena, born and developed through scientific methods such as observation and experimentation, and requires a scientific attitude such as curiosity, openness, honesty, and so on. According to Susanto [1] argues that science or science is a human effort to understand the universe through observations that are right on target, using procedures, and explaining by reasoning to get a conclusion. Science is part of the subjects that are developed based on the achievement of three aspects, namely knowledge, attitudes, and skills. Biology as part of natural science is a study related to how to find out about nature systematically so that biology is not only the mastery of a collection of knowledge in the form of facts, concepts, and principles but also a process of discovery [2].

Augmented reality is a technology that combines virtual or virtual world objects with the real world. This technology is generally developed on desktop PCs along with technological advances, many applications adopt this technology into a smartphone applications. The application of this technology can facilitate learning by displaying objects. 3D and animation that students are expected to better understand the material [3]. Advances in multimedia technology, especially learning multimedia, can help increase interest in learning because everything is completely digital, one of which is by utilizing mobile devices. The combination of the world of

technology and education can provide benefits to the learning process because it has great potential for teaching development, such as through Augmented reality technology. The use of Augmented reality technology in learning has advantages, one of which is that it can train the creativity and imagination of students [4].

From the results of Qumilaila's research [5] Augmented reality technology as a medium of learning has several advantages. AR allows digital content (audio, video, 2D and 3D objects) to appear to merge with the real world through a device. Augmented reality also enables content learning in three dimensions (3D), so that it can visualize things that are hard to see. In addition, AR's ability to present virtual objects to the real world in real time can activate students' sense of presence, immediacy, and immersion.

1.1 Learning Outcomes of Biology

Biology is one of the sciences that is part of the Natural Sciences (IPA) group. IPA is a science that studies natural phenomena, both material and abstract. The results of this natural science can come from experiments, direct observations, or with other scientific methods [6]. Whereas biology material that is integrated with science, many of which are not supported by a complete laboratory, cannot be practiced directly and media is needed to visualize it. Such material is for example the circulatory system, digestive system, heart work, excretory system, cell structure in both plants and animals, the solar system, the movement of animals on the seabed, and so on.

According to Qumilaila [7], biological material can be seen as something simple, but it can also be seen as something complicated and complex. Based on the results of Haka's research and discussion [8], interactive multimedia integrated with scientific values are very suitable for use as a biology learning medium.

Many materials in biology subjects cannot be visualized directly, so they need certain media. Visualization through learning media is one way that can be done to make it look real. Biology materials will be more interesting and easy to learn by packaging them in a digital format. For example, material about cells, and cell images presented in the form of ordinary/two-dimensional (2D) images will certainly be different from cells presented in the form of three-dimensional (3D) animation.

Jayawardana [9] said that teachers need to design classroom learning activities that combine the use of various kinds of technology. With a large number of digital tools and various kinds of information in them, education is needed to guide and foster all of these digital learning processes (e-learning). Digital The teaching and learning process is often faced with abstract material beyond the daily experiences of students so that the material becomes difficult for teachers to teach and difficult for students to understand.

1.2 The Nature of Augmented Reality
Arnaldi [10], Augmented Related is a technology that has sparked many fantastic and interesting things but delivered little. Many have made a virtual reality into the real world and created sensational videos that make the audience believe that what is experienced feels real, then visualize the sofa that will be purchased next as if it is presented in the living room, even

as if it is presenting someone who has died. In short, real life would be exactly like a television series.

Augmented Reality technology is a technology that can combine real and virtual situations displayed in real-time [11]. Augmented reality is the result of technology that can combine the virtual world and the real world that can be used by a teacher because he can project something abstract so that it can be interactive [12]. This technology can combine the real world and the virtual world that raises or projects in all directions according to what we want. The application of this technology aims to enable students to conduct experiments thereby increasing learning motivation. Its use can communicate objects in the virtual world to be used as a positive display so that it looks quite good, it will feel like you are in a much more concrete experience because of the limitations of the five human senses which cannot see directly abstract natural phenomena.

According to Qumilaila [13], Augmented Reality technology as a medium of learning has several advantages. AR allows digital content (audio, video, 2D and 3D objects) to appear to merge with the real world through a device. Augmented Reality also enables content learning in three dimensions (3D) so that it can visualize things that are hard to see. In addition, the ability of Augmented Reality to present virtual objects to the real world in real time can activate a sense of presence, closeness, and immersion in students.

Augmented reality, is a technology that can incorporate a 3D object into a real environment using webcam media. This technology has now been widely developed in developed countries for various purposes including educational purposes [14].
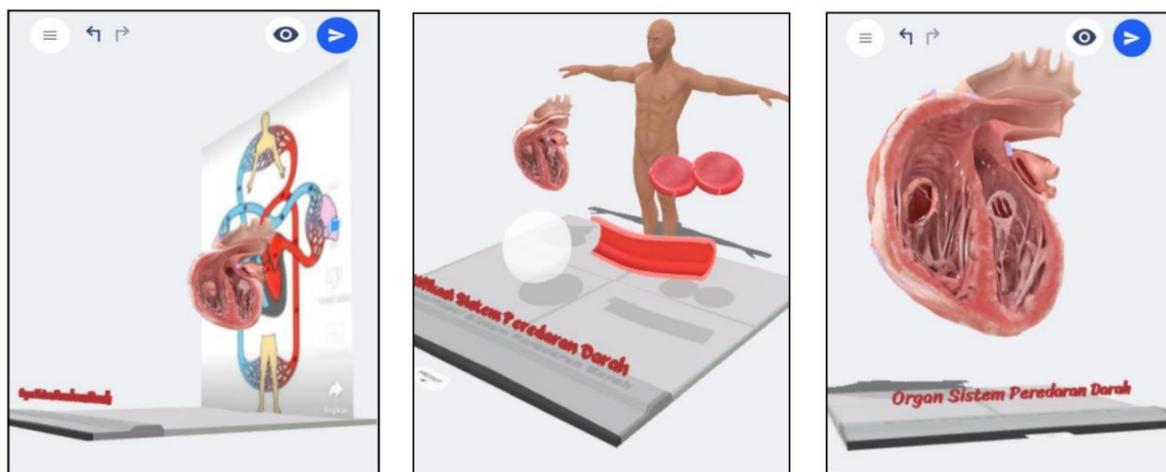

Figure 1. Augmented Reality organs of the human circulatory system

1.3 Discovery Learning Models
Discovery learning is learning based on discovery (inquiry-based), constructivism, and the theory of how to learn. According to the view of constructivism, learning is an active process of students constructing meaning, discourse, dialogue, and physical experience in which the assimilation process occurs and connects the experience or information that has been learned [15]. The discovery learning model is also known as an inquiry approach starting from a belief in the

development of students independently. This model requires active participation in scientific investigations.

The discovery learning model is a learning theory that is defined as a learning process that occurs when students are not presented with lessons in their final form but are expected to organize themselves. In Bruner's opinion [16], "Discovery learning can be defined as the learning that takes place when the student is not presented with the subject matter in the final form, but Raher is required to organize it himself", where

Bruno's basic idea is the opinion from Piaget which states that children must play an active role in learning in the classroom.

According to Khasinah [17], the Discovery learning method is an active and direct learning style developed by Jerome Bruner in the 1960s. Bruner emphasized that learning must be done while doing or learning by doing. With this method, students actively participate and not just passively receive knowledge. Discovery learning denotes a general instructional approach that represents the development of constructivist learning for school-based learning environments. The procedures for applying the Discovery learning model, namely: (1) Stimulation; (2) Problem Statements; (3) Data Collections; (4) Data Processing; (5) verification; and (6) Generalizations.

Table 1. Discovery Learning model steps in Biology learning

No	Syntax	Activities on Circulatory System Material and Digestive System
1	stimulation Providing stimulation	a) Given a problem that has no solution, for example viewing and observing 3D images of diseases related to the circulatory system motivates them to investigate and solve the problem. b) Students see and observe 3D images, namely pictures of food and digestive organs c) The teacher facilitates them by giving questions, and directions for reading books or texts.
2	Problem Statements Problem Identification	a) Students are given the opportunity to identify as many problems as possible related to the circulatory and digestive systems and their mechanisms, then formulated in the form of hypotheses or temporary answers to the problems set.
3	Data Collections Data Collection	a) Students explore to collect relevant data or information by reading, literature, and observing objects, such as the structure of the heart organ, blood vessels, and small and large blood circulation mechanisms through 3D Augmented Reality objects b) Students try to answer questions or prove the truth of the hypothesis. c) Students seek information about food substances and their functions, digestive organs, and their working mechanisms through 3D Augmented Reality objects.
4	Data Processing Data Processing	a) Students process the data or information they obtained in the previous stage and then analyzed, and interpreted it. All information, both from readings, interviews, and observations, is processed, classified, and tabulated by discussing with the group to get solutions to problems, for example formulating the mechanism of the circulatory and related systems and the digestive system of food and compiling a healthy menu.
5	Verification	a) Students carry out careful verification to test the hypothesis set by alternative findings, linked to the results of data processing. b) Think critically to solve problems by discussing ways to look for ways to prevent disease in the human circulatory system and digestive system and answer the questions.
6	Generalizations Draw conclusions	a) Draw conclusions that can be used as general principles and apply to all the same incidents or problems, taking into account the verification results. b) Presenting the results of group discussions around the circulatory organs, large and small circulatory systems, and efforts to prevent circulatory system diseases. c) Presenting the results of group discussions regarding the digestive organs, food substances, their functions, and healthy menus.

Hosnan [18] and Haerullah and Hasan [19], stated the advantages of the discovery learning model, namely as follows: (1) Helping students to improve and enhance cognitive skills and processes; (2) The knowledge obtained through this model is very personal and powerful because it strengthens understanding, memory, and transfer; (3) Can improve students' ability to solve problems; (4) Helping students strengthen their self-concept, because they gain trust in working with others; (5) Encouraging student involvement; (6) Encourage students to think intuitively and formulate their hypotheses; (7) Train students to learn independently; and (8) Students are active in teaching and learning activities because they think and use their abilities to find the final result.

Based on the problems, the research problems are formulated as follows: (1) How is the development of Augmented reality learning media products based on Discovery learning on the circulatory system material; (2) How is the feasibility of Augmented reality learning media products based on Discovery learning on circulatory system material; (3) How is the effectiveness of Discovery learning-based Augmented reality learning media products on the circulatory system material.

2. METHOD
This research is Research and Development (R & D) development research. Research methods are used to produce certain products, and test their effectiveness of these products. According to Sugiono [20], to be able to produce certain products, research that needs analysis is used, and to test the effectiveness of these products so that they can function widely in society. The model used in development research is the ADDIE development model, namely Analysis, Design, Development or Production, Implementation or Delivery, and Evaluations.

Figure 1. ADDIE Model Process Flow [21]

Research and development (research and development) is a series of processes or steps to develop a new media or improve existing media so that it can be accounted for [22].

The research location is Sedayu Nusantara Plus High School, which is located at Jl Marelan Pasar III Timur, Rengas Pulau, Medan City, where this school is a cadet school. When the research was conducted in odd semesters.

The research subjects were students of class XI SMA Plus Sedayu Nusantara where the study population consisted of all students of class XI SMA Plus Sedayu Nusantara which consisted of 7 classes, namely class XI Mia 1 to class XI Mia 7 class with a total of 205 students. Students as respondents consisted of two classes, namely the control class and the experimental class. The control class (XI Mia 1) had 30 respondents, while the experimental class (XI Mia 2) had 30 respondents. The object of this research is Augmented Reality learning media based on Discovery Learning on the human circulatory system.

Table 2. Expert Validation Questionnaire Assessment Qualification Criteria, and Student Response Instruments to Discovery Learning-based Augmented Reality Learning Media

percentage of Achievement Level	Eligibility	Description
81,26% ≤ X < 100%	Very good/Valid	No need for revision
62,6% ≤ X < 81,25%	Good/Valid	No Revision Required
43,76% ≤ X < 62,25%	Invalid	Revision
25% ≤ X < 39 %	very Invalid	Revision

Source : (Akbar, [23])

Based on the quantitative data from the results of the validator by material experts, media experts, and student response questionnaires, the next step is to analyze the data and calculate the percentage level of achievement based on the formula:

$$P = \frac{\sum x}{\sum x_i} \times 100\,\%$$

Keterangan:
x: The answer score from the validator
x_i: Score the highest answer
P: Presentation of eligibility level

The feasibility and effectiveness criteria achieved for use in media development are described in the following table.

Table 3. Media Eligibility Criteria

No	Score in Percentage (%)	Eligibility Category
1	80 ≤ P < 100	Very Eligible
2	60 ≤ P < 80	Eligible
3	40 ≤ P < 60	Adequate
4	21 ≤ P < 40	Inadequate
5	P < 21	Very Inadequate

The developed Augmented Reality learning media gets a positive response from students if the percentage obtained from the student response questionnaire reaches a score of ≥60%, then the learning media is categorized as feasible and effective.

Product Effectiveness Test Data Analysis Techniques
The effectiveness test aims to obtain information about whether or not the product development being tested is effective in the learning process.

Based on the formulation of the first problem, namely whether the Discovery Learning-based Augmented Reality Learning Media developed is feasible to use. Discovery Learning-based Augmented Reality Learning Media can be said to be feasible to use based on the results obtained from expert validation regarding suggestions and improvements related to the Discovery Learning-based Augmented Reality Learning Media that was developed. The next step is to carry out individual trials of 3 students, and small group tests of 9 students to find out the response to the Discovery Learning-based Augmented Reality Learning Media that was developed. Then a field trial was carried out with 26 students to find out the responses to the Augmented Reality Learning Media based on Discovery Learning that was made.

Based on the formulation of the next problem, namely whether the Discovery Learning-based Augmented Reality Learning Media developed is effective for improving Biology learning outcomes. Learning is said to be effective if there are significant differences in learning outcomes between classes that are given treatment and classes that are not given treatment. The hypothesis uses the mean difference test or t-test. The t-test is the average difference to find out whether there is a significant difference at the 0.05 significance level with Microsoft Excel 19.

The hypothesis formulated is:
Ho: μ1 = μ2 (there is no average difference between the treated and untreated classes).
Ha: μ1 ≠ μ2 (there is an average difference between the treated and untreated classes).
Decision-making Ho is accepted if the significance is greater than 0.05. The following is the calculation using the 2 difference test on the population average according to Sudjana (2009):

$$t = \frac{\bar{X}1 - \bar{X}2}{s\sqrt{\frac{1}{n1} - \frac{1}{n2}}}$$

Where:
$\bar{X}1$ = total average score of the experimental class sample.
$\bar{X}2$ = total average score of the control class sample.
s = standard deviation

3. RESULTS AND DISCUSSION
3.1 Results
The results of the assessment by media experts, material experts, individual trials, small group trials, and limited field trials for all aspects of the assessment are determined by the average score. The results of the assessment are then analyzed and determined whether or not it is appropriate to develop Discovery Learning-based Augmented Reality Learning Media on the material of the human circulatory system. The average percentage of the results of the assessment of media experts, material experts, individual trials, small group trials, and field trials is as follows:

Table 4. Average Percentage of Assessment Results on Discovery Learning-based Augmented Reality Learning Media on Human Circulatory System Material

No	Categorization	Percentage of average score %	Criteria
1.	Media Expert Validation	89,21	very feasible
2.	Material Expert Validation	95,83	very feasible
3.	Individual Trial	86,74	very feasible
4.	Small Group Trial	88,62	very feasible
5.	Field Test	88,03	very feasible
	Total Average	89,69	very feasible

Augmented Reality Learning Media based on Discovery Learning on the material of the human circulatory system from the validation of experts along with trials showed a percentage of 95.83% in media validation, 89.21% in material validation, 95.83% in individual trials, 86.74% small group trials, 88.03% in field trials. Overall, the average percentage is included in the "Very Eligible" category, which means that the use of Augmented Reality Learning Media based on Discovery Learning on the human circulatory system material meets the needs of students.

The data that has been obtained and processed, results that the highest score for the experimental class is 94 and the lowest score is 44. The mean value is 70.90, the median (Me) value is 72.22 and the mode value (Mo) is 83 and the standard deviation value is 14.56. For more details, see the following data.

Table 5. Learning Outcomes Using Discovery Learning-based Augmented Reality Learning Media on the material of the human circulatory system

Class	Interval Class	F. Absolut	F. Relatif %
1	44 – 52	5	16,67
2	53 – 61	6	20,00
3	62 – 70	2	6,67
4	71 – 79	6	20,00
5	80 – 88	8	26,67
6	89 - 97	3	10,00
	TOTAL	30	100

From the data obtained, is then calculated and the result is that the highest value in the control class is 78 and the lowest value is 17. The average value is 55.37, the median value (Me) is 55.56 and the mode value (Mo) is 56.61, 67, 72 and the standard deviation value is 6. For more details, see the following data:

Table 6. Student Learning Outcomes Using Augmented Reality learning media based on Discovery Learning

Class	Interval Class	F. Absolut	F. Relatif %
1	17-26	1	3,33
2	27-36	2	6,67
3	37-46	6	10,00
4	47-56	7	23,33
5	57-67	8	26,67
6	68–77	6	20,00
	TOTAL	30	100

The analysis requirements test performed is the normality and homogeneity tests. Testing was carried out using the Liliefors test. A summary of the normality of the two samples can be seen in Table 7 below:

Table 7. Summary of Data Normality Test with Liliefors

No.	Data	Class	L_{count}	L_{table}	Conclusion
1	Post-tes	Class XI student learning outcomes on the circulatory system using Augmented Reality learning media based on Discovery Learning	0,083	0,161	Normal
2		Class XI student learning outcomes on the circulatory system without using Augmented Reality learning media based on Discovery Learning	0,069	0,161	Normal

Thus the pre-test values for the experimental class and control class Lcount < Ltable, while the post-test values for the experimental class and control class Lcount <Ltable, it is synthesized that the two sample group data are normally distributed.

Homogeneity test analysis using the F test is to prove the largest variance and the smallest variance with the formula:

$$F = \frac{Varian\ terbesar}{Varian\ terkecil} = \frac{S_1^2}{S_2^2}$$

A summary of the homogeneity of the two samples is seen in Table 8 below:

Table 8 Summary of Data Homogeneity Test

No.	Data	Class	F_{count}	F_{table}	Conclusion
1	Post-test	Class XI student learning outcomes on the circulatory system using Augmented Reality learning media based on Discovery Learning	0.078	1.080	Homogen
2		Class XI student learning outcomes on the circulatory system using expository-based Augmented Reality learning media			

So it can be seen that Fcount <Ftable at a significant level of α = 5% states that the data of the two samples have a

homogeneous variance and it can be concluded that the research data meets the requirements for hypothesis testing.

Hypothesis testing uses the t-test with the formula, namely:

$$t = \frac{\bar{x}1 - \bar{x}2}{s\sqrt{\frac{1}{n1} + \frac{2}{n2}}}$$

The following is the formulation of this statistical hypothesis, namely:

Ho   : $\mu A1 \leq \mu A2$
Ha   : $\mu A1 > \mu A2$

Information:

$\mu A1$: average student learning outcomes taught using Augmented Reality learning media based on Discovery Learning

$\mu A2$: average student learning outcomes taught using expository-based Augmented Reality learning media

The t-test is used as a hypothesis-testing tool because the research data is normally distributed and homogeneous. The hypothesis in the research is:

Ho: Augmented Reality Learning Media based on Discovery Learning is not effective in improving Biology learning outcomes.

Ha: Augmented Reality Learning Media based on Discovery Learning effectively improves Biology learning outcome

From the results of testing the hypothesis obtained t count = 1,699 and t table 2,045 so that t count = 1,699 < t table 2,045, so that H0 is accepted, it can be concluded that there is a significant difference in the learning achievement of class XI students using Augmented Reality learning media based on Discovery Learning on the circulation system blood compared without using Augmented Reality learning media based on Discovery Learning.

3.2 Discussion

Augmented Reality learning media products based on Discovery Learning on the circulatory system were developed through the student trial stage and the validation stage of material experts and learning media experts and with product revisions that have been carried out based on the advice of experts, students, and teachers, Augmented learning media products are produced Reality-based Discovery Learning on the circulatory system that is feasible and effective for students and teachers to use in the learning process for high school class XI students. This is also in line with the results of Mauludin's research [24], conducting research on the effectiveness of Augmented Reality learning media products which shows that the Android version of the Augmented Reality application is considered attractive and acceptable to students and teachers as an alternative learning media for the human excretory system, as well as being effective. in assisting the learning process of biology in class.

Using this product will greatly assist teachers and students in understanding the material in the learning process. The advantages of this learning media product are (1) Displaying 3D object images on the organs displaying the structure of the human body, namely the circulatory system in the form of red blood cells, white blood cells, blood vessels, and especially the 3D view of the heart organ. and displays the circulatory mechanism of the circulatory system. (2) In addition to seeing 3D views and images as if they were real, it provides an interesting learning experience for students. And according to the results of Qumilaila's research [25] the opportunities for Augmented Reality to be accessed by the community including students are open. (3) The material for this learning media product also links the circulatory system material with the Discovery Learning learning method so that it stimulates students' analytical thinking skills to make their conclusions from the material that has been obtained. According to Khasinah [26] that the Discovery learning method is an active learning style, that learning must be done while doing or learning by doing. And in the research process, students tend to study actively and have group discussions after using Discovery Learning-based Augmented Reality learning media. (4). Learning is independent so learning tends to be Student-Centered Learning (focusing on student activity).

According to the research results of other researchers, Widyaningsih & Wulandari [27], Ainni & Budi [28], Lauryn [29] & Mauluddin [30], the use of Augmented reality as a learning medium can help students to understand concepts and theories, stimulate students to think conceptual and feel 3D, improve the image (representation and perception, create an interactive and attractive and more fun learning atmosphere). From the results of this study, it is evident that Discovery Learning-based Augmented Reality learning media products in the circulatory system is feasible and effective and are used as learning media that can help students understand concepts and theories, stimulate students to think conceptually and experience 3D learning to improve the picture in thinking and create an interactive and attractive learning atmosphere and fun.

Discovery Learning-based Augmented Reality learning media products in the circulatory system are products that are effectively used based on evidence of analysis of student learning outcomes in the control class and experimental class. It was found that the experimental class that used Discovery Learning-based Augmented Reality learning media on the circulatory system had higher learning outcomes than the control class that did not use Discovery Learning-based Augmented Reality learning media products on the circulatory system. In addition, effectiveness is also evidenced by an assessment of student responses showing a motivated attitude and providing a fun, active, and innovative learning experience. Data analysis was also carried out in testing the effectiveness of Augmented Reality learning media products based on Discovery Learning on the circulatory system fulfilling the criteria for effective use which is shown from the results of product effectiveness testing using the t-test on different independent test tests. The average value of the two groups of students, namely the control class and the experimental class, was then tested with the t-test [31]. The average result of the assessment in the experimental class was that the average score of the learning outcomes test was higher than the average result of the control class students who did not use learning media.

4. CONCLUSION

Augmented Reality learning media based on Discovery Learning on the circulatory system is very suitable to be used as a learning resource in learning material for the Circulatory System Class XI SMA. This feasibility has been assessed and validated by experts in their fields, namely material experts and learning media experts, and then with suggestions from experts, revisions to learning media products are carried out to produce better and ready-to-use products. In addition, the teacher's assessment of the biology study of the learning

media products developed was followed by product revisions according to suggestions. Based on the trials of class XI students at SMA Plus Sedayu Nusantara in individual trials, small group trials, and large group trials which were carried out, the results of this product were very feasible to use and developed to increase student learning outcomes, strengthen discussion learning (student-centered learning), to conclude (Discovery Learning), and improve students' thinking skills.

Augmented Reality learning media based on Discovery Learning on the circulatory system is very effective to use. This is evidenced by the assessment of student responses showing a motivated attitude and providing a fun, active, and innovative learning experience. In addition, in testing the effectiveness of Augmented Reality learning media products based on Discovery Learning in the circulatory system, it meets the criteria for effective use, which is shown from the results of product effectiveness testing using the t-test on different independent test tests. The average result of the assessment in the experimental class was that the average score of the learning outcomes test was higher than the average result of the control class students who did not use learning media.

5. REFERENCES

[1] Ahmad, Susanto. 2013. Teori Belajar dan Pembelajaran di Sekolah Dasar. Jakarta: Kencana Prenada Media Group.

[2] Abdjul, T., 2019. Buku Model Pembelajaran Releac. Cetakan Pertama. Gorontalo: Politeknik Gorontalo.

[3] Widyaningsih, & Wulandari, W. 2019. Edukasi Tata Surya Menggunakan Teknologi Augmented Reality. Jurnal Sains Komputer dan Teknologi Informasi. Palangkaraya: Prodi Teknik Informatika. Vol.2 (1), 29-39.

[4] Ainni & Budi, 2020 Ainni, L & Budi, A., (2020). Pembuatan Aplikasi Augmented Reality sebagai Media Pembelajaran Mengenai Tata Surya Berbasis Android untuk Sekolah Dasar. Jurnal Multimedia dan IT. Vol.4 (2).

[5] Qumilaila Q, Susanti B & Zulfiani Z. 2017. Pengembangan Augmented Reality Versi Android sebagai Media Pembelajaran Sistem Ekskresi Manusia. Jurnal Cakrawala Pedidikan. Vol.36 (1).

[6] Sugiarto, A. 2021. Pengembangan Media Pembelajaran IPA Tiga DImensi Pada Materi Sistem Peredaran Darah Menggunakan Augmented Reality Assemblr Edu di Kelas VIII Madrasah Tsanawiyah Negeri (MTsN) Batu. Tesis. Malang:Universitas Muhammadyah Malang.

[7] Qumilaila Q, Susanti B & Zulfiani Z. 2017. Pengembangan Augmented Reality Versi Android sebagai Media Pembelajaran Sistem Ekskresi Manusia. Jurnal Cakrawala Pedidikan. 36(1).

[8] Haka, N. B., Suryaasih, P. A., Anggoro, B. S., & Hamid, A. 2020. Pengembangan Multimedia Interaktif Terintegrasi Nilai Sains Sebagai Solusi Peningkatan Kemampuan Berpikir Kritis Kelas XI Mata Pembelajaran Biologi Di Tingkat SMA/MA. Quagga: Jurnal Pendidikan Dan Biologi, Vol. 13(1), 1. https://doi.org/10.25134/quagga.v13i1.3202

[9] Jayawardana, H. B. A. 2017. Paradigma Pembelajaran Biologi Di Era Digital. Jurnal Bioedukatika, Vol. 5(1), 12. https://doi.org/10.26555/bioedukatika.v5i1.5628

[10] Arnaldi, B.,Guitton, P., & Moreau, G. 2018. Virtual Reality and Augmented Reality. Great Britain and the United State: Briley.

[11] Irma, A., 2015. Pengembangan Media Pembelajaran Interakttif Berbasis Teknologi Augmented Reality pada Materi Ikatan Kimia. Skripsi tidak diterbitkan. Jakarta : Prodi Pendidikan Kimia, UIN Syarif Hidayatullah.

[12] Sugiarto, 2021 Sugiarto, A. (2021). Pengembangan Media Pembelajaran IPA Tiga DImensi Pada Materi Sistem Peredaran Darah Menggunakan Augmented Reality Assemblr Edu di Kelas VIII Madrasah Tsanawiyah Negeri (MTsN) Batu. Tesis. Malang:Universitas Muhammadyah Malang.

[13] Qumilaila Q, Susanti B & Zulfiani Z. 2017. Pengembangan Augmented Reality Versi Android sebagai Media Pembelajaran Sistem Ekskresi Manusia. Jurnal Cakrawala Pedidikan. Vol.36 (1).

[14] Nisa, S. K., Nurmiyati, N., & Rinanto, Y. 2019. Pengembangan Media Pembelajaran Laboratorium Virtual Berbasis Discovery Learning pada Materi Sistem Ekskresi untuk Kelas XI MIPA. BIO-PEDAGOGI, Vol. 8(2), 120. https://doi.org/10.20961/bio-pedagogi.v8i2.39434.

[15] Haerullah, A & Hasan, S., 2017. Model dan Pembelajaran Inovatif (Teori dan Aplikasi). Yogyakarta : Lintas Nalar.

[16] Muliana. 2015. Model-model Pembelajaran Inovatif. https://ainamulyana.blogspot.com/2015/09/model-model-pembelajaran-inovatif-dan.html.

[17] Khasinah, S. 2021. Discovery Learning: Definisi, Sintaksis, Keunggulan dan Kelemahan. Jurnal MUDARRISUNA,Vol.11 (3). http://dx.doi.org/10.22373/jm.v11i3.582

[18] Hosnan. 2014. Pendekatan Saintifik dan Kontekstual dalam Pembelajaran Abad 21. Bogor : Ghalia Indonesia.

[19] Haerullah, A & Hasan, S., 2017. Model dan Pembelajaran Inovatif (Teori dan Aplikasi). Yogyakarta : Lintas Nalar.

[20] Sugiyono. 2013. Metode Penelitian Kuantitatif Kualitatif dan R & D. Bandung : Penerbit ALfabeta.

[21] Rusmayana, 2021 Rusmayana, T. (2021) Model Pembelajaran ADDIE Integrasi Pedati. Bandung: Penerbit Widina.

[22] Sugiyono. 2013. Metode Penelitian Kuantitatif Kualitatif dan R & D. Bandung : Penerbit ALfabeta.

[23] Akbar, Sa'dun. 2015. Instrumen Perangkat Pembelajaran. Bandung: PT Remaja Rosda Karya

[24] Mauludin R, Sukamto A & Muhardi H,. 2017. Penerapan Augmented Reality sebagai Media Pembelajaran SIstem Pencernaan pada Manusia dalam Mata Pelajaran Biologi. Jurnal Edukasi dan Penelitian Informatika (JEPIN), Vol.3(2).

[25] Qumilaila Q, Susanti B & Zulfiani Z. 2017. Pengembangan Augmented Reality Versi Android sebagai Media Pembelajaran Sistem Ekskresi Manusia. Jurnal Cakrawala Pedidikan. Vol.36 (1).

[26] Khasinah, S. 2021. Discovery Learning: Definisi, Sintaksis, Keunggulan dan Kelemahan. Jurnal MUDARRISUNA,Vol.11 (3). http://dx.doi.org/10.22373/jm.v11i3.582

[27] Widyaningsih, & Wulandari, W, 2019. Edukasi Tata Surya Menggunakan Teknologi Augmented Reality. Jurnal Sains Komputer dan Teknologi Informasi. Palangkaraya: Prodi Teknik Informatika.Vol.2 (1). 29-39.

[28] Ainni, L & Budi, A., 2020. Pembuatan Aplikasi Augmented Reality sebagai Media Pembelajaran Mengenai Tata Surya Berbasis Android untuk Sekolah Dasar. Jurnal Multimedia dan IT. Vol 4 (2).

[29] Lauryn M, 2020. Aplikasi Pengenalan Hewan Bermetamorfosis dengan Menggunakan Augmented Reality Berbasis Androi. JIKA (Jurnal Informatika). Vol.4 (3).

[30] Mauludin R, Sukamto A & Muhardi H,. 2017. Penerapan Augmented Reality sebagai Media Pembelajaran SIstem Pencernaan pada Manusia dalam Mata Pelajaran Biologi. Jurnal Edukasi dan Penelitian Informatika (JEPIN), Vol.3 (2).

[31] Meiryani. 2021. Memahami Uji-t dalam Regresi Linear. https://accounting.binus.ac.id/2021/08/12/memahami-uji-t-dalam-regresi-linear/.

Facial Expression Recognition Based on Improved Densenet Network

Tao JunChen
Chengdu University of
Information Technology
Chengdu, China

Abstract: As a branch of image recognition, facial expression recognition helps to carry out medical, educational, security and other work more efficiently. This article combines deep learning knowledge and conducts research on expression recognition based on DenseNet121, a dense convolutional neural network that integrates attention mechanisms for multi-scale feature extraction. Firstly, in response to the insufficient ability of DenseNet121 to extract complex facial expression features, multi-scale feature extraction dense blocks were introduced to replace DenseBlocks used to extract features of different sizes; Secondly, using multi-scale feature extraction convolutional blocks to replace the large convolutional kernel at the head of DenseNet121 further enriches feature extraction; Finally, in order to extract more important features from the channel dimension, we consider combining ECA channel attention mechanism to help improve model performance. The experiment proves that the model proposed in this chapter has improved recognition accuracy by 2.034% and 3.031% compared to DenseNet121 on the FER2013 and CK+datasets, respectively. It also has certain advantages compared to other commonly used classification models.

Keywords: Facial expression recognition, Deep Learning, Convolutional Neural Network,DenseNet

1. INTRODUCTION

There are various ways for humans to convey emotions, such as language, voice, facial expressions, and body movements. Research has shown that among various emotional expression methods, facial expressions carry the most abundant emotional expression and can comprehensively reflect the emotions and psychological status of the speaker at that time. Facial expressions, as an important way of conveying information between people in our daily lives, also convey a rich and efficient amount of information. Therefore, studying facial expression recognition has significant social significance and practical value for our human emotional analysis and psychological assessment. The field of facial expression recognition, as an important component of image recognition technology, is showing an unprecedented development trend. Currently, large-scale software and hardware foundations have been fully popularized. With the support of hardware foundations, the application market of facial expression recognition is very broad, and the demand in various fields is also very large. This technology can be applied to multiple fields, such as medical treatment, education, driving safety, etc. Traditional facial expression recognition algorithms mainly use manually designed feature extraction algorithms, such as LBP, HOG, and other feature operators, which have average extraction results and low efficiency. After the emergence of deep learning technology, it can perform end-to-end feature extraction and classification work, opening up a new direction for the field of facial expression recognition. This article takes into account the shortcomings of traditional feature extraction algorithms, and uses the DenseNet convolutional neural network model in deep learning to conduct research. In order to enrich its feature level extraction capabilities, multi-scale feature extraction convolutional blocks, multi-scale feature extraction dense blocks, ECA channel attention mechanism are introduced, and an improved DenseNet model is constructed and obtained, The improved DenseNet model has better feature extraction capabilities in the field of facial expression recognition.

2. RELATED WORK

The research on facial expression recognition was first proposed by British biologist Darwin in the 19th century. Early facial expression recognition work belongs to the field of psychology. By studying the relationship between facial expressions and psychological states, it helps to discover the relationship between facial expressions and psychological states and explore some observation and treatment methods. In 1971, psychologist Ekman[1] and his research partner Friesen proposed a facial action encoding system by dividing facial muscle action units and analyzing the corresponding relationships between different expression categories and facial muscle units. The system marked the facial muscle action units and provided corresponding formulas for the expression types obtained from different combinations of markers, which had a certain impact on subsequent facial expression recognition work. Until 1978, Suwa et al. [2] were the first to use computer technology for facial expression recognition, with the aim of achieving efficient automatic recognition of facial expressions. In 1991, Mase et al. [3] proposed an optical flow method for facial expression recognition using optical flow as the expression feature of expressions. The development of computer technology has driven the maturity of facial expression recognition technology. Traditional methods commonly use handmade features, and Gabor filters [4] have shown certain advantages in feature extraction for facial expressions, improving the recognition rate of facial expressions, but the algorithm complexity is relatively high. Other traditional manual feature algorithms include

extracting local grayscale features using local binary patterns [5] to assist in recognition, and finally using SVM [6] for classification, LBP-TOP [7] on three orthogonal planes, etc. Luo et al. [8] proposed an improved principal component analysis facial expression recognition algorithm for static expression recognition. The grayscale features extracted from local binary patterns were used to assist in the global grayscale processing of facial expression recognition, and SVM was used for classification. Simulation experiments showed that this method can effectively classify different expressions.

In recent years, with the development of modern information technology and the continuous improvement of computer computing power, deep learning technology has rapidly developed and widely applied in the field of image processing. The focus of facial expression recognition has gradually shifted from traditional manual feature operators to generalization and strong classification capabilities in neural networks. By using computer vision algorithms and deep learning algorithms, visual observations similar to those of the "human eye" can be achieved in some scenes in daily life, Recognition functions, such as facial recognition for entering and exiting large venues, facial expression recognition for analyzing emotions, and scene recognition for the environment, can be applied. The rapid development of deep learning has led foreign scholars to use this technology to carry out facial expression recognition work on large datasets, and have made certain progress. For example, Lopes et al. [9] preprocessed facial expression data, constructed multi-layer convolutional neural networks for feature features, and inputted the extracted features into a classifier for facial expression classification. Experiments have shown that good recognition rates have been achieved on the CK+dataset. Later, the development direction of deep learning in classification tasks was mainly to improve recognition rate by deepening the depth and width of the network. A series of models began to emerge, such as AlexNet, VGGNet, ResNet, MobileNet, DenseNet, etc. The models had better performance in classification tasks.

3. PROPOSED METHOD

3.1 Multi scale feature extraction of dense blocks

In response to the lack of complex feature extraction ability in the DenseNet121 model, in order to make the features extracted by DenseNet121 richer and more significant, this paper proposes a multi-scale feature extraction dense block, which improves the structure of the DenseLayer single branch in the original DenseNet121 dense block Denseblock. As shown in Figure 1,
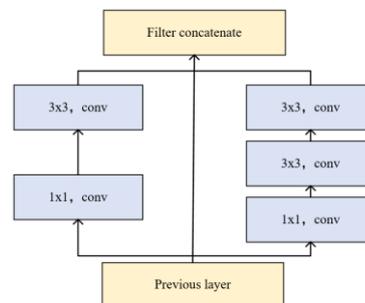

Figure. 1 Multi scale feature extraction of dense blocks

the number of convolutional kernels in the original Denselayer is reduced by half, At the same time, another convolution branch is developed in the feature extraction process. The new branch is composed of two convolutions of size 3 and one convolution of size 1. The purpose of this design is to provide

different scales of Receptive field for feature extraction, compared with the 3 provided by the original single trunk branch 3 × 3 Receptive field, multi-scale feature extraction dense block can provide 3 × 3 and 5 × 5 With two sizes of Receptive field, this structural design is convenient to better capture targets of different sizes and improve the effectiveness of feature extraction.

3.2 Multi-scale feature extraction convolutional blocks

The first layer of DenseNet121 is a convolutional layer with a size of 7 and a step size of 2. This large convolutional kernel is used to extract the boundary information of the image, which consumes a lot of computation. In order to further improve the feature expression ability, this paper proposes a multi-scale feature extraction convolution block to convolution with a size of 7. Compared to methods such as enhancing channels, the effect is better. The specific structure of the multi branch feature extraction convolution block is shown in Figure 2,The specific structure of the multi branch feature extraction convolution block first uses a convolution with size 3 for dimensionality reduction, followed by two branch structures for processing. The first branch uses a convolution with size 1 and a convolution with size 3 for feature extraction, and the second branch uses a maximum pooling with size 2. This strategy similar to combination pooling can effectively enrich the feature layer.
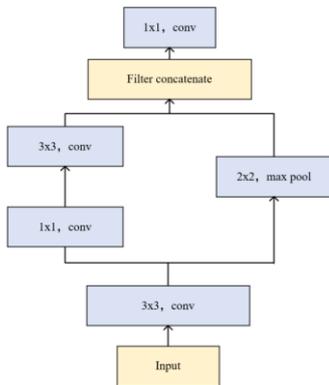

Figure. 2 Multi scale feature extraction of convolutional blocks

3.3 Lightweight Channel Attention Mechanism ECA-Net

In response to the problem of weak correlation between channels caused by the deepening of Dense121 layers, distinguishing different channels in the channel dimension will be helpful for the overall feature extraction work in facial expression recognition. Research has shown that the channel attention mechanism can help improve the feature extraction ability of convolutional neural networks. However, some current attention modules are designed more complex to achieve better performance indicators, making the overall model more complex. Here, an effective and lightweight channel attention module ECA-Net[10] is used, which has fewer overall parameters and excellent performance, and has better recognition performance gain for the overall model, ECA-Net is improved based on SENet. The core of ECA-Net is to use a 1D convolution and set an adaptive Kernel size to replace the

channel dimensionality reduction of the fully connected layer, as shown in Figure 3.
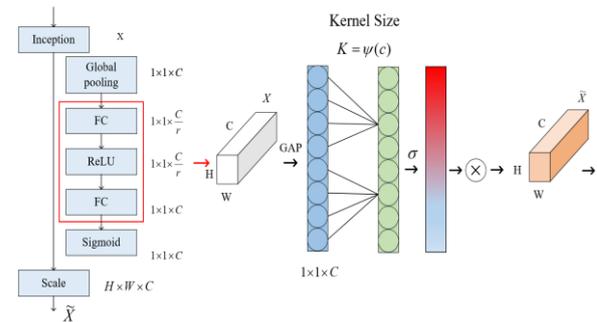

Figure. 3 Lightweight Channel Attention Mechanism ECA-Net

3.4 Improved DenseNet Network Architecture

On the basis of DenseNet121 model of dense convolutional neural network, in order to further improve its performance for facial expression recognition, the following work has been done: 1. A multi-scale feature extraction dense block is proposed to replace Denseblock in DenseNet121. The multi-scale feature extraction dense block has a multi branch feature extraction architecture, providing Receptive field of different sizes, It can extract richer and more refined features. 2. A multi-scale feature extraction convolutional block has been proposed to replace the top level convolution with a size of 7 in DenseNet121. Multiscale convolutional blocks can reduce computational consumption and better enhance feature expression capabilities, with better performance compared to methods such as channel enhancement. 3. On the basis of the improved model, the ECA channel attention mechanism was introduced, which increased the correlation between channels and improved the overall performance of the model in recognition work. Finally, the overall structure of a multi-scale dense convolutional neural network model integrating attention mechanism was proposed, as shown in Figure 4.
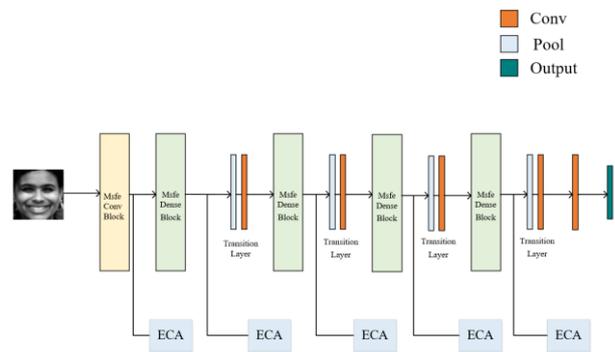

Figure. 4 Improved DenseNet Network Architecture

4. ANALYSIS OF EXPERIMENTAL RESULTS

4.1 Analysis Dataset

4.1.1 FER2013 Dataset

We proposed a facial expression recognition and classification technology based on convolutional neural network.ln order to

improve the robustness of fer task, we used multiple Data sets for training to improve the practicability of the network in expression recognition task. The first Data set is FER2013[11], a large facial image database randomlycollected by Google.After rejecting incorrectly marked frames and adjusting the clipping area, all images are registeredand adjusted to 48 pixels.FER2013 Dataset includes 35886 facial images, including 28709 training images, 3589verification images and 3589 test images. All images may correspond to one of the seven expression classification labels:0 anger, 1 Disgust, 2 fear;, 3 happy, 4 sad; 5 surprised; 6 Normal. The specific example is shown in Figure 5.

Figure. 5 FER2013 Dataset

4.1.2 CK+ Dataset

The second Data set is CK+ Data set[12].CK+ Data set recorded the facial movements of 210 adults with two hardware synchronized cameras in the laboratory environment. The participants are aged between 18 and 50 years old.CK+ Data set records 593 facial movement image data,which are based on the labels of the subjects' seven basic emotion categories: fear, happiness, disgust, sadness, anger, contempt and surprise.

4.2 Experimental pretreatment

The experiment compared the results on the FER2013 dataset and the CK+dataset. In terms of data preprocessing, considering the large number of image datasets, in order to accelerate the convergence speed of model training, improve training efficiency, and shorten training cycles, the processing of the dataset is divided into the following steps: image grayscale processing, image size normalization processing, facial expression localization and cropping processing, The processed dataset images are uniformly presented as 48 × The pre processed grayscale image of 48 is shown in Figure 6.

Figure. 6 Preprocessed images

4.3 Experimental configuration

The model proposed in this paper is based on pytorch deep learning architecture and runs on windows10 operating system. The GPU used is NVIDIA TiTan Xp. In the experiments of FER2013, the number of iterations is set to 250, the initial leaming rate is 0.01, and the learming rate is attenuated by 10 times every 20 iterations.In the CK + dataset experiment, the number of iterations is set to 60, the initial leaming rate is 0.01, and the learning rate is attenuated by 5 times every 5 iterations. Multiple datasets are set to 128 batches.

4.4 Experimental results and analysis

4.4.1 Experimental results on FER2013 Dataset

The confusion matrix obtained from the experiment of the improved model and DenseNet121 model on the Fer2013 dataset is shown in Figure 7 and Figure 8.
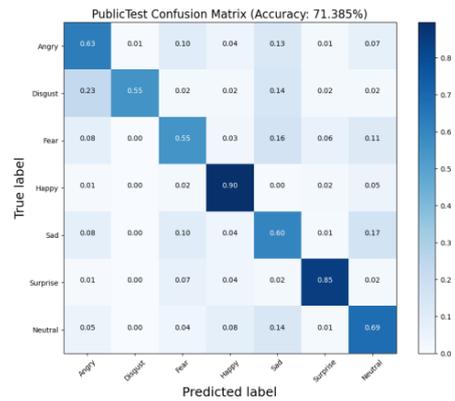

Figure. 7 The confusion matrix of this improved model on FER2013 dataset
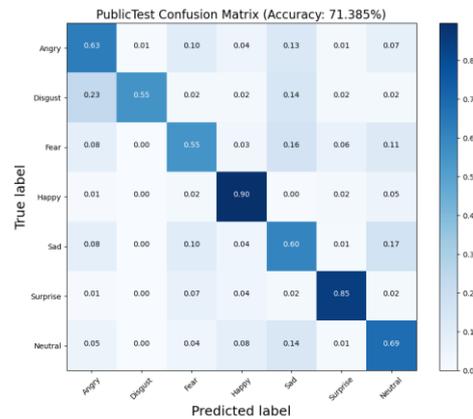

Figure. 8 The confusion matrix of DenseNet121 on FER2013 dataset

In order to further verify the progressiveness of the improved DenseNet model in the computer vision field of static expression recognition, we compared the accuracy with DenseNet121 and the public work of some common models of static expression recognition work scenes. The comparison results are shown in Table 1.

Table 1. The comparison results of the improved model and common network models on the FER2013 dataset in this article

Network mode	Recognition rate（%）
VGG19-BN	70.102
DenseNet121	71.385
ResNet32+CBAM	72.232
ResNet50+SVM	72.552
This paper	73.419

4.4.2 Experimental results on CK+ Dataset

The confusion matrix obtained from the experiment of the improved model and DenseNet121 model on the Fer2013 dataset is shown in Figure 7 and Figure 8.
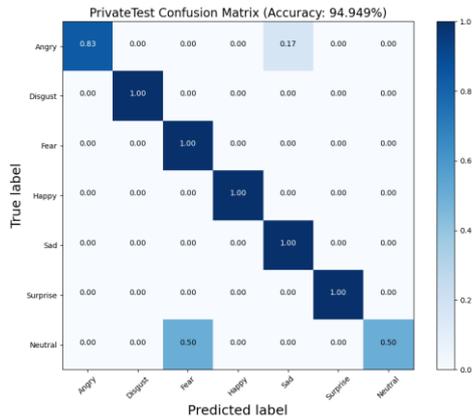
Figure. 8 The confusion matrix of this improved model on CK+ dataset
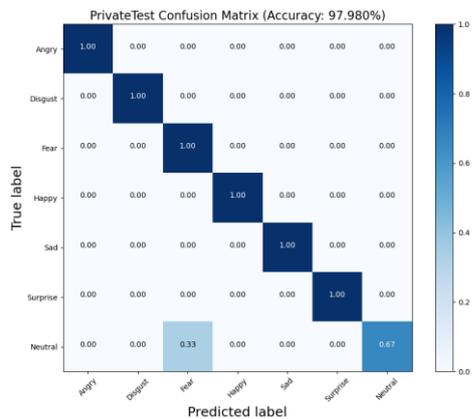

Figure. 9 The confusion matrix of this improved model on CK+ dataset

The comparison results between the improved model and common network models in this article on the CK+dataset are shown in Table 2.

Table 2. The comparison results of the improved model and common network models on the FER2013 dataset in this article

Network mode	Recognition rate（%）
VGG19-BN	93.816
DenseNet121	94.949
ResNet32+CBAM	94.512
ResNet50+SVM	95.183
This paper	97.980

4.4.3 Analysis of experimental results

By analyzing the experimental results of the improved DenseNet network and DenseNet121 on the FER2013 and CK+datasets, we found that due to the limitations of the network structure, DenseNet has average ability to extract complex facial expression features. After improvement, the DenseNet network improved its feature extraction ability, achieving the highest accuracy of 73.419% on the FER2013 dataset, and an improvement of 2.034% compared to DenseNet121; The accuracy rate of 97.980% was achieved on the CK+dataset, and 3.031% was increased compared with DenseNet121. Observing the confusion matrix of the improved

DenseNet network on the above dataset, it can be found that the model has a very high recognition rate for happy and surprised expressions, almost reaching 90% - 100%, but for some more complex expressions, such as fear and neutral expressions, the recognition rate is relatively general, On the FER2013 dataset, the recognition rate is only 55% -69%, indicating that the network has good recognition performance for facial expressions with obvious features and high discrimination. However, there is still room for improvement in recognition performance for facial expressions with insignificant category features, such as neutral facial expressions.

Analyzing the recognition performance of the improved DenseNet model and commonly used network models for facial expression recognition on the public datasets FER2013 and CK+, it was found that the model proposed in this chapter has a certain leading recognition rate compared to the improved VGG19-BN, ResNet32+CBAM, ResNet50+SVM models on the two facial expression datasets FER2013 and CK+, reflecting the advantages of the model proposed in this chapter in facial expression recognition tasks.

5. CONCLUSION

This article selects DenseNet121, a dense convolutional neural network, for expression recognition work. It is found that DenseNet121 has a relatively average ability to extract complex facial features. After analyzing the structural characteristics of DenseNet, we consider improving the model structure of DenseNet121. Firstly, we improve the Denseblock in DenseNet by replacing the original Denseblock with dense blocks extracted from multi-scale features, Provide multi-scale Receptive field to facilitate the extraction of features of different sizes; Using multi-scale feature extraction convolutional blocks instead of the top level convolution with size 7 in DenseNet121, and introducing the ECA-Net channel attention mechanism to enhance the model's feature extraction ability from the channel dimension, an improved DenseNet model was finally proposed. The effectiveness of the proposed network improvement on DenseNet121 was verified through experiments on FER2013 and CK+datasets, And progressiveness compared with other network models in facial expression recognition.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Ekman P, Friesen w V, Tomkins s S. Facial Affect Scoring Technique: A First Validity Study[U]. Semiot ica,1971,3(1)37-58.

[2] Suwa M, Sugie N, Fujimora K.A preliminary note on pattern recognition of human emotional expression[U]. 1978.

[3] Mase K.Recognition of facial expression from optical flow[J]. IEICE Transactions on Information and Systems,1991,74 (10): 3474-3483.

[4] Gabor，D. Theory of communication. Part l: The analysis of information[J]. Electrical Engineers Part llI Radio & Communication Engineering Journal of the Institution of,1 946,93(26): 429-441.

[5] Prabhakar S, Sharma J,Gupta S. Facial expression recognition in video using adaboost and SVM[U]. International Journal of Computer Applications,2014,104(2):1-4.

[6] Huang D, Shan C, Ardabilian M, et al. Local binary patterns and its application to facial image analysis: a survey. IEEE Transactions on Systems, Man, and Cybernetics. 2011,41(6);765~781.

[7] Shan C, Gong S, Mcowan P.W. Facial expression recognition based on Local Binary Patterns: A comprehensive study[J]. Image and Vision Computing, 2009,27(6):803-816.

[8] Sun W, Ruan Q .Two-Dimension PCA for Facial Expression Recognition[C]l/ International Conference on Signal Processing.IEEE, 2007.

[9] Lopes A T, De Aguiar E, De Souza A F, et al.Facial expression recognition with convolutional neural networks: coping with few data and the training sample order[J].Pattern Recognition,2017,61:610-628.

[10] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.

[11] Khaireddin, Y. , and Z. Chen . "Facial Emotion Recognition: State of the Art Performance on FER2013." (2021).

[12] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.