

Proposal for AI Video Interview Using Image Data Analysis

Jong-Youel Park †, Chang-Bae Ko ††

† Professor, Dept. of Smart IT, Baewha Women's University, Korea

†† Professor, Department of Business Administration, Kyungdong University, Korea
pjy@baewha.ac.kr, kcb2013@kduniv.ac.kr

Abstract

In this paper, the necessity of AI video interview arises when conducting an interview for acquisition of excellent talent in a non-face-to-face situation due to similar situations such as Covid-19. As a matter to be supplemented in general AI interviews, it is difficult to evaluate the reliability and qualitative factors. In addition, the AI interview is conducted not in a two-way Q&A, rather in a one-sided Q&A process.

This paper intends to fuse the advantages of existing AI interviews and video interviews. When conducting an interview using AI image analysis technology, it supplements subjective information that evaluates interview management and provides quantitative analysis data and HR expert data. In this paper, image-based multi-modal AI image analysis technology, bioanalysis-based HR analysis technology, and web RTC-based P2P image communication technology are applied. The goal of applying this technology is to propose a method in which biological analysis results (gaze, posture, voice, gesture, landmark) and HR information (opinions or features based on user propensity) can be processed on a single screen to select the right person for the hire.

Keywords: *AI Interview, video interview, P2P, HR Information, Image Data Analysis*

1. Introduction

At a time when AI and big data technologies, which are important factors in the era of the industry 4.0, are becoming important, we are aiming for a gradational daily recovery, but the recruitment market is rapidly shrinking due to the continuous spread of COVID-19. Along with this, non-face-to-face interview methods are being introduced in talent acquisition.

Automated systems using artificial intelligence (AI), a representative technology of the industry 4.0, and untact services, which had been gradually growing due to the increase in single-person households, are rapidly spreading around the world in the wake of the COVID-19. Blind recruitment is also spreading in order to apply artificial intelligence technology and maintain fairness in talent acquisition. Efforts are being made to solve barriers between countries and regional problems to recruit excellent talents. Major domestic companies are also actively considering non-face-to-face recruitment using artificial intelligence technology [1].

Intact is a compound word of interactive and untact while it refers to an online video interview in which an actual human resources manager conducts an interview using a remote video program and evaluates an interviewer. Major domestic companies are adopting online video interviews to ensure fairness and in-depth interviews with human resources managers, but they remain at the level of video conference, requiring

Manuscript Received: April. 25, 2022 / Revised: May. 1, 2022 / Accepted: May. 3, 2022

Corresponding Author: kcb2013@kduniv.ac.kr(Chang-Bae Ko)

Tel: Fax: +82-31-869-9909

Professor, Department of Business Administration, Kyungdong University, Korea

technical tools to detailedly identify interviewers.

In order to develop AI image analysis technology and recruit excellent talents, the private sector-led human resource development program in the new technology field is operated through education and training, corporate internships, and mentoring [2]. In order to hire the right talents for the company, non-face-to-face interview services applying AI technology are applied to overcome the national-to-region boundaries.

In this study, AI video interview using image data analysis is proposed as the final form in which the results of biometric analysis (gaze, posture, voice, gesture, and landmark) plus Human Resources (HR) information is displayed on one screen. Video-based multimodal AI image analysis, biometric HR analysis, and Web Real-Time Communication (WebRTC) based Peer-to-Peer (P2P) video communication are proposed as technical elements while a method of receiving video data, analyzing the interview information by AI, and providing it to the personnel manager when job applicants and HR personnel conduct AI video interviews.

2. Related Studies

2.1 WebRTC Technology and Image Analysis Technology

WebRTC is an API designed to allow web browsers to communicate with each other without the help of plug-in. Javascript API enables real-time communication without plug-ins. Using this technology, heterogeneous terminals may be linked to a video conference solution [3].

If there is data of various images, it is possible to continuously learn and analyze cloud-based intelligent images. Using a cloud-based intelligent image security platform, it is possible to provide intelligent image analysis services for people by utilizing image data [4][5]. Using the WebRTC protocol, plug-in-free live streaming is possible[10].

2.2 Trends of AI Interview Technology

AI is expected to replace HR jobs within the next 10 years, and the job that will be most affected by AI among HR jobs is expected to be in the field of the recruitment process. The reason for using AI interviews has the advantage of reducing human prejudice and minimizing potential discrimination such as language and appearance. In addition, fair recruitment is possible through job analysis when hiring [6].

It was found that AI interviews have a greater effect than the perception of potential convenience and fairness. The acceptance of a new technology requires verification of the usefulness and ease that it provides [7].

The introduction of artificial intelligence is also increasing in public services. And it is predicted that computers will supplement or replace service providers' labor, while that artificial intelligence robots will automatically replace service provision to perform non-face-to-face tasks [8][9].

3. AI Video Interview Configuration

3.1 Image Analysis Processing and HR Analysis Configuration by Job

Using WebRTC-based P2P communication technology, AI video interviewees and interviewers can safely share voice, video, and screen content. Template development for HR analysis proceeds with conversation classification, data, and program development. In the HR conversation classification, it is classified into 14 job groups and 34 jobs based on the National Competency Standards (NCS).

Figure 1 is a WebRTC screen sharing structure diagram. The Extension APP delivers the screen stream ID to peer1 and peer2 and allows peer1 and peer2 to check signaling on the signaling server and share media if it

matches. The three important APIs used in getUserAPI are MediaStream API, RTCMediaStream API, and RTCDataChannel API. The MediaStream API accesses data streams such as cameras and microphones. The RTCPeerConnection API connects encryption and rental width management, audio or video. The RTCDataChannel API provides general data P2P communication.

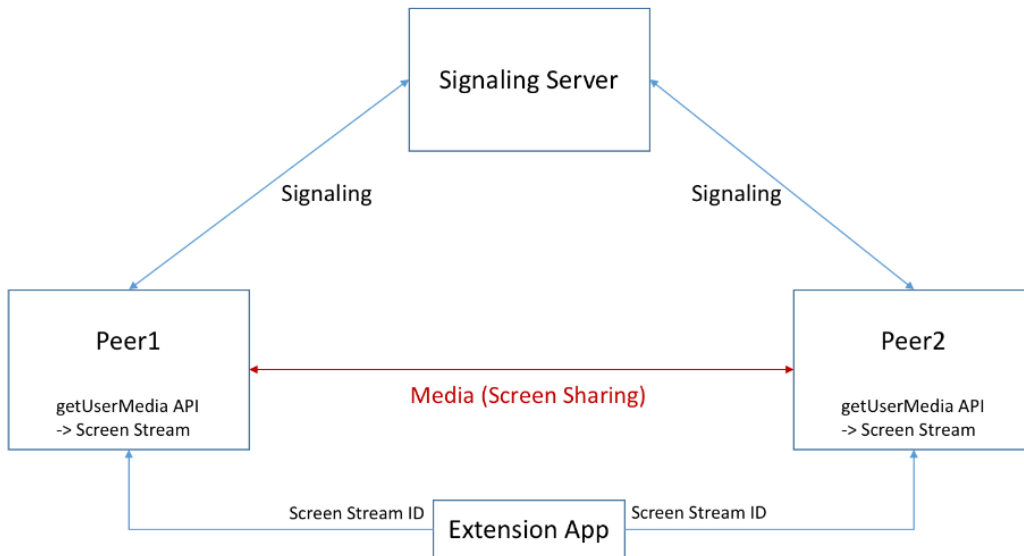


Figure 1. WebRTC screen sharing flow

Figure 2 is a diagram of keyword derivation for job competency. It consists of at least three trees, and extracts the topic and key words of the answer to the Q1. The Q2 derives keywords related to A1, selects keywords related to the Q2's questions related to the A1, and processes topics and questions based on relevance. The Q1 asks questions to understand job competency, while the Q2 is configured to be a related question to the A1 for Q1. In program development, AI learning type sample data is produced and data preprocessing is performed.

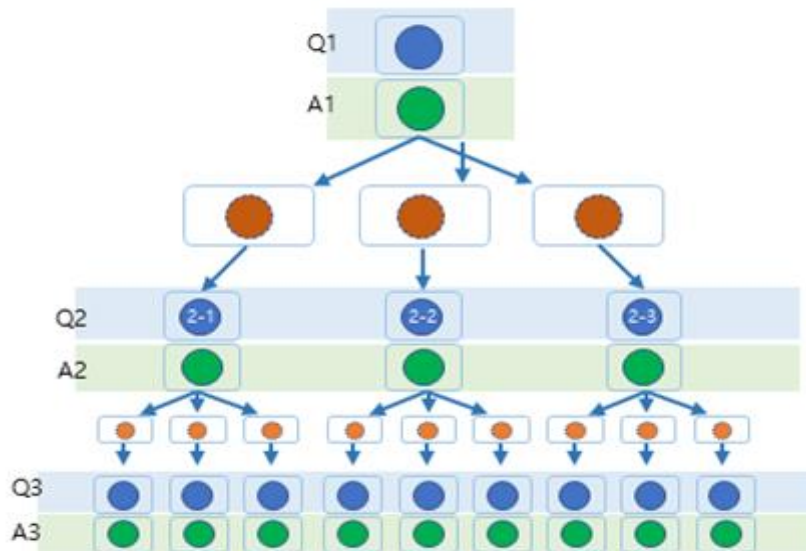


Figure 2. Structure diagram of keyword derivation for job competency

3.2 Configuring AI Conversational Interview Processing

The processing of the interactive system, which is the center of AI conversational interview processing, proceeds with extraction of image voice, and composition of three types of speech quality while model learning is carried out based on speech-composition data. The formula for scoring and processing the model learning, which is an important factor in the AI interactive interview, derives the most appropriate response using the softmax probability distribution, and scores the response for the part that fits the job competency.

Figure 3 asks questions to the interviewer and constitutes three speech information (entity extraction, utterance embedding, bag of words). Data learning is conducted by applying this information to the LSTM encoder and deep learning recurrent neural network algorithm in the conversation system. Data clustering for Q&A is performed, while the softmax probability distribution is used to obtain the most appropriate response. For relevant questions about the Q2, AI expected questions are extracted with similar parts of keywords and topics. When the answers are completed, the data is clustered, while the response to the job competency is scored to derive the output result as the system result value.

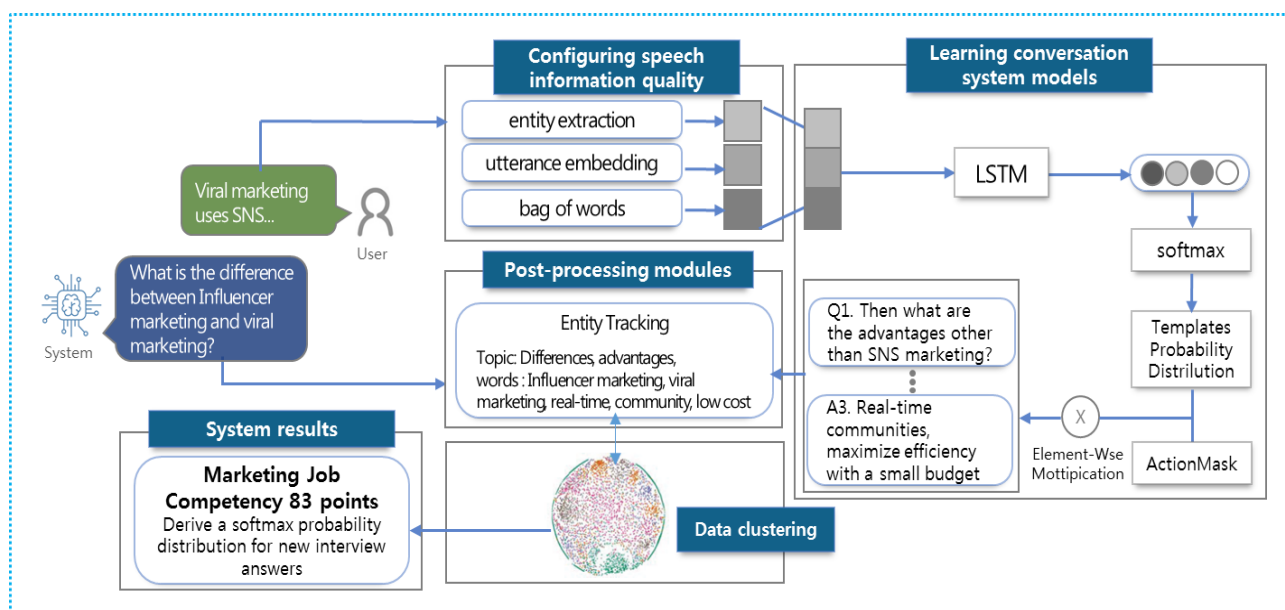


Figure 3. AI conversational interview processing structure diagram

4. AI Video Interview Implementation Plan

Interview customized analysis AI technology extracts multimodal biometric information such as facial expressions, voices, and words from interview video data of job applicants, and derives basic information that can analyze job group suitability and keyword propensity as well as output of objective biometric information data.

Figure 4 checks gaze processing, upper body movement, and gestures in image analysis. Facial expression analysis is processed by applying facial expression analysis techniques and gender and age information analysis techniques of six types (absence of expression, joy, sadness, anger, surprise, and fear) of Koreans and foreigners who have completed data learning. AI real-time image analysis is implemented by emotion analysis and landmark detection through face recognition. It is processed using 6 types of emotion analysis, 28,719 data learning, eye tracking, gender, and age recognition.

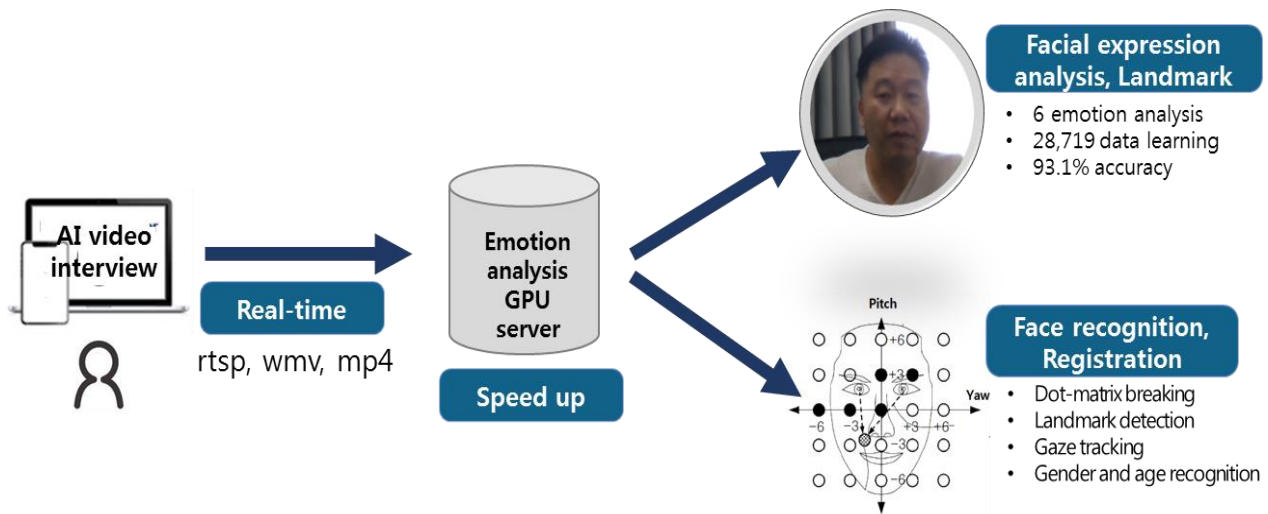


Figure 4. AI real-time image analysis

Figure 5 checks the voice tone, volume, and speech speed in voice analysis. It uses voice recognition technology that performs DNN-based feature enhancement, and derives the speaker's emotional state through nine emotion map algorithms based on the extracted bio-signal information. The algorithm for constructing the emotion map is processed with a weight ratio. Emotion maps are processed according to pitch frequency and volume decibels. The nine emotion maps consist of tiredness, joy, excitement, relaxation, comfortable, tension, lethargy, stability, and anxiety.

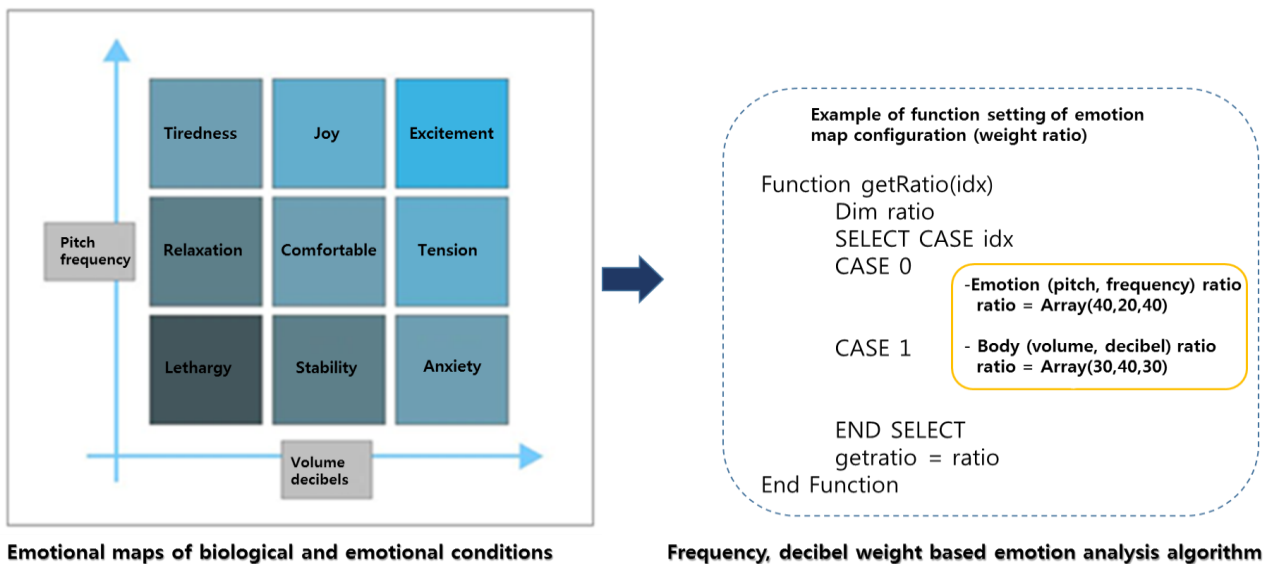


Figure 5. Implementation method of AI speech analysis algorithm

Figure 6 derives HR information of the interviewee based on video information such as gaze processing, movement, and gesture extracted through multimodal AI biometric analysis and voice information such as voice tone, voice volume, and speech speed. The algorithm for constructing the emotion map is processed with a weight ratio. Based on the interviewee's propensity information, the job group suitability based on the job description table is supposed to be matched.

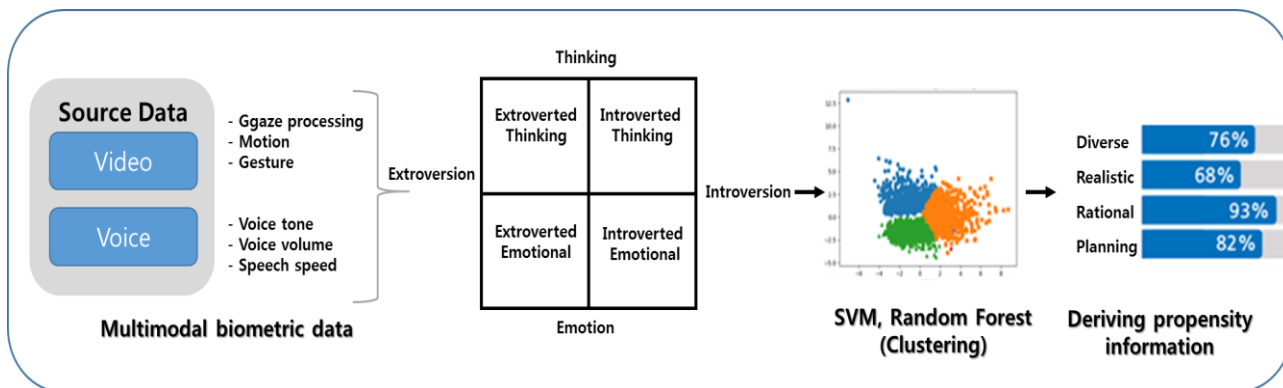


Figure 6. Implementation plan for job group suitability based on propensity information

Figure 7 shows the sample program for the AI video interview proposed in this study. The AI video interview service was implemented by analyzing the existing video interview service and AI interview service. The top screen is for the interviewer whereas the bottom one is for the interviewee. Up to five interviewees are allowed to be examined, while only one interviewer is available. On the interviewer screen at the bottom, frequently used words are analyzed in real time, and the answers to the questions are analyzed and AI-recommended questions are placed on the left side of the screen to help the interviewer proceed with the interview.

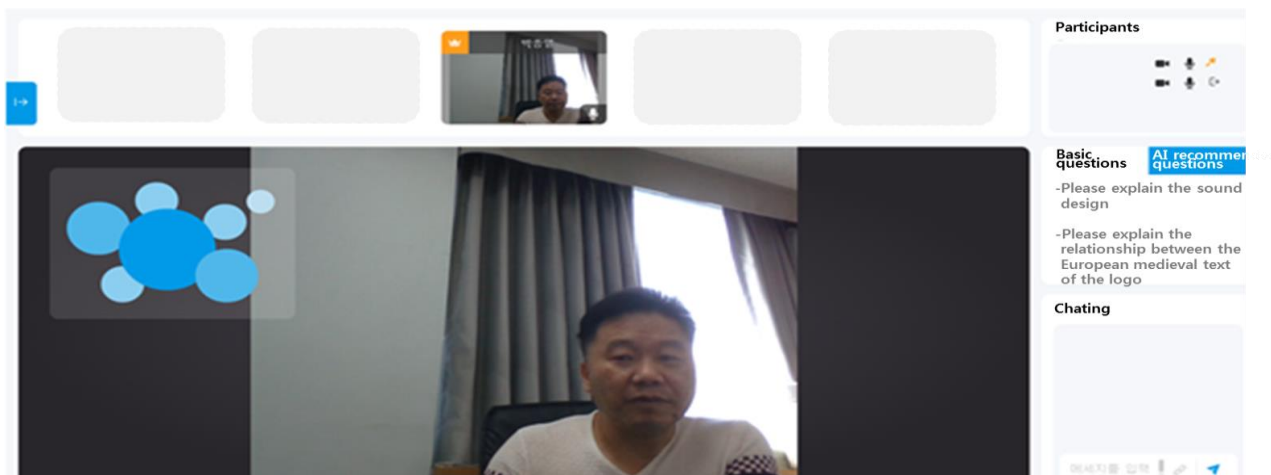


Figure 7. AI video interview implementation

5. Conclusion

It is a service that analyzes video/voice data transmitted through sensors such as microphones and cameras while answering questions presented by AI in an online environment, providing fair interview opportunities and quantitative AI analysis data, whereas it is used as means of fair assistance evaluation to efficiently recruit talents. AI video interview using video data analysis presented in this study is a technology that receives the applicant's video data and provides information about the interview to the HR manager when the job applicant and the HR manager conduct face-to-face interviews through this service. Biometric analysis results plus HR information according to biological analysis is configured on a single screen. Applying the results of this study, biological signals composed of attitudes, gazes, postures, voices, gestures, and landmarks that are difficult to find in online video interviews can be analyzed to derive new results such as psychological status and apply objective results when conducting AI video interview. Through various analyzes in subjective evaluation and AI video interview, the interviewer of the company scores and provides the objective evaluation, thereby

deriving a result that can recruit excellent talents.

Acknowledgement

This research was supported by Kyungdong University Research Fund, 2021.

References

- [1] Jinhung Cho, Hwansoo Kang, Woochang Yoo and Kyutae Park, "Prototype Design and Development of Online Recruitment System Based on Social Media and Video Interview Analysis," *Journal of Digital Convergence*, Vol. 19, No. 3, pp. 203-209, Dec 2019.
DOI: <https://doi.org/10.14400/JDC.2021.19.3.203>
- [2] Sun-Jeong Jeong and Hwi-Eun Nam, "Status and Implications of Private-led Human Resources Development Program in New Technology Field," *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 23, No. 2, pp. 322-330, Feb 2022.
DOI: <https://doi.org/10.5762/KAIS.2022.23.2.322>
- [3] Sung Jin Baek, Rang Hyuck Lee and Chang Seok Yi, "Design and Development of A Systemic Structure to Ensure the Interoperability between the WebRTC-based Video Conferencing Systems and Heterogeneous Terminals," *KIISE Transactions on Computing Practices*, Vol. 23, No. 4, pp. 238-243, Apr 2017.
DOI: <https://doi.org/10.5626/KTCP.2017.23.4.238>
- [4] Yun-jung Choi, Yoojin Chung, Ho Young Yoon, Minjung Kim, Na Young Kim, Chen Lu, Juyeon Sin, Ju-hee Lee, Na-young Kim, Eun Yeo and Jewon Kang, "Analysis of the 19th Presidential TV Debate Using Deep Learning Based Video Processing Algorithms : Analysis of the frequency, facial expression and gaze," *Korean Society For Journalism And Communication Studies*, Vol. 23, No. 4, pp. 238-243, Oct 2017.
DOI: <https://doi.org/10.20879/kjics.2020.64.5.009>
- [5] Changki Jang and WookJoon Sung, "A Study on Policy Acceptance Intention to Use Artificial Intelligence-Based Public Services: Focusing on the Influence of Individual Perception & Digital Literacy Level," *Informatization Policy*, Vol. 29, No. 1, pp. 060-083, Oct 2022.
DOI: <https://doi.org/10.22693/NIAIP.2022.29.1.060>
- [6] Jungmin Kwon, Jieun Kwon and Youngsun Lee, "Possibility of using AI in the Admissions Officer System in Korea," *The Journal of Korea Elementary Education*, Vol. 31, Supplement, pp. 1-15, Jan 2021.
DOI: <http://doi.org/10.20972/kjee.31..202101.1>
- [7] Namin Shin and Sejin Chang, "A Study on High School Students' Perceptions of AI Interview for University Admission," *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 22, No. 7, pp.242~251, Jul 2021.
DOI: <https://doi.org/10.5762/KAIS.2021.22.7.242>
- [8] Namin Shin and Sejin Chang, "A Study on High School Students' Perceptions of AI Interview for University Admission," *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 22, No. 7, pp.242~251, Jul 2021.
DOI: <https://doi.org/10.5762/KAIS.2021.22.7.242>
- [9] Ki-chaе Min and Kyu-sung Lee, "A Study on Non-Facing Services of National Pension in the era of the 4th industrial revolution," *The Journal of the Convergence on Culture Technology*, Vol. 4, No. 3, pp.139-147, Aug 2018.
DOI: <http://dx.doi.org/10.17703/JCCT.2018.4.3.139>
- [10] Hansol Cho, Minsung Choi, Younwook Choi and Beongku An, "Live Streaming Service using Webrtc," *Proceedings of Symposium of the Korean Institute of communications and Information Sciences*, pp.475-476, 2019.