Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 8, July 2021: 140-161

Development of Genetic Algorithm and Diagnosis Techniques for Covid-19 Virus Detection

Khalil Ibrahim Mohammad Abuzanouneh[1], Khalil H. A. Al-Shqeerat[2]

Abstract

The rapid detection of the SARS-CoV-2 is a critical phase in preventing and controlling the covid-19 disease. This paper presents a novel approach for the diagnosis and analysis of SARS-CoV-2 gene samples at the early detection stage using a Multi-Objective Genetic Algorithm (MOGA). We apply MOGA and the diagnostic model to analyze and detect the genes of swabbing specimen nucleic acids, which are the most discriminative features of coronavirus. The Genetic Operators of MOGA are performed to search and diagnosis the suspected objects that have the same features of virus genes using the current population (samples swabs). In an early step, the amplification technique will be performed using the selective amplification of swabs to improve the diagnosing testing. In the next stage, MOGA will examine and explore the structure changes and new properties of SARS-CoV-2 using new fitness values to detect the nucleic acids that are performed by molecular diagnostic assays based on new chemical and physical characteristics of coved -19 swab. The proposed algorithm is an efficient technique to diagnose and explore gene mutations of the SARS-CoV-2 virus.

Keywords: *Genetic Algorithm, Diagnosis Techniques. Colorimetric Validation, Amplification Curve, Cycle Threshold.*

Introduction

The COVID-19 (SARS-CoV-2) epidemic was detected in Wuhan, China in December 2019, causing a pandemic, bringing illness, and death to our lives, it rapidly spread globally, caused by coronavirus 2. The COVID-19 disease has expanded to be a public health emergency of global concern. SARS-CoV-2, which is widely present in nature, has a positive sense and a

[1] Associate professor, Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia, ka.abuzanouneh@qu.edu.sa

[2]Associate professor, Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia, kh.alshqeerat@qu.edu.sa

single-stranded RNA virus. The human coronavirus has caused outbreaks of respiratory diseases in animals and humans (Zhu, Zhang, Wang, Li, Yang, Song, Zhao, Huang, Shi, Lu, Niu, Zhan, Ma, Wang, Xu, Wu, Gao, and Tan, 2020).

Globally, according to the World Health Organization (WHO) report on 14 April 2021, there have been 136,996,455 confirmed cases of the disease, including 2,951,865 deaths, and 733,287,398 vaccine doses have been administered. At present, the effective methods of disease control and prevention are quarantine management and prompt diagnosis. There is still an urgent need for a sensitive detection measure to identify infected patients with Covid-19, especially asymptomatic carriers. Rapid isolation of infected people during the period of infection is critical. Therefore, a large-scale diagnostic method is needed to detect the spread of the virus in the population very quickly, comprehensively, and accurately (Phua, Weng, Ling, Egi, Lim, Divatia, Shrestha, Arabi, Ng, Gomersall, Nishimura, Koh, and Du, 2020). In addition, the development and promotion of repeated testing of large groups of people are required as a long-term strategy to control the new outbreak until effective antibodies are produced and immunity against COVID-19 diseases is provided. The qualified test should be used to diagnose SARS-CoV-2 infection in patient samples to detect viral genomes and viral antigens (Mizumoto, Kagaya, Zarebski, and Chowell, 2020), according to "Novel Coronavirus Pneumonia Diagnosis and Treatment Plan (Provisional; 7th Edition Revisions). Testing and analysis processes for the detection of SARS-CoV-2 antigens depend on the specificity, sensitivity, and diagnostic speed of the antibodies, as well as specific oligonucleotides required for the detection of viral RNA (Laboratory testing for 2019 novel coronavirus (SARS-CoV-2) in suspected human cases, 2020).The SARS-CoV-2 diagnostic consists of three steps:

1. Oropharyngeal swab specimens or collecting nasopharyngeal.
2. Isolation of total RNA.
3. Specific detection of the viral genome by RT-qPCR.

RT-qPCR includes a reverse transcriptase (RT) step, which converts the viral RNA to DNA, using oligonucleotides specific and DNA polymerase chain reaction to make viral cDNA (qPCR). In this step, many pieces of the viral genome are amplified and detected using a specific oligonucleotide probe known as a fluorescent dye. The test procedure requires many steps to process the samples, and the diagnostic laboratory needs 3 to 24 hours for the detection process when the number of samples very huge. Increasing the number of RT-qPCR testing capabilities depends on SARS-CoV-2 RNA diagnostics (Laboratory biosafety

guidance related to the novel coronavirus (2019-nCoV), 2020). Therefore, new techniques and methods are required to improve diagnostics in the detection of viral genomes and RNA isolations.

Research tools for genetic algorithms (GAs) based on various techniques such as representation and fitness analysis, mutation, recombination, selection, and evolutionary computation are emerging as optimization methods.Image analysis based on image processing technology in the medical field is very important. This research diagnoses and describes the development of COVID-19 diseases at an early stage. SARS-CoV-2 species are easily detected by changing the color using microscopic images of the assay samples.Analyzing the results by changing colors through evolutionary computation in sample portions using multiple objectives genetic algorithm (MOGA) as a methodology and strategy of evolutionary genetic algorithm is very significant. MOGA is a type of machine intelligence to derive its behavior using amplification processes of SARS-CoV-2 in nature.A polymerase chain reaction (PCR) test is used to detect genetic material from a specific organism, such as some virus genes. Testing the sample detects the presence of a virus if infected at the time of the test.

Image Processing Techniques

Digital image processing (DIP) is a part of artificial intelligence that is used to analyze images created by computers. It can create visualization and interpretation processes in the medical field. DIP has new and powerful tools for analyzing, detecting, transmitting, storing, and displaying the results of medical images. Image processing techniques are not expensive and do not require complex testing and lab equipment. The results extraction of assay samples is a challenge to innovate and develop an integrated system from design to implementation. An accurate result is a measure of the successful test that is used in the analysis of the SARS-CoV-2 sample. The steps of COVID-19 analysis contain information extraction, discovery, diagnosis, monitoring, and evaluation of the result.

The cell morphology techniques of DIP can be used to detect and diagnose many medical diseases such as coronavirus and WBCs, for example, white blood cells play the main role in the diagnosis of different diseases, as information extraction is important for a hematologist's analysis. There are some complications in extracting color pixels from the test in samples of similar colors and sizes. Moreover, the color contrasts between cell boundaries based on the conditions of the color capture process and lighting.

This research uses a hybrid deep system that focuses on the colorimetric validation process for assay samples using MOGA and a diagnosis system that has been developed in the last few years. Use of MOGA with the colorimetric validation technique (CVT) using a set of constraints to extract the data required to classify and identify different types of samples. A deep analysis system using diagnostic technology is introduced for the entire sampling process based on amplification stages, temperature rate, and other factors related to swab samples to be reference information stored in the sample slide database system.The proposed system focuses on the analysis and diagnosis of covid-19 disease and its features in microscopic samples based on changes in gene geometry, texture, statistical analysis, and color contrast of genes. Therefore, the microarray technology can be used to make a robust genomic system that can use it to analyze the behavior of a huge number of genes simultaneously. The deep analysis of samples that were obtained from the microarray technology can be helped in the detection and diagnosis of most diseases.An automated testing system has been developed for analyzing and detecting positive and negative cases of assay samples accurately. The proposed algorithm performs a set of operations such as deep analyzing of valid and invalid samples, extracting accurate information of locating genes in the samples and getting effective diagnostic.

Genetic Algorithm

Charles Robert Darwin invented the genetic algorithm as the algorithm of natural selection by taking a set of numbers and generating huge solutions as outputs to select the best values. Genetic Algorithms were formed to represent processes in complex natural systems, which are important to evaluate and perform effective large-field research for a set of optimal solutions.

In the color-processing test of samples, MOGA can provide accurate information related to the analysis sample data compared to the traditional optimization systems. The fitness function of the genetic algorithm depends on the individuals of gene samples, moreover, MOGA has a set of operating methods such as reproduction, crossover, and mutation that can be used to generate new solutions and use them to discover the optimal solutions for the new mutation of samples that contain new chromosomes (Mahony, 2008). MOGA has a range of solutions to improve the diagnostic process of the test sample and select the best solution that satisfies multi-objective and conditions. The basic stages of the studied MOGA are illustrated as follows.

a) Initial Population

The first process begins with a set of individuals called a Population (set of test samples). Each individual (sample) is a solution to the problem that we want to solve. An individual (Sample) is characterized by a set of parameters (variables) known as genes such as ORF1ab, N, E gene of SARS-CoV2. A set of genes are joined into a string to form a Chromosome (one target or sample). In the proposed algorithm, an individual's gene set is represented using binary values (a sequence of 1s and 0s) to encode genes in the chromosome (ORF1ab, N, E gene).

b) Fitness Function

The fitness function determines how well an individual is fit and able to compare with other individuals. It evaluates the fitness value for each individual (the sample's genes), and the probability of selecting an individual for the RT-PCR test based on its fitness value (valid or invalid).

c) Selection Process

The selection process aims to check and select the fittest individuals (valid genes) and allow them to pass their genes to the next stage of the RT-PCR test. The individuals will be selected based on their fitness values. Individuals with high fitness (valid gene) have a good opportunity to be selected for the next process.

d) Chromosomal Crossover

The crossover process is the most significant stage in the MOGA algorithm. In the test sample and for each pair of parents (two samples) that are mated, a crossover pointed is created between two samples to get the highest values that satisfy diagnostic conditions based on MOGA and chosen genes of a valid sample.

e) Genetic Mutations of SARS-CoV2

In certain new offspring formed from SARS-CoV2, some of its genes can be mutated with a low change probability, meaning that some bits (gene values) in the bit string (sample) can be changed.

f) Termination of Genetic Algorithm

The execution of the algorithm is terminated if the solutions are convergent, satisfactory, and compatible with the characteristics of the SARS-CoV2 genes. In this case, the genetic algorithm performs a set of diagnostic and analysis operations for the genes of the sample.
Some comments are taken into consideration when dealing with a genetic algorithm:
- The population (number of samples) has a fixed size;
- New generations are formed and tested;
- Individuals with least fit are dying (invalid samples);
- Space is made for new offspring (new samples);
- The sequence of diagnostic stages is repeated to test new individuals at each new stage.

MOGA is an effective color-contrast detection method for test samples used to diagnose and analyze the quality of test results by contrasting natural colors at different scales of assay samples. MOGA is a technology for the systematic search for solutions using its genetic methods to improve and solve complex problems based on deep and rigorous analysis in the medical fields. It can be used in genetic color gradient analysis to accurately determine the features of test samples. The image color processing technology has some problems such as color interference, which is reflected in the real color, the color convergence is not suitable for the human system, and the color distribution in the examination samples is not suitable for the normal visual limits of human perception (Chinese Center for Disease Control and Prevention, 2020). It is not enough to apply one technique to be suitable for one type of RT-PCR test sample analysis. MOGA can extract the desired color values, choose the appropriate features for the diagnostic test results, and analyze using the heuristic process of the proposed algorithm to solve analysis problems effectively (Yu, Wu, Hao, Dong, Mao, Pelechano, Chen, and Yin. 2020).

RT-PCR Analysis

There are a set of factors that can be used in this research to analyze results of color pixels for the SARS-CoV-2 virus at a low level of copy amplification. Experimental testing of samples can be checked whether the required factors (time, isothermal amplification, freezing loop-mediated, temperature) have negative or positive results. This paper proposes a system that can be applied to diagnose and analyze Covid_19 disease based on MOGA technology. The purpose of this work is to accelerate SARS-CoV-2 detection, improve detection accuracy, and assist clinicians and healthcare alike in reducing time and effort in achieving output. An automated diagnostics system can be developed, using available rule-based tools to cover a

wider range related to the SARS-CoV-2 containing different types and mutations. In order to find a more accurate analysis, and distinguish between COVID-19 patients, SARS-CoV-2 virus traits, and normal people, the complex classification problems of SARS-CoV-2 can be formulated into the pattern recognition problems of virus genes to build a robust diagnostic system (Abuzanouneh andAl-Shqeerat, 2021).MOGA will be applied based on colorimetric features to describe the colors and results of the sample genes, which are the most discriminative features of sample analysis of RT- PCR. A sample gene diagnosis is a target object, the MOGA selection operator is used to detect the test values for genes with the same features as the SARS-CoV-2 virus traits of the current population (test samples) that will be used in the next stage (amplification cycles) (Abuzanouneh, 2017). The convergence of the identification process can be completed and achieved under the required iterations, and the number of amplification processes is performed to diagnose and detect coronavirus using a minimum amplification cycle for each generation of population. In the next step of the test sample, the results represented the intensity of colorimetric features that can be calculated and detected from examining the samples. The colorimetric process is a digital color processing where pixel values are modeled as analysis results. The mutation and selection process based on the proposed algorithm is used to store new data from new mutations of the virus and integrate it into the system dataset to accelerate and improve the research ability of individuals and the new environment (Abuzanouneh. 2017).The fitness function of the proposed algorithm is used to evaluate the probability of color pixel levels reflecting the test results. The fitness functions are used to obtain robust solutions and results using a simulation system for diagnosis and analysis as much as possible with reliable results (Hahanov and Abu Zanouneh, 2001).The best individuals from the test samples will be selected after the amplification process to create new tables containing deep analysis (details information) and optimal solutions (accuracy results) to be as references in the new diagnosis processes of new mutations of virus (SARS-CoV2).

Specimen Collection Diagnostic

The test of SARS-CoV-2 nucleic acid is used for qualitative detection by targeting N, E, and ORF1ab region proteins of the viral genome. These Three target regions for the SARS-CoV-2 test using specific primers to amplify the conserved segments of the protein regions are used.
Sample test results help in the laboratory diagnosis process as a basis for obtaining final results such as correct, invalid, negative or positive sample (Weissleder, Lee, Ko, and Pittet, 2020; Corman, Landt, Kaiser, Molenkamp, Meijer, Chu, Bleicker, Brünink, Schneider,

Schmidt, Mulders, Haagmans, Veer, Brink, Wijsman, Goderski, Romette, Ellis, Zambon, Peiris, Goossens, Reusken, Koopmans, and Drosten, 2020). MOGA-based procedure model of sample test is validated as follows.

- Conditions of Specimen Collection:

1. Specimens collection should avoid possible contamination in transportation and storage.
2. Specimens must be assumed contagious and processed according to health regulations.
3. The swabs of collection should have a synthetic head, such as Dacron or nylon, and plastic or an aluminum shaft.
4. Calcium alginate swabs are not accepted and the wooden shaft cotton swab is not recommended.
5. After completion of the sampling phase, the swabs should be saved in the Sample Storage Reagent (SSR) directly.

- Storage and Delivery Conditions of Specimens:

1. Specimens must be processed immediately.
2. If the Specimens are stored at 4℃, they should be tested within 24 hours or less.
3. Specimens should be tested within 24 hours if stored at 4℃.
4. Specimens should be stored at -70℃ minimum if they are not tested within 24 hours.
5. Specimens should be stored at -20℃ for no more than 10 days.
6. Nucleic of acid specimens should be stored at -30 ℃ for a maximum of 15 days.
7. Multiple thawing and freezing cycles should be avoided as much as possible.
8. Specimens transported must be in a sealed frozen box.

- Laboratory Conditions of Sample Extraction:

1. Clinical specimens should be kept in a sample storage reagent (SSR).
2. Sample processing with SSR RNA rapid release technology should be used, which is usually provided in the kit.
3. In a 1.5 ml, EP tube a pipette should add 200 µl of the sample, and centrifuge for 5 minutes (12,000 rpm).
4. The supernatant fluid should be discarded carefully.
5. Precipitation at the bottom should avoid removal.
6. For each tube, add 50 µL SRR using vortex for five seconds.
7. In the last step, the lysed sample should be added directly to the rRT-PCR reaction.

Results Interpretation and Discussion

If the curve is S-shaped and two targets or more (ORF1ab, N, E gene) are detected, and the exponential growth phase is very large, then the analysis will give a positive result for SARS-CoV-2 nucleic acid.Chromosomes in genetic algorithms (also called genotypes) are a set of parameters that determine a proposed solution to the problem that the GA is trying to solve. Chromosomes are represented as binary sequences, although different types of other data forms can be used (Wee, Sivalingam, andYap, 2020; Lamb,Bartolone, Ward, and Chancellor, 2020).The fitness value of standard curve f (c) is positive (1) when the curve is S-shaped, while the fitness value is negative (0) if the curve is not S-shaped, as shown in equation (1).

$$f(\mathrm{c}) = \begin{cases} 1 & if\ c = S \\ 0 & if\ c \neq S \end{cases} \quad (1)$$

For the gene sample, the chromosome is represented by three target sequence values (ORF1ab, N, and E protein genes) of the Virus Gene (Vg) for SARS-CoV-2. A set of specific primers were designed to amplify conservative segments of these target regions.

If the number of targets (ORF1ab gene, N, E gene) greater than or equal to two, The result of targets will be positive and detected. So, the fitness value of ORF1ab f(O) should be equal to 1, where O stands for ORF1ab gene.

$$f(\mathrm{O}) = \begin{cases} 1 & if\ \ \mathrm{ORF1ab} = \mathrm{Vg} \\ 0 & if\ \ \mathrm{ORF1ab} \neq \mathrm{Vg} \end{cases} \quad (2)$$

$$f(\mathrm{N}) = \begin{cases} 1 & if N = \mathrm{Vg} \\ 0 & if N \neq \mathrm{Vg} \end{cases} \quad (3)$$

$$f(\mathrm{E}) = \begin{cases} 1 & if E = \mathrm{Vg} \\ 0 & if E \neq \mathrm{Vg} \end{cases} \quad (4)$$

Sample analysis must be repeated when one target result is detected. If two targets are detected after repetition, and if the exponential growth phase is very large and the virus curve is S-shaped, then the analysis result will give a positive result for SARS-CoV-2 nucleic acid (Shani-Narkiss, Gilday, Yayon, Landau, 2020).If the detected target is still one in the sample, the sample results are suspected, so the result will be positive, and the sample detection must be repeated and the sequence determined again using other sequencing methods.The negative result will be detected when the three channels (ORF1ab, N, and E gene) are not detected, the IC detection result is positive, and the analysis result will give a negative SARS-CoV-2 DNA result (Shani-Narkiss, Gilday, Yayon, and Landau, 2020).

Positive and Negative Test Control

Positive controls (Pc) are used to monitor whether RT-PCR processes are working properly and are used in each detection of SARS-CoV-2 DNA results. There are two cases of positive control (valid = 1, invalid = 0).

Fitness functions of the proposed algorithm are used to check positive and negative values; f (Pc) and f (Nc) are used to evaluate the result of solution quality [18].

$$f\ (\mathrm{P}_c) = \begin{cases} validif P_c = 1 \\ invalidif P_c = 0 \end{cases} \qquad (5)$$

Negative Controls f (Nc) are used to monitor whether there are contaminations for RT-PCR processes and are used in each detection of SARS-CoV-2 nucleic acid results. There are two cases of the negative control (valid =0, invalid=1)

$$f(\mathrm{N}_c) = \begin{cases} validif \mathrm{N}_c = 0 \\ invalidif \mathrm{N}_c = 1 \end{cases} \qquad (6)$$

Internal controls (ICs) are used to target the RNAse gene and monitor the sample collection, and to handle the RT-PCR process for each sample amplification (Fomsgaard andRosenstierne, 2020). The positive control and negative control for each run are illustrated in Table 1 and Table 2.

Experimental Results

MOGA examines test materials using control values before interpretingpatient results. Patient results cannot be reliable when positive or negative controls are not valid based on their fitness values (Schmid-Burgk, Schmithausen, Li, Hollstein, Shmuel, Israeli, Weiss, Paran, Wilbring, Liebing, Feldman, Słabicki, Lippke, Sib, Borrajo, Strecker, Reinhardt, Hoffmann, Cleary, Hölzel, Nöthen, Exner, Ludwig, Regev, and Zhang 2020).Generally, diagnostic laboratories choose a cut-off value for RT-PCR amplification assays, if the value rate more than the cycle threshold (Ct),the value is considered false (Ribeiro da Silva, Pardee, and Pena, 2019).High Ct value is interpreted as contaminations or fluorescence artifact that can lead to positives or negatives results, or fault amplification. The execution of Ct cut-off might be acceptable, but its selection and justification should be based on experiments and proofs (Broughton, Deng, Yu, Fasching, Servellita, Singh, Miao, Streithorst, Granados, Sotomayor-Gonzalez, Zorn, Gopez, Hsu, Gu, Miller, Pan, Guevara, Wadford, Chen, and Chiu, 2020;Huang,Wang, and Li, 2020).

In this study,we selected Ct cutoffs grouped in analytical and epidemiologic approaches of COVID-19. Analytical strategies use criteria collected in the assay development including reaction end-cycle, fluorescence threshold, and detection limit. The Ct cut-off value is set to 40, the fluorescent curves before final interpretation should be reviewed by End-user before final interpretation, and the amplification curves should be S-shaped to be a positive value (Klinkenberg and Wallinga, 2020; Wu and Leung, 2020; Mizumoto and Chowell, 2020).

Table 1

Positive Control Interpretation for each run

	Positive Control(+)				MOGA	
Test	**ORF1ab gene**	**N gene**	**E gene**	**PC**	**F (Pc)**	**Actions**
0	1	1	1	1	Valid	Result interpretation
1	1	1	1	0	Invalid	RT-PCR contaminated
2	1	1	0	1	Negative	
3	1	1	0	0	Invalid	RT-PCR contaminated
4	1	0	1	1	Negative	
5	1	0	1	0	Invalid	RT-PCRcontaminated
6	1	0	0	1	Negative	
7	1	0	0	0	Invalid	RT-PCRcontaminated,
8	0	1	1	1	Negative	
9	0	1	1	0	Invalid	RT-PCR contaminated
10	0	1	0	1	Negative	
11	0	1	0	0	Invalid	RT-PCR contaminated
12	0	0	1	1	Negative	
13	0	0	1	0	Invalid	RT-PCR contaminated
14	0	0	0	1	Negative	
15	0	0	0	0	Invalid	RT-PCR contaminated

Table 2

Negative control Interpretation for each run.

	Negative Control(-)				MOGA	
Test	**ORF1abgene**	**N gene**	**E gene**	**Nc**	**F4(Nc)**	**Actions**
0	1	1	1	1	Invalid	RT-PCR contaminated
1	1	1	1	0	Negative	
2	1	1	0	1	Invalid	RT-PCR contaminated
3	1	1	0	0	Negative	
4	1	0	1	1	Invalid	RT-PCR contaminated
5	1	0	1	0	Negative	
6	1	0	0	1	Invalid	RT-PCR contaminated
7	1	0	0	0	Negative	
8	0	1	1	1	Invalid	RT-PCR contaminated
9	0	1	1	0	Negative	
10	0	1	0	1	Invalid	RT-PCR contaminated

11	0	1	0	0	Negative	
12	0	0	1	1	Invalid	RT-PCR contaminated
13	0	0	1	0	Negative	
14	0	0	0	1	Invalid	RT-PCR contaminated
15	0	0	0	0	Valid	Result interpretation

The fitness method for gene analysis F(A) is described in Equation (7).

$$F(A) = \mathrm{F(C)} + \mathrm{F(O)} + \mathrm{F(N)} + \mathrm{F(E)} \qquad (7)$$

1. If the positive control equals 1, the result of F(Pc) is valid, otherwise it is invalid.

2. If the negative control is 0, the result of F (Nc) is valid, otherwise invalid.

3. If the positive and negative control is valid and any value of the analysis gene is negative, the final result of the fitness method F(A) is negative.

4. If the positive and negative control are valid and all genes of the analysis are positive, the final result of the fitness method F(A) is positive as shown in Table 4.

Table 3 shows F(A) fit analysis based on internal controls (positive, negative) for the proposed algorithm for each run shown in Table 4.

Table 3

Fit analysis based on internal controls (positive, negative)

Test	ORF1ab	N	E	Pc	Nc	F (Pc)	F (Nc)	F (A)
0	1	1	1	1	1	Valid	Invalid	Invalid
1	1	1	1	1	0	Valid	Valid	Positive
2	1	1	1	0	1	Invalid	Invalid	Invalid
3	1	1	1	0	0	Invalid	Valid	Invalid
4	1	1	0	1	1	Valid	Invalid	Invalid
5	1	1	0	1	0	Valid	Valid	Negative
6	1	1	0	0	1	Invalid	Invalid	Invalid
7	1	1	0	0	0	Invalid	Valid	Invalid
8	1	0	1	1	1	valid	Invalid	Invalid
9	1	0	1	1	0	Valid	Valid	Negative
10	1	0	1	0	1	Invalid	Invalid	Invalid
11	1	0	1	0	0	Invalid	Valid	Invalid
12	1	0	0	1	1	valid	Invalid	Invalid
13	1	0	0	1	0	valid	Valid	Negative
14	1	0	0	0	1	Invalid	Invalid	Invalid
15	1	0	0	0	0	Invalid	Valid	Invalid
16	0	1	1	1	1	valid	Invalid	Invalid
17	0	1	1	1	0	valid	valid	Negative
18	0	1	1	0	1	Invalid	Invalid	Invalid
19	0	1	1	0	0	Invalid	Valid	Invalid
20	0	1	0	1	1	Valid	Invalid	Invalid
21	0	1	0	1	0	Valid	valid	Negative

22	0	1	0	0	1	Invalid	Invalid	Invalid
23	0	1	0	0	0	Invalid	Valid	Invalid
24	0	0	1	1	1	Valid	Invalid	Invalid
25	0	0	1	1	0	Valid	Valid	Negative
26	0	0	1	0	1	Invalid	Invalid	Invalid
27	0	0	1	0	0	Invalid	Valid	Invalid
28	0	0	0	1	1	valid	Invalid	Invalid
29	0	0	0	1	0	Valid	Valid	Negative
30	0	0	0	0	1	Invalid	Invalid	Invalid
31	0	0	0	0	0	Invalid	Valid	Invalid

Table 4

Invalid Results of sample based on Positive and Negative Control and fitness method of MOGA.

Test	**ORF1ab**	**N**	**E**	**Pc**	**Nc**	**F (Pc)**	**F4(Nc)**	**F (A)**
0	1	1	1	1	1	Valid	Invalid	Invalid
2	1	1	1	0	1	Invalid	Invalid	Invalid
3	1	1	1	0	0	Invalid	Valid	Invalid
4	1	1	0	1	1	Valid	Invalid	Invalid
6	1	1	0	0	1	Invalid	Invalid	Invalid
7	1	1	0	0	0	Invalid	Valid	Invalid
8	1	0	1	1	1	valid	Invalid	Invalid
10	1	0	1	0	1	Invalid	Invalid	Invalid
11	1	0	1	0	0	Invalid	Valid	Invalid
12	1	0	0	1	1	valid	Invalid	Invalid
14	1	0	0	0	1	Invalid	Invalid	Invalid
15	1	0	0	0	0	Invalid	Valid	Invalid
16	0	1	1	1	1	valid	Invalid	Invalid
18	0	1	1	0	1	Invalid	Invalid	Invalid
19	0	1	1	0	0	Invalid	Valid	Invalid
20	0	1	0	1	1	Valid	Invalid	Invalid
22	0	1	0	0	1	Invalid	Invalid	Invalid
23	0	1	0	0	0	Invalid	Valid	Invalid
24	0	0	1	1	1	Valid	Invalid	Invalid
26	0	0	1	0	1	Invalid	Invalid	Invalid
27	0	0	1	0	0	Invalid	Valid	Invalid
28	0	0	0	1	1	valid	Invalid	Invalid
30	0	0	0	0	1	Invalid	Invalid	Invalid
31	0	0	0	0	0	Invalid	Valid	Invalid

Table 5showsthe negative results of sample analysis based on a valid value of positive and negative controlsusing the proposed algorithm.Table 6showsthe positive results of sample analysis based on genes analysis using the proposed algorithm.A sample result check is an evaluation of the sample test results that should be made based on the positive and negative

control values tested and confirmed as acceptable and valid. Patient sample results cannot be interpreted if control values are invalid.

Table 5

Negative results of sample analysis using the proposed algorithm

Test	ORF1ab	N	E	Pc	Nc	F (Pc)	F (Nc)	F (A)
5	1	1	0	1	0	Valid	Valid	Negative
9	1	0	1	1	0	Valid	Valid	Negative
13	1	0	0	1	0	valid	Valid	Negative
17	0	1	1	1	0	valid	valid	Negative
21	0	1	0	1	0	Valid	valid	Negative
25	0	0	1	1	0	Valid	Valid	Negative
29	0	0	0	1	0	Valid	Valid	Negative

Table 6

Positive results of sample analysis using the proposed algorithm

Test	ORF1ab	N	E	Pc	Nc	F (Pc)	F (Nc)	F (A)
1	1	1	1	1	0	Valid	Valid	Positive

Table 7 shows the interpretation of the result with respect to control values (Wang, Hu, and Hu C., 2020; WHO, 2020). For example, the cycle threshold (Ct) values for this group are set to 40 and the fluorescent curve values for end-users must be reviewed and evaluated before the final results are interpreted. In addition, all positive curve values should be represented by S-shaped amplification curve values and without plateau-shape values for incomplete positive samples, in this case, the cycle threshold can oscillate between 38 and 40.

Table 7

Positive and Negative Control Values.

Pc(+)	**Nc(-)**	**Pc ∧ ~ Nc**
+	+	-
+	-	+
-	+	-
-	-	-
where (+) stands for valid test and (-) stands for invalid test		

Table 8. shows the test table of the Multi-target gene (O, N, E), where O stands for ORF1ab gene.

Table 8

Test table of Multi-target gene

O	**N**	**E**	**O ∧ N ∧ E**
+	+	+	+
+	+	+	-
+	-	-	+
+	-	-	-
-	+	+	+
-	+	+	-
-	-	-	+
-	-	-	-
where (+) stands for positive test and (-) stands for negative test.			

Diagnosis and Analysis Methods

MOGA is used to research and describe test models and Lab procedures for SARS-CoV-2 nucleic acid. The structural logical description of lab test behavior, and technical diagnostic methods for control and detection of invalid results. In order to realize the program, Python was used to diagnose the proposed algorithm.

Solving problems in laboratory diagnostics is best implemented using the quadrant parameters <T, (M, D), R>: <test, (model, defect), reaction>, which is involved in defining the technical condition of the sample analysis. A sample analysis diagnosis can be created, where M identifies a valid state of the sample analysis, and defects D for the invalid state of a sample analysis that does not take a common sample test model (M, D) from the matrix of permissible components (M x D) (Hahanov, Hanko, and Abu Zanuneh, 1999). The general equation of lab test model in the presence or absence of faults:

$$f(M,T,R) \oplus f^{*}(M,T,D,R) = 1 \quad (8)$$

The formalizing of the problem can be illustrated as follows:

1. Creatinglab test model for test generation and defect modeling as functional regimes and relevant reactions at faults absence.

$$M(D = \emptyset) = f^{1}(T,D,R) (9)$$

2. Determining the part of suspected defects in the sample test according to its reaction on test regimes in one swab test.

$$D(R_{I}^{*} \in R) = f^{1}(T,D,R) \quad (10)$$

3. Reducing and detecting the suspected defect part in the swab sample of its reaction on test regimes unacceptable results.

$$D(R^* \in R) = f^2(T, M, R^*) \quad (11)$$

The diagnostic system test strategy is determined by using one instrument as a test control (Tc) and analyzer in good condition, and the rest of the sample test equipment is an object of diagnostics. The proposed algorithm for detectinga fault in the test analysis based on which algorithms are developed: structural analysis of fault multi ciphered tables (DMT) which minimize the suspected fault area; back watch and half division using probe element which provides the given diagnosed depth of samples.There are two indicators of the test sample valid and invalid, where {TV, TI} = {0,1}; 1 (0)– indicator of valid or invalid testing.

$$T = Tc \wedge O1 \wedge E1 \wedge N1 \vee O2 \wedge E2 \wedge N2 \vee O2 \wedge E2 \wedge N2$$
$$\vee \ldots \wedge Oi \wedge Ei \wedge Ni \quad (12)$$

Where Tc means Test Control of the sample (0,1).

A structural algorithm for diagnosis that is not strongly differentiated in control and error-finding procedures is described by the general fitness equation for defect diagnosis:

$$\boldsymbol{F^*(F,T,D)_{F,T,g^*} = \bigcup_i (g(T,F) \bigcap g^*(T,F,D_i)) = \varnothing} \quad \boldsymbol{(13)}$$

where *g, g**– standard and the experimental reaction of the watched lines;*F, T,* – functional which check the test on defects given beforehand.

$$\boldsymbol{P_i = \begin{cases} P_i^+ \Leftarrow \left[g^*(T,F,D_i) \bigcap g(T,F) = g(T,F)\right] \\ P_i^- \Leftarrow \left[g^*(T,F,D_i) \bigcap g(T,F) = \varnothing\right] \end{cases}} \quad (14)$$

The use of logical analysis determines the elemental check*Pi=g*(T, F, Di)* as receivinga reaction to the sample test chosen when the*T*test is performed. A *Pi* check is considered positive if at the control point *i* the standard reaction is done at the absence of defects in the scheme and it is equal to the experimental one: *SM = SR*. If *SM* is not equal to *SR*, then*Pi* checking is negative. In the general case, probing state is determined by the expression:

$$\boldsymbol{D_s = \begin{cases} \{D_0, D^+\} \Leftarrow P_i = P_i^+ \\ D^- \Leftarrow P_i = P_i^- \end{cases}} \quad (15)$$

Where *Pi* control point is accessible for examining the line of valid or invalid test.

Two logical test results give alternative ways of fault search according to the structure and or technical condition of the object state, where F∑ = F0 ∪ F; F+ ∪ F- = F; F+ ∩ F- = ∅; F0 –

satisfy technical condition;F – fault presence of supposed test; F+ (F–) – fault of test which is determined by positive (negative) check F+ (F–).

The model faults of the sample test are determined a set of conditions and a multi-objective genetic algorithm as follows:∑F = {F0, F1, F2, F3, F4, F5, F6, F7, ……},whereF0 – Sample test is functioning;F1 – All test genes have a valid state;F2 – Error results in the gene examination;F3 – Sample contamination;F4 – Specimens is stored at 4℃ and tested within 24;F5 – Specimens are stored at -70℃ if not tested.F6 – Multiple freezes of Specimens. F7 – Specimens are transported in a sealed frozen pitcher.

The algorithm check for errors shows the effect check according to the structure of the sample test directed to the successive checks of the lines of the logic model to recover them to work properly.Tests for each gene from thesample are:

1. Test O genes {O1 ... O8} = {L1,L4, L7, L10,L13, L16, L19, L22};

TO= L1∨ L4 ∨L7 ∨ L10 ∨ L13 ∨ L16 ∨ L119 ∨ L22.

2. Test E genes {E1 ... E8} = {L2, L5, L8, L11, L14,L17, L20, L23};

TE= {L2 ∨ L5 ∨ L8 ∨ L11 ∨ L14 ∨ L17, L20, L23};

3. Test N genes {N1 ... N8} = {L3,L6, L9, L12, L15,L18,L21, L24,};

TN= {L3 ∨ L6 ∨ L9 ∨ L12 ∨ L15 ∨ L18,L21, L24};

4. Test status of genes:

T1(O1 ∨N1 ∨E1) =L25; T2(O2 ∨N2 ∨E2) = L26;

T3(O3 ∨N3 ∨E3) =L27; T4(O4 ∨N4 ∨E4) = L28;

T5(O5 ∨N5∨E5) =L29; T4(O6 ∨N6 ∨E6) = L30;

T7(O7 ∨N7 ∨E7) =L31; T8(O8 ∨N8 ∨E8) = L32;

5. Test status of each sample:

Si ={L33, L34, L35, L36, L37, L38, L39, L40};

6. Test status of all samples:

Si = {L33, L34, L35, L36, L37, L38, L39, L40};

7. The final status of the swabs samples: Si={L41}.

Figure 1. illustrates the Logical Diagnostic Scheme for Sample Test.The correct behavior of the sample test is modeled.It gives the vector along all numbered lines=1.

The input test case is T_{41} = (111111111111...111111111111). Table. 9shows faults finding method based on the diagnosis scheme.

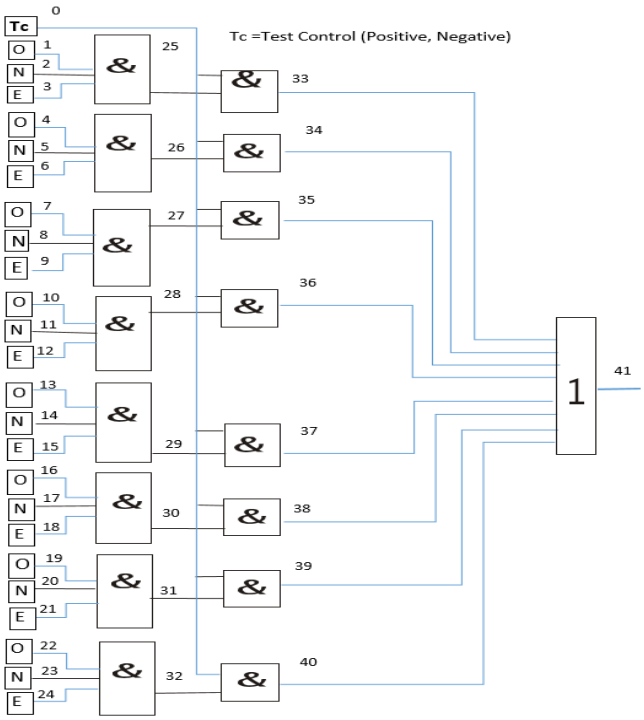

***Figure 1.*Logical Diagnostic Scheme for Sample Test.**

Table 9

Faults finding method based on the diagnosis scheme.

Pi	41	40	39	38	37	36	35	34	33
P+	F0	39	38	37	36	35	34	33	F1
P-	40	32	31	30	29	28	27	26	25

Pi	32	31	30	29	28	27	26	25	24
P+	F2	F2	F2	F2	F2	F2	F2	F2	23
P-	F4	21	18	15	12	9	6	F3	F5

Pi	23	22	21	20	19	18	17	16	15
P+	22	F8	20	19	F8	17	16	F8	14
P-	F6	F7	F5	F6	F7	F5	F6	F7	F5

Pi	14	13	12	11	10	9	8	7	6
P+	13	F8	11	10	F8	8	7	F8	5
P-	F6	F7	F5	F6	F7	F5	F6	F7	F5

Pi	5	4	3	2	1
P+	4	F8	2	1	F8
P-	F6	F7	F5	F6	F7

Figure2 showsfitness values and generation numbers using 100 SARS-CoV-2test samples.

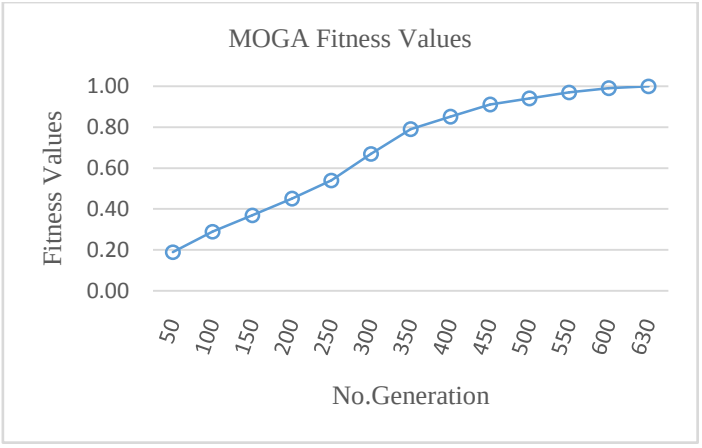

***Figure 2.*fitness values and generation number using 100 samples.**

Figure 2 shows the fitness values and the number of generations using MOGA, where the number of samples =100, the maximum value of generation =630, and the fitness value is 1.0. The result shows the ability of the MOGA algorithm to efficiently complete the diagnosis process for all samples. It can be seen that the fitness value (1.0) obtained for generation (630) is ideal. The program of the proposed system is implemented to develop MOGA in the medical field, with a diagnosis method for the SARS-CoV-2 detection. The application can be used to prepare deep diagnostic simulation experiments in the early stage of samples testing automatically. The proposed system can be used to analyze assay samples to be more accurate diagnose for covid-19 disease.

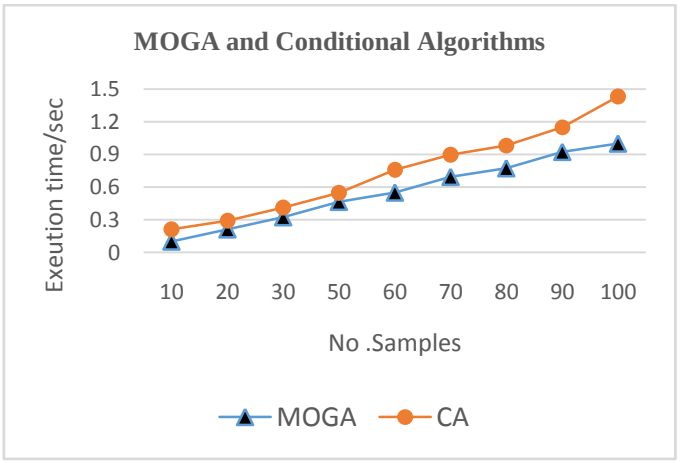

***Figure 3.*Results of experiments on SARS-CoV-2 sample test models.**

The result of experiments on SARS-CoV-2 test samplesaccording to the strategy which combines successive fulfillment of the diagnosis algorithms using MOGA and conditional algorithms (CA) for search faults as shown in Figure 3.It illustrates a significant reduction in the maximum number of probes compared to conditional algorithms.

Conclusion

This research developed a multi-objective genetic algorithm (MOGA) and diagnosis techniques to detectthe COVID-19 virus. MOGA is used as a specificresearch methodology to analyze and control the covid-19 disease. We provide a novel approach forSARS-CoV-2 gene sample diagnosis and analysis at the early detection stage.

In this study, we applied MOGA and the diagnostic model to detect the genetic status of SARS-CoV-2, which are the most discriminative features of coronavirus.MOGA Genetic Operators are performed to diagnose the suspected objects with the same virus-genefeatures using the existing population (samples swabs). MOGA and diagnosis techniques are used to examine and explore the structural changes of SARS-CoV-2 using different fitness values to detect the nucleic acids that are performed by molecular diagnostic assays based on new chemical and physical characteristics of COVID-19 swab.The structural logic of the technical diagnostic method is implemented to describe thebehavior of a laboratory test to diagnose the final results of the SARS-CoV-2gene sample.The proposed algorithm is an efficient technique to diagnose and explore gene mutations of the SARS-CoV-2 virus.

Acknowledgment

The authors gratefully acknowledge Qassim University, represented by the Deanship of Scientific Research, on the material support for this research under project number: coc-2020-1-1-L-10020 during the academic year 1441 AH/ 2020 AD.

References

1. N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan. (2020).*China Novel Coronavirus Investigating and Research Team, A novel coronavirus from patients with pneumonia in China*, N. Engl. J. Med. 382, 727–733.
2. J. Phua, L. Weng, L. Ling, M. Egi, C.-M. Lim, J. V. Divatia, B. R. Shrestha, Y. M. Arabi, J. Ng, C. D. Gomersall, M. Nishimura, Y. Koh, B. Du. (2020) *Intensive care management of coronavirus disease 2019 (COVID-19): Challenges and recommendations*. Lancet Respir. Med. 8, 506–517.
3. K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell.(2020). *Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship*, Yokohama, Japan,. Euro Surveill. 25, 2000180.
4. *Laboratory testing for 2019 novel coronavirus (SARS-CoV-2) in suspected human cases*.World Health Organization(2020)
5. *Laboratory biosafety guidance related to the novel coronavirus (2019-nCoV)*. Interim Recommendations WHO; 28 February (2020).

6. JB.Mahony.(2008).*Detection of Respiratory Viruses by Molecular Methods. Clinical Microbiology Reviews,* 21(4):716-747.
7. Chinese Center for Disease Control and Prevention. (2020).*Laboratory guidelines for novel coronavirus infection in pneumonia (Second Edition)*.
8. L. Yu, S. Wu, X. Hao, X. Dong, L. Mao, V. Pelechano, W.-H. Chen, X. Yin. (2020).*Rapid detection of COVID-19 coronavirus using a reverse transcriptional loop-mediated isothermal amplification (RT-LAMP) diagnostic platform*. Clin. Chem. 66, 975–977.
9. Khalil Ibrahim Mohammad Abuzanouneh, Khalil Hamdi Ateyeh Al-Shqeerat. (2021). *Development and Improvement of Web Services, Selections Using Immigrants Scheme of Multi-Objective Genetic Algorithm*. Advances in Science, Technology and Engineering Systems Journal Vol. 6, No. 2, 401-411.
10. Khalil Ibrahim Mohammad Abuzanouneh. (2017). *New Image Processing Techniques Using Elitism Immigrants Multiple Objective of Genetic Algorithms for Disease Detection*. International Journal of Computer Science and Information Security (IJCSIS), Vol. 15, No. 12.
11. Khalil Ibrahim Mohammad Abuzanouneh. (2017).*Hybrid Multi Objectives Genetic Algorithms and Immigrants Scheme for Dynamic Routing Problems in Mobile Networks*. International Journal of Computer Applications 164(5): 49 -57.
12. V.I . Hahanov, Babich A.V., Abu Zanuneh I.M. Halil. (2001). *Designing of LAN models for diagnosis problem solving – The Experience of Designing and Application of CAD System in Microelectroncs*. Proceeding of the VI-th Conference CADSM 2001, Lviv,Slavsko, 307-309.
13. R. Weissleder, H. Lee, J. Ko, M. J. Pittet (2020). *COVID-19 diagnostics in context*. Sci. Transl. Med. 12, eabc1931.
14. V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. W. Chu, T. Bleicker, S. Brünink, J. Schneider, M. L. Schmidt, D. G. J. C. Mulders, B. L. Haagmans, B. van der Veer, S. van den Brink, L. Wijsman, G. Goderski, J.-L. Romette, J. Ellis, M. Zambon, M. Peiris, H. Goossens, C. Reusken, M. P. Koopmans, C. Drosten. (2020). *Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR*. Euro Surveill. 25, 2000045.
15. S. K. Wee, S. P. Sivalingam, E. P. H. Yap.(2020).*Rapid direct nucleic acid amplification test without RNA extraction for SARS-CoV-2 using a portable PCR thermocycler*. https://doi.org/10.3390/genes11060664.
16. L. E. Lamb, S. N. Bartolone, E. Ward, M. B. Chancellor. (2020).*Rapid detection of novel coronavirus/Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by reverse transcription-loop-mediated isothermal amplification*. PLOS ONE 15, e0234682.
17. H. Shani-Narkiss, O. D. Gilday, N. Yayon, I. D. Landau.(2020). *Efficient and practical sample pooling for high-throughput PCR diagnosis of COVID-19*. medRxiv 2020.04.06.20052159. https://doi.org/10.1101/2020.04.06.20052159.
    A. S. Fomsgaard, M. W. Rosenstierne. (2020). *An alternative workflow for molecular detection of SARS-CoV-2* – escape from the NA extraction kit-shortage, Copenhagen, Denmark, March 2020. Euro Surveill. 25, 2000398.

18. J. L. Schmid-Burgk, R. M. Schmithausen, D. Li, R. Hollstein, A. Ben-Shmuel, O. Israeli, S. Weiss, N. Paran, G. Wilbring, J. Liebing, D. Feldman, M. Słabicki, B. Lippke, E. Sib, J. Borrajo, J. Strecker, J. Reinhardt, P. Hoffmann, B. Cleary, M. Hölzel, M. M. Nöthen, M. Exner, K. U. Ludwig, A. Regev, F. Zhang (2020).*LAMP-Seq: Population-Scale COVID-19 Diagnostics Using Combinatorial Barcoding*. bioRxiv 2020.04.06.025635 https://doi.org/10.1101/2020.04.06.025635.
19. S. J. Ribeiro da Silva, K. Pardee, L. Pena. (2019). *Loop-mediated isothermal amplification (LAMP) for the diagnosis of Zika virus: A review*. Viruses 12, 19.
20. J. P. Broughton, X. Deng, G. Yu, C. L. Fasching, V. Servellita, J. Singh, X. Miao, J. A. Streithorst, A. Granados, A. Sotomayor-Gonzalez, K. Zorn, A. Gopez, E. Hsu, W. Gu, S. Miller, C.-Y. Pan, H. Guevara, D. A. Wadford, J. S. Chen, C. Y. Chiu. (2020). *CRISPR–Cas12-based detection of SARS-CoV-2*. Nat. Biotechnol. 38, 870–874.
21. Huang C .Wang Y .Li X.(2020). *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China,*Lancet, 395: 497-506.
22. Backer JA Klinkenberg D Wallinga J. (2020). *Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan*, China, 20–28.
23. J. T. Wu, K. Leung, and G. M. Leung. (2020).*Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study*. Lancet, vol. 395, no. 10225, pp. 689–697.
24. Mizumoto K Chowell G. (2020).*Estimating the risk of 2019 novel coronavirus death during the course of the outbreak in China*. medRxiv. (published online Feb 23.) (preprint). DOI:10.1101/2020.02.19.20025163
25. Wang D, Hu B, Hu C (2020). *Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan*. China,JAMA. (published online Feb 7.) DOI:10.1001/jama.2020.1585
26. WHO (2020). *Coronavirus disease 2019 (COVID-19)*. SituationReport, 41. https://www.who.int/docs/default-source/ coronaviruse /situation-reports/20200301-sitrep-41-covid-19.pdf? sfvrsn=6768306d_2
27. V.I Hahanov, V.V Hanko and I.M. Halil Abu Zanuneh. (1999).*Network fault models and formulation of their diagnostics tasks*. Radio electronics and informatics, no. 4, pp. 49-55.