European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 05, 2020

Propagation of Situational Information about Covid-19 and Detection with Ranking-Based Diversity Reduction Strategy

Dr. Ch. Ashok kumar[1], R. Lakshmi Priya[2], Rajesh Kannan K[3]

[1]*Assistant Professor, Department of Information Technology, PSNA College of Engineering and Technology, dindigul, Tamilnadu, India.*

[2]*Associate Professor, Department of CSE, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India.*

[3]*Assistant Professor, Department of Information Technology, Chaitanya Bharathi Institute of Technology (Autonomous) , Gandipet, Hyderabad, Telangana, India.*

Abstract: In the present situation COVID-19 is the critical problem facing in all over the world. Therefore for detecting and diagnosing this virus effectively, quickly and with a low cost on a lot of tests, several countries are facing the problem of lacking the medical resources. In general, COVID-19 causing the pulmonary opalescence in frosted glass and bilateral pulmonary parenchyma, sometimes with a rounded morphology and peripheral pulmonary distribution. Therefore, main aim of this paper is rapidly extracting the small regions from the chest which can characterize the features of COVID-19 through the X-Ray images. For diagnosing these X-ray images to detect the virus the segmentation method of marine predator's algorithm is mostly used. Hence to get a best solution within a few iterations, a strategy of Ranking based Diversity Reduction (RDR) is proposed in this paper which improves the performance of the marine predators algorithm. Moreover, to this ongoing outbreak of corona virus disease (COVID-19), it is important to provide valuable situational information for the public and authorities to know in which manner it is being spreading over society. Hence, to it is required to provide appropriate information. So, the appropriate information publishing strategy of Weibo data is used to classify information related to the COVID-19 epidemic into seven types of situational information are proposed in this paper.

Keywords: **COVID-19, ranking-based reduction diversity, information propagation, marine predators algorithm.**

1. INTRODUCTION:

In several countries, the testing of COVID-19 is doing on a few numbers of populations only because the existing diagnosis tools are in limited number [1, 2]. In many countries requiring the medical resources is a main problem faced in regarding to a great effort of finding an efficient method to detect COVID-19. Therefore for effectively recognizing and diagnosing COVID-19, it is essential to find a quick and inexpensive tool. It was found by Guan in [2] that computed tomography analysis incorporated the bilateral pulmonary parenchymal ground-glass and consolidative pulmonary opacities, sometimes with rounded morphology and peripheral pulmonary distribution when performing computed tomography scans on chest from the COVID-19 infected patients of 21 members in China. Hence for the extraction of main disease characteristics, the diagnosis of COVID-19 is presented as a problem of image segmentation. An algorithm is developed in order to solve the problem of segmentation

which is capable of extracting similar smaller regions that indicates a COVID-19 virus infection.

The ability of segmenting the image with much similar regions which observes with a high threshold values of an image have been not validated by any method that represented in the literature even though there exist many methods for segmenting medical images. As a result, these algorithms may not be the better option to search in medical images for smaller homogeneous regions which might include the COVID-19 disease characteristics. This challenge inspires to monitor some state-of-art algorithms performance for tackling the ISP proposed in the literature. Also this paper proposed a new robust method of RDR that moves the worst solutions positions nearly close to the better solution for improving performance of Marine Predators Algorithm.

The ongoing outbreak of the COVID-19 (corona) virus disease has been caused in Wuhan, China which leads a regional and global public health crisis [3]. Public utilize a platform of social media in the crises period i.e., like COVID-19 epidemic for getting the related necessitate information and can share opinions on it. In a social media platform different types of information's are present. The information that will help authorities or interested people in an emergency for understanding the situations including usable information such as seeking help, the number of people affected [5] is called as situational information that is helpful to the public and the authorities to guide their responses [4], [6]. In order to perceive the public mood, the gaps in information between the authority and public and the need of information for public, the above type of information is identified and predicted the extent of its dissemination which will give advantage to the associated authorities. It would therefore support the authorities in developing appropriate emergency strategies [4]. The prediction of propagation scale on the situational information was ensured that different types of situational information's can be published by the related authorities depending on the people requirements. So finding the key features is important since that features plays a crucial role in the prediction process. Meeting this need for information is critical in case of sudden epidemics like COVID-19.

2. MARINE PREDATORS ALGORITHM (MPA)

In finding the prey of marine predators for simulating the best possible aging mechanism, marine predator's algorithm (MPA) has been proposed. If low concentration of prey presented in predators then it employs approach of Lévy and also Brownian movements is employed if abundant prey is presented. The trade off in between approaches of Lévy and Brownian is represented as the velocity ratio (v) in between prey and predators.

1. The most suitable approach of predators for moving towards Lévy steps is at low-velocity of v<0.1 in spite of whether the prey move towards Brownian or Lévy.
2. The predators move towards Brownian at a unit velocity of v=1 when the prey move towards Lévy steps.
3. Lastly, the approach of predators to stay motionless is at high-velocity >10 in spite of whether prey move towards Brownian or Lévy.

Following gives mathematical model for marine predators algorithm:

In the first step initializes a group of the prey with the help of equation $\vec{X} = \vec{X}min + rand(0,1) * (\vec{X}\max - \vec{X}min)$, in the search space. In that $rand\ (0,1)$ indicates a random number within a (0, 1) range, $\vec{X}min$ and $\vec{X}max$ are the vectors with the upper and lower bounds of each dimension in the optimization problem for the search space.

A best value of fitness is found and considered as a top predator by calculating the fitness value for each and every predator whenever the initialization phase of prey is completed. The best method in foraging is top predator according to its fitness survival, thus a matrix called $Elite$ it constructed by using it. The following represents the formulation of this elite matrix.

$$Elite = \begin{bmatrix} X^I_{1,1} & X^I_{1,2} & \dots & X^I_{1,d} \\ X^I_{2,1} & X^I_{2,2} & \dots & X^I_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ X^I_{n,1} & X^I_{n,2} & \dots & X^I_{n,d} \end{bmatrix}$$

Where $\overrightarrow{X^I}$represents the top predator vector which replicates in *n* times for building an Elite matrix of $n\times d$ order, n indicates the number of individual preys in the population and d indicates the number of dimensions.

In addition, for updating the positions of predators one more matrix called as *Prey* is used. This prey is also having the same dimensions as of *Elite* and thereby formulation of this prey matrix is given by:

$$Pray = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,d} \\ X_{2,1} & X_{2,2} & \dots & X_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & X_{1,1} & \dots & X_{n,d} \end{bmatrix}$$

According to the velocity ratio, the process of optimization is separated in a three phases in major iterations of the Marine Predators Algorithm.

3. SITUATIONAL INFORMATION IN SOCIAL MEDIA

Public most commonly use the platform of social media in various circumstances for sharing the related information. Hence, the users of social media generate well-off situational information in the crisis period like COVID-19 epidemic [5]. Nevertheless, in different types, the situational information is categorized by different researchers [9]. For example, Rudra in [9] defined the situational information as those notifications of the casualties or injured/stranded people or helping relief operations and categorized sympathizing with the victims, praising or criticizing the relief operation, post analysis of the reasons the crisis happens, and donation-related information into non situational information. The non situational information is categorized into situational information by Rudra in the study of Vieweg (2012). Situational information is particularly categorized into physical environment, built environment and social environment information's.

The posts that provide "tactical, actionable information that can aid people in making decisions, advise others on how to obtain specific information from various sources, or offer immediate post-impact help to those affected by the mass emergency" is the situational information defined by the Vieweg. For understanding and guiding the people in the crisis period, sharing of this information help to the related authorities and individuals [9]. Hence the situational information related to COVID-19, is categorized into seven types according to the above definition. Those content types are shown in figure (1) such types are: i) caution and advice; ii) notifications and measures been taken; iii) donations of money, goods, or services; iv) emotional support; v) help seeking; vi) doubt casting and criticizing; and vii) counter-rumour.

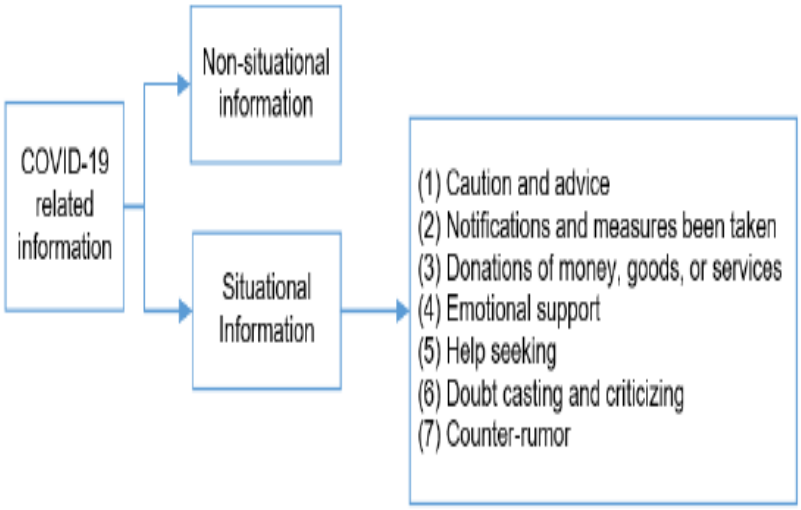

Fig. 1: INFORMATION TYPES RELATED TO COVID-19 ON SINA WEIBO

The information relating to concern and guidance information of the harm full virus inform to particularly help the people for protecting themselves. The inadequate information concern is made easy with the help of information from the notifications such as situational updates or casualties and damages related to COVID-19 that tells about these epidemic details to the people. The people which require help to know about what and which types of helps are available is given by the Donation information i.e. offering help. As result of this epidemic, a positive effect is shown on victims to recover from the emotional harm by using the Emotional support information and sharing that kind of information helps others to get a cooperative support and understanding. Help seeking posts are disclosed the information on immediate help or assistance in the crisis and this information sharing helps authorities and people get help or support. In addition, doubt casting and criticizing information frequently discuss about the sociopolitical reasons, suggestions and responsibilities for disaster, and sharing of this information will help others to improve their understanding of the information about this epidemic.

In addition, doubt casting and criticizing information frequently discuss about the socio-political reasons, suggestions and responsibilities for disaster, and sharing of this information will help others to vault the validity of information or improve their understanding about this epidemic. Counter-rumor information helps the people to understand the fact and reduces the confusions originated by rumors. Finally, these seven types of informations are categorized as situational information and thereby concluding that sharing of those information types gave benefit of efficiency in relief processes of crisis or disaster epidemic.

4. PROPOGATION OF SITUATIONAL INFORMATION AND DETECTION WITH RDR STRATEGY

A. Prediction of Propagation Scale for Situational Information

The five groups of features that are reviewed in the literature are extracted for finding key features which may be used to predict the circulation of each type of situational information precisely. Following are those features which are extracted.

1) Emotional factors: affect, negative emotion (negemo), positive emotions (posemo), anxiety (anx), anger and sadness words (sad) in the posts.

2) Perception-related factors: see, hear, percept and feel.

3) Affiliation-related factors: Affiliations, achieve, power, reward and risk.
4) User-related features: amount of followers (Followers (log)), amount of followees (Followees (log)), near the event or not (NearCity), exist in developed city or not (BigCity) and verified users or not.
5) Content-related factors: whether contained hashtag or URL (URL, Hash), the post length (*length*) and the publishing timing of the post (*hours*).

Specially, most widely used tool of extracting the linguistic information from content is the Linguistic Inquiry and Word Count (LIWC) that extracts the affiliation, perception and emotional factors. The linguistic patterns can be expose the perceptions, emotions and their necessity since the public use words and linguistic patterns in their text for achievement, power, affiliation and so on. In order to evade zeros, more quantity of groups are log-transformed as 1. From the data sets, it can be directly obtain the verified status of users. The definitions attained for Locations of Users are as follows:
1) The location of users near the disaster/crisis is defined as assign 1 to the variable if users located in province of "Hubei" and assign 0 otherwise.
2) The location of users near the developed region is defined as, assign 1 to the variables if users located in province of "Beijing, Shanghai, Shenzhen, and Guangzhou" or else assign 0.

In the content-related factors of URL or Hashtag, 1 is assigned when at least one URL / Hashtag is contained and while 0 is assigned in other condition. The post length is equal to total count of word in that post. The number of hours from December 30, 2019 to the post time of any information is defined as the post time. With the help of methods of linear regression, RF, and stepwise, fist the features for every situational information type are selected by use of above defined features. The method of stepwise approach and RF feature selection are chosen after comparing those selected features performances for information result of Type 1 to Type 6 and for Type 7 of counter-rumor information respectively and RF feature selection is chosen because of its high goodness of fitness (R2*)*. Also, the selected features of each situational information type use 85% of the data as training data which is randomly sampled by carrying out the regression model of multiple linear and negative binomials for predicting the log-transformed reposted amount of each type of information. Root mean square error (RMSE) of every prediction model is calculated. Therefore, from that calculation it was considered that linear regression model is used for predicting the COVID 19-related amount of reposted information since model of linear regression has less RMSE. Additionally, the effect of features selected on total seven situational information types is given in below. They particularly show the selected and potential features along with the effects of them for each type of situational information. For an instance, mostly six types of features called see, achieve, verified, followers (log), Big City, and hours are selected for Type 1 or caution and advice information.

1) The hash tag would be used to increase the reposted amount of the reposts of informations related to the Type 2 of notifications and measures to be taken, Type 3 of money donations, supplies or services, and Type 5 of seeking help. Generally publishing it was responsibility of related authorities for those information types. The use of hash tags is the best option when authorities wish to repost the information related to those types in vast quantity.
2) There would be a vast quantity of reposted amount of posts related to COVID-19 by the unverified users and with almost all types of this crisis-related information. Therefore, it was necessary to keep deep concentration by the authorities on the unverified users reliability for their huge influence. Ignoring the reliability of unverified users may leads to harm for related authorities.

3) The use of few negative words and more number of followers of users especially from the developed cities would be lead to have larger reposted amount for the Type 6 information of doubt casting and criticizing.
Hence, it might be necessary to keep concentration on these rational criticizers who have a large number of followers (also they might be the opinion leaders) by the related authorities and should verify their views as are valuable or not. The authorities give some advices to improve present crisis response techniques if there may be invaluable views. In future, during a crisis time the information of doubt casting type will become more popular comparing to other situational types of information.
4) The use of more words or containing more words in the content will ensure the larger reposted amount for all situational types of informations except the Type 7 of counter-rumor information. The propagation scale of all situational types of informations except Type 7 information can be increased if enlarging their length or else opposite techniques is used. It is necessary to increase the sample data if we want to verify that counter rumor information type is better or not to use lesser words in such a case of enlarging its propagation scale.
5) The reposted amount will be increased if the reposts related to Type 7 of counter-rumor informations are coming from developed areas and users having a larger number of followers. The authorities have to focus on the users who are having more number of followers through replying or stating them in order to increase the propagation scale of counter-rumor.

Finally, from outcome of this paper it can be said that the it gives about the understanding of people attitude towards the present crises reposes techniques of related authorities for identifying and increasing the notification, help seeking and donation informations which might be necessary to the public, and for identifying and countering the challenges of blames or rumors, for enhancing the crisis information publishing techniques of the authorities in the future.

B. Proposed Model

The image segmentation problem of multi thresholding is solved by developing the standard and improved MPA in this segment. The next segment shows the adopted steps.

i. Initialization

The prey number N and threshold number are predefined at this stage. After that, in between 0 to 255 gray levels of 8 bit image randomly initializes every threshold which is given by the equation (1).

$$P_{i,j} = L_{min} + r * (L_{max} - L_{min}) \text{ ----- (1)}$$

Where, L_{min} and L_{max} indicates a histogram image upper and lower bounds of the gray level values, and r signifies as a random number that randomly generated within the range of [0, 1].

ii. Ranking based Diversity Reduction Technique (RDR)

The number of iterations is terminated before reaching the best solution for a few numbers of particles that need a more time to find the number of iterations which are far away from the optimal solution. Thus, for every particle which cannot able to find a best solution the consecutive number of iterations is calculated by the proposed algorithm**.** This proposed algorithm is given as Algorithm 1 and the particles, within a subsequent number of iterations which are identified as a particle that cannot able to identify a best solution are updated in the algorithm for finding the better solution by minimizing the distance from the optimal solution with the help of Equation (2).

$$\overrightarrow{P_b} = \overrightarrow{P_b} + r * (\overrightarrow{P_b} - \overrightarrow{P_i}) \text{ ---- (2)}$$

Here, the particle which can't able to identify a better solution in a subsequent number of iterations is represented by $\overrightarrow{P_i}$ signifies as worst particle, $\overrightarrow{P_b}$ indicates a best solution in vector form and r represents the random number that is randomly generated within a [0, 1] range. Finally, the technique that minimizes the distance for an optimal solution of a worst particle in order to identify a better solution in a subsequent number of iterations is called the rank-based diversity reduction technique (RDR). The steps of this RDR algorithm are shown in Algorithm 1 below.

Algorithm 1: Ranking-based Diversity Reduction

```
P: number of prey
CR: a vector of size N and contain 0's value in the start
i = 0
perIter = 3
    while (i < N)
        if (fit (Pi) > fit Local (Pi))
            CRi++
        else
            CRi=0
        end if
      i++
      end while
    for each i particle
       if (CRi> perIter)
       Update Pi toward the best one using Eq. 3
       end if
    end for
```

C. Improved Marine Predators Algorithm (IMPA)

The segmentation problems of multi thresholding are solved by employing the IMPA along with RDR in steps by step. Prey numbers are initialized randomly in the first step called initializing step, since first step in all meta-heuristic algorithms consider as an initialization step. The Top Predator Best and Top Predator Position are defined as the value of highest fitness and its position respectively by calculating the fitness values for each and every prey within the step of initialization. Subsequently, at a rate of prey and current iteration current positions would be start updating by first primary optimization step with the help of any one of the updated multi thresholding equations. The memory saving in MPA is accomplished when the process of optimization at first stage is completed along with the completion of fitness value calculation for each prey. In the final step of optimization process i.e., the second stage methodology of FADs is implemented. For the organizing of local optima and thereby to find best solution, FADs helps to MPA. Lastly, the diversity among prey is reduced by calling the technique of RDR for a number of iterations after selecting the number of iterations.

The previous solution is replaced with present solution when this present solution is better than previous by the memory saving in the proposed marine predators algorithm or else, the previous one is used in the population to change direction and find better solutions. But if the previous solution may better than the predator should remain in its state i.e., motionless, and the distance through the better solution should not change. There is reduction of probability to find the better solution till in a condition that the particles are far from the better solution. Consequently, it would be neglected a significant number of iterations. Although the position updated is not better than the previous one, the Ranking-based Diversity Reduction is employed for solving this problem by slowly moving the particle in the subsequent number of

iterations on the best solution, that particle not able to find a better solution. Even though other regions may have a best solution, this helps the particle to explore. There is a decrease in the diversity among the population members as the particles move towards it since the best solution is uniform for all members. As a result, many better solutions can be created through exploring multiple regions by particles that have not been able to identify the best solution in the subsequent number of iterations.

5. RESULTS

The total Corona virus cases in India have crossed the 1.06 lakh while COVID-19 related deaths are over 3300 as of end of May, 2019. In India, there are 61,149 corona virus positive cases. So far 42,297 COVID 19 positive patients have recovered. As many as over 3300 people have died due to the deadly virus. The recovery rate in India is now over 38 per cent, as per details shared by the Ministry of Health and Family Welfare. Figure (2) shows the cases that are confirmed, recovered, deaths and active case in India from January to April of 2019.

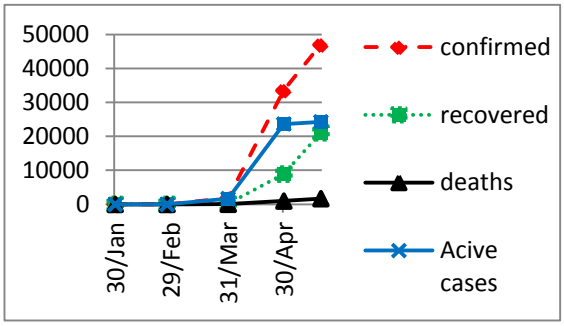

Fig. 2: COVID-19 CASES IN INDIA SHARED BY THE MINISTRY OF HEALTH AND FAMILY WELFARE

6. CONCLUSION

Today, most of the countries all over world are in the crises period of the COVID-19 epidemic. So for detecting this COVID-19 quickly on a vast number of tests a novel hybrid method called improved marine predators algorithm (IMPA) along with the ranking-based diversity reduction (RDR) approach is used in this paper. This model obtained the regions which may have effected with the COVID-19 by using x-ray images to extract similar small regions is called as image segmentation. This algorithm will also be utilized in color image segmentation for different medical applications in the future. In addition this paper found the need of using various approaches of COVID-19-related situational information publishing and selected the features for different types of situational information that helps to related authorities to understand and guide the people from this COVID-19

7. REFERENCES

[1] H. Liu, “Clinical and CT imaging features of the COVID-19 pneumonia: Focus on pregnant women and children”, Journal of Infection, 2020.
[2] C.S Guan, “Imaging Features of Coronavirus disease 2019 (COVID-19): Evaluation on Thin-Section CT”, Academic Radiology, 2020.

[3] G.M. Leung, J.T. Wu and K. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study," Lancet, volume- 395, number-10225, pp. 689-697, February, 2020.
[4] Martinez A.J. Pedraza and L. Yan, "Social media for disaster management: Operational value of the social conversation," Prod. Oper. Manage., volume-28, number-10, pp: 2514-2532, October, 2019.
[5] R. Beck and A. Mukkamala, "The role of social media for collective behaviour development in response to natural disasters", 26th Proceedings Eur. Conference on Information Systems (ECIS), 2018, pp: 1-18.
[6] J.C. Rubio Romero, M. Martínez Rojas and M.D.C. Pardo Ferreira, "Twitter as a tool for the management and analysis of emergency situations: A systematic literature review," Int. J. Inf. Manage., volume-43, pp: 196-208, December, 2018.
[7] J. A. Hendler, Q. Zhang, R. Breiger, F.Y. Wang, and D. Zeng, "Brokers or bridges? Exploring structural holes in a crowdsourcing system," IEEE Computers, volume- 49, noumber-6, PP: 56-64, June, 2016.
[8] J. Kuruvilla, "A review on image processing and image segmentation", IEEE, in 2016 international conference on data mining and advanced computing (SAPIENCE), 2016.
[9] Rudra.k, P. Goyal, N. Ganguly, and S. Ghosh "Extracting situational information from microblogs during disaster events: A classification-summarization approach", 24th Proceedings of ACM International Conference on Information Knowledge Management (CIKM), pp: 583-592, year: 2015.
[10] P. Burnap, "Tweeting the terror: Modelling the social media reaction to the woolwich terrorist attack," Social Network Analysis Mining, volume-4, number-1, PP: 1-14, December, 2014.
[11] P. Ghamisi, "Multilevel image segmentation based on fractional-order Darwinian particle swarm optimization", IEEE Tran. on Geo science and Remote sensing, PP: 2382 – 2394, **52**(5), 2013.
[12] H. Wang, Z. Ch, S. Jajodia and Gianvecchio . S, "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?" IEEE Transactions, Depe. Secure Computers, volume-9, number-6, PP: 811 – 824, November, 2012.
[13] K. Starbird, S. Vieweg, L. Palen and A. L. Hughes "Microblogging during two natural hazards events," in Proceedings of 28th International Conference Hum. Factors Computer Systems (CHI), PP: 1079, Year: 2010.
[14] Niels J. Noordhoek, Carsten Meye, Robert Manzke, Jochen Peters, Vivek Y. Reddy, Jürgen Weese, Olivier Ecabert, Raymond C Chan and Aravinda Thiagalingam, "Automatic Segmentation of Rotational X-Ray Images for Anatomic Intra-Procedural Surface Generation in Atrial Fibrillation Ablation Procedures", IEEE Transactions on Medical Imaging, Volume: 29, Issue: 2, Year: 2010.
[15] Francisco Serrano, Edward J. Delp, Kevin S. Lorenz and Paul Salama "Segmentation and registration based analysis of microscopy images", 2009 16th IEEE International Conference on Image Processing (ICIP),Year: 2009**.**