

Received September 6, 2021, accepted October 31, 2021, date of publication November 8, 2021, date of current version November 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125799

Fuzzy Consensus With Federated Learning Method in Medical Systems

DAWID POŁAP^{ID}, (Member, IEEE)

Faculty of Applied Mathematics, Silesian University of Technology, 44-100 Gliwice, Poland

e-mail: dawid.polap@polsl.pl

This work was supported by the Silesian University of Technology under Grant BKM-512/RMS2/2021 and Grant 09/020/RGJ21/0008.

ABSTRACT Large-scale group decision-making (LSGDM) is one of the main open problems where a decision is made by many different results. Moreover, there is also a problem with how to make the decision when there is no all information. This uncertainty can be very problematic for many different solutions in artificial intelligence. In this paper, we propose to extend a federated learning (FL) approach to not only a training process but also for making a decision using many different classifiers. This solution is applied in LSGDM, where many different results are intended for the classification of various data and can be used for deciding, even when some of the data are missing. For this purpose, we propose a fuzzy consensus that can be used in these problems. The contribution of this paper is the new way of using FL and extending its operation to many different classifiers. Our proposition was described for medical purposes and evaluated to show the advantages of the proposal. The proposal obtained 89,12% of accuracy on HAM10000, which is one of the best results compared to state-of-art.

INDEX TERMS Federated learning, LSGDM, fuzzy, neural network, healthcare industry 4.0.

I. INTRODUCTION

The fourth industrial revolution is a broad field that is designed to contribute to the betterment of our lives. A particularly important element is the automation of activities and operations in selected problems. It is based on modeling and implementing solutions in the field of automation, information collecting, data processing, and analysis. And here, artificial intelligence methods find their wide application through their information analysis capabilities. Recent years have shown that the fourth industrial revolution is slowly becoming more and more visible in the world around us, where we deal with products, systems and things with the prefix intelligent or smart [1], [2].

Increasingly wide automation, or even extending the functionality of various systems, contributes to the collection of a huge amount of data. Such data is initially processed to extract only meaningful information. However, there are situations when certain data may be used by many different systems. In practice, the opposite situation is also distinguished, i.e. when the products of many subsystems are used to obtain other data, i.e. indirect processing. It can be also equated with large-scale group decision-making (LSGDM), where many

persons or devices take part in making a decision. Unfortunately, in mentioned cases, there may be situations where the data is incomplete or uncertain. In the case of using artificial intelligence methods, the lack of specific information may result in incorrect analysis. Hence, a very important element is the analysis or control of not only data but also decisions made.

In this paper, we propose a fuzzy consensus in LSGDM using the FL idea for training different artificial intelligence solutions. FL is used for training solutions, but we re-modeled this solution for making two things – training a general model for the same classifiers and for making a decision using any information. In the case of using many different data for making a final decision, it is done by a fuzzy controller which can decide with some uncertainty. Such an approach allows not only to omit a given classifier during classification but also to make decisions based on the lack of all relevant information. Our main contribution are:

- a model of using artificial intelligence methods in practical application for medical purposes, that allows creating a machine model solution and applying it in practice without any previous preparation. The proposed operation model is based on the use of various simple classifiers that can precede the initial process of implementing

The associate editor coordinating the review of this manuscript and approving it for publication was Sunil Karamchandani^{ID}.

and training deep networks, such as convolutional neural networks.

- a fuzzy consensus method in LSGDM, through which the possibility of training smaller classifiers, as well as the division of classification tasks into different centers, increases the preservation of data privacy and the possibility of faster implementation of deep classifiers. The application of the fuzzy approach contributes to the reduction of the number of operations and great freedom in extending the models to further classification tasks.
- improving the FL approach by extending its functions to making a decision and providing the policy of using additional simple classifiers for the time of increasing the accuracy of deep classifiers.

The proposed model has main two purposes that met the current assumptions and requirements of the technology market. The first one is not creating intelligent solutions for a specific database, but enabling the use of artificial intelligence methods, in particular an artificial neural network, implementation to the problem. This approach makes the solution much more flexible when it comes to practical implementation. It is visible when the data can be modified during the implementation or the use of a given model. The second important feature is the possibility of implementing the software despite the lack of a large amount of data and the possibility of automatic change of classifiers depending on the obtained accuracy. Our solution has been subjected to classification tests for various parameters under the influence of accuracy and time. This allows confronting the results with other existing solutions to demonstrate the superiority of the proposed method.

II. RELATED WORKS

The last years brought many areas, where different artificial intelligence techniques can be used and automatize some process [3], [4]. The application and development of these solutions brought new challenges, especially in terms of uncertainty analysis, whether decision-making is based on a lot of information or even indirect decisions. In [5], the authors analyze the problem of decision-making with applications in medical problems and healthcare area. They focused on the use of fuzzy logic for analyzed problems and show that it has already great achievements. Again in [6], the non-cooperative behavior was analyzed for making a decision. The research shows that linguistic preference ordering can be very helpful in analyzing the data. Having many different results and take them into consideration in a process of making a decision is a very difficult task in computer science. One of such solutions for it is the use of clustering algorithm like k-means and based on the probability make a decision [7]. Another proposition was proposed in [8], where the scientist described making a decision process as an optimization task using the connections and relationships between experts. Similar idea with considering relationship was shown in [9], the theoretical model of analyzing experts was proposed. The authors described the creation of a similarity matrix and analyzing it. Again in [10], the consensus

was reached by merging experts into groups according to their opinions. This approach gives better scalability and manages polarized conflicting opinions. One of the latest research that uses fuzzy logic in decision-making tasks was shown in [11], where the authors define a new score function and show the practical advantages of the proposed model.

Artificial intelligence methods are solutions that are used quite often in all applications. This is especially visible in the analysis of the collected data, which are then presented to experts or doctors. They can also be used for classifying and decision making. One of the most universal tools is artificial neural networks and their derivatives. However, they have one disadvantage, which is the very long training process. In order to minimize the training process and possible implementation of neural networks, collaborative training called federated learning (FL) was proposed [12], [13]. It is based on using many workers for training on smaller, private databases and creating one aggregated model. using many workers has a problem with communication cost between them and a server. In [14], the authors proposed using a heuristic algorithm for optimization task which is communication cost. Another issue which is improved in it is the privacy of models which can be improved and make the data, and models much safer is blockchain [15], [16]. It is also used in edge-cloud cooperation environment [17], and this proves that federated learning can be used not only in different solutions but with different software. In [18], the use of three methods like neural network, heuristic and fuzzy logic was marge for damage detection. Another approach was combining neural network with decision tree for decision making in breast cancers [19].

All of the mentioned solutions, methods are mainly used in medical systems that can speed up the exchange of information, perform faster analysis, and highlight certain data. In [20], the authors discuss the possibilities of fog computation for medical purposes. This kind of research is very important, because of two reasons. The calculation of used tools must be done somewhere and in many cases, this task needs a lot of computation power. Many different solutions are investigated in that terms of the use of cloud and fog computing. These solutions offer the possibility of using large, external computing systems as well as data storage. The practical aspects are visible in our lives, where we use different sensors (like smartwatches, bands, etc) for analyzing our heartbeat, steps, etc. All of this data might be helpful for better analysis of our health and condition because of constant measurements what was analyze din [21]. The presented research shows that different sensors are very valuable in constantly monitoring our health, especially with the increasing possibilities and implementation of the Internet of Things concept. A similar topic was shown in [22], where the possible application of machine learning algorithms was presented. Again in [23], the aspect of machine learning and big data was considered. The enormous amount of data needed to train many different methods of AI has been highlighted. Not only laboratory experiments are performed, but some practical analysis and implementation are made [24].

In that paper, the authors have shown a framework that can be applied in elderly patients' healthcare.

III. FEDERATE LEARNING METHOD IN LSGDM

In this section we define our proposition of medical system. We focus on defining federated learning approach with our modification for LSGDM. This is a continuation of our research on the use of intelligent solutions in the Internet of Medical Things (IoMT) [25], [26].

A. ARCHITECTURE MODEL

The main idea is based on creating a complex service between many different hospitals, clinics, laboratories, etc. A patient may have a doctor in hospital A, but do some examination on laboratory 1 and 2. In the meantime, there will be another health problem that will be treated at Hospital B. The doctor at Hospital B will perform some examinations. After receiving the results, he is not sure if the new information is related to some other diseases or complications treated by other specialists. We assume, that a doctor can send a query to the server which makes calculations and makes a decision based on all available information. A simple visualization of this is presented in Fig. 1.

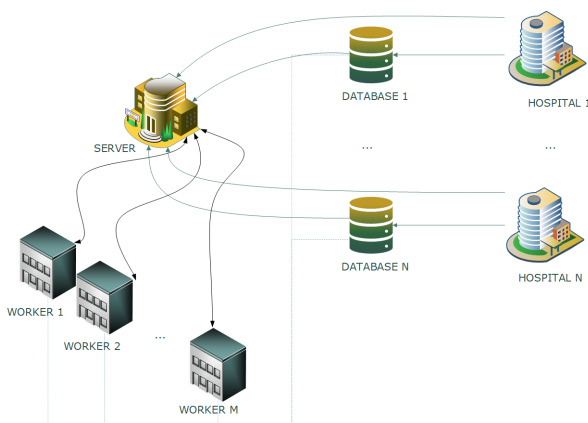


FIGURE 1. A visualization of the general idea. The results from hospitals are collected in databases. Each of the hospitals can request a decision that is made on the server. The server is responsible for training individual classifiers depending on the variety of data provided by hospitals. Depending on the number of workers, the server carries out training using federated learning.

The above problem formulation can be considered from a practical point of view, i.e. how exactly individual objects, such as workers, server, or even medical objects works. Suppose we have N of various medical facilities such as hospitals, clinics, and laboratories. Each of these facilities has its own private database where some information are stored like orders, research results, as well as encrypted confidential data.

After the medical examination is performed, the data is saved in a database. Then the doctor analyzes them. However, a very important element is the additional auxiliary analysis performed by artificial intelligence in the form of a decision support system. In fact, such a system performs its own

analysis when a sample is added to the system. The result of the automatic analysis is entered in the database, which the doctor can see and take into account during his own analysis. If his decision is the same as the system, it should be approved in the database. Otherwise, it is changed. Such action fills the database with subsequently approved records that can be used in the process of training artificial intelligence methods that perform this analysis. As a result of such activities, the accuracy of the methods is increased. This description shows the action of a specific classification, where one classifier is connected to the medical facilities. However, there may be a situation, when many different examinations may be symptoms for other diseases. Here it is necessary to consider the problem of LSGDM, where multiple results will determine the diagnosis. The doctor can send a request to the server to analyze all medical data for a given patient, including those in other databases. The server makes a diagnosis based on all the classifier's results.

B. SERVER'S OPERATION

The server in our proposal has two main tasks - performing the training process using the federated learning approach with the classification of a given sample, as well as making a decision based on the current state of knowledge in all databases. A pseudocode of the server's operation is presented in Alg. 1.

1) FEDERATED LEARNING AND THE ABILITY TO CLASSIFY A SINGLE SAMPLE

The basic operation of the server is to analyze the number of databases, problems as well as the number of workers. The worker is an external server or cloud computing that enables training of the selected classifier. Let us assume that we have N databases marked as D_s where $s \in \{0, 1, \dots, N - 1\}$ and m workers. Having N different databases, we assume that we need N trained classifiers. So, for each database, one classifier must be prepared. Initially, the number of workers that can be assigned to individual databases is calculated. If $m < 2 \dots N$, this means that a cascade training is performed, i.e. each database is trained consecutively. In the opposite case, the number of workers for a single database is determined as $\lfloor m/N \rfloor$. The rest of them are assigned randomly. Such action ensures that the federated learning approach makes sense, as there is training on several workers who can create one common model.

Then, each database k is split on equal parts on all workers which were assigned to this database. Then a basic model is generated as random and marked as θ_k^0 and upload to all workers. After a workers upload the trained models θ_i , a server aggregate all of them and send to all workers for another round. It is made by T_{fl} iterations. The aggregation can be formulated as follows:

$$\theta^{t+1} = \theta^t - \sum_{i=0}^{k-1} \frac{|D_i|}{|D_s|} \theta_i^t, \quad (1)$$

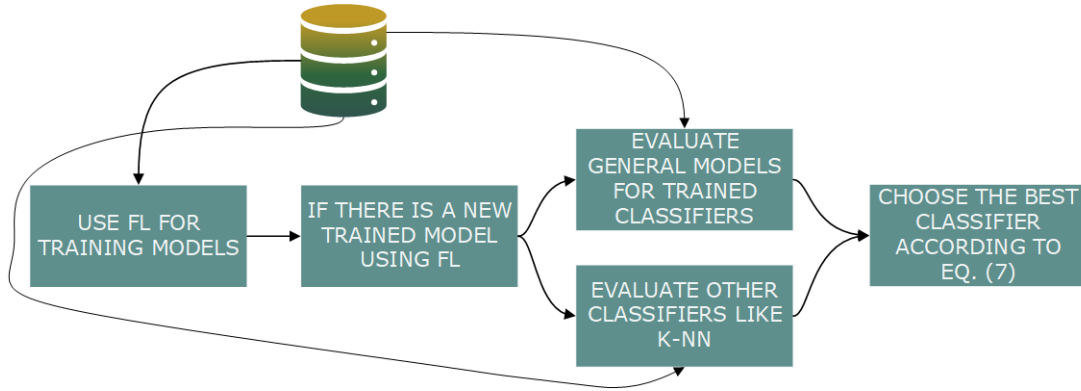


FIGURE 2. Visualization of the classifier selection carried out after each generation of a new model using FL.

where t means the current iteration, $|D_s|$ is the number of all samples in database D_s , and k is the number of workers assigned to this database. The whole described idea of FL can be described by the following equation:

$$\min_{\theta} \hat{L}(\theta) = \sum_{i=0}^{k-1} \frac{|D_i|}{k} L_i(\theta), \quad (2)$$

where $L_i(\cdot)$ is a loss function defined as:

$$L_i(\theta) = \frac{1}{|D_s|} \sum_{j \in \xi_i} l_j(\theta), \quad (3)$$

where $l_j(\cdot)$ is a square error function for a given sample $\xi_i \in D_s$.

In this way, the server is responsible for training classifiers for a given databases connected to individual medical facilities. When a specific facility sends a sample with a request for classification, the server retrieves the currently trained model for the given database and uses it to classify. The result is sent to the doctor who can approve or reject the result. The rejection will most often take place when the classifier will not be sufficiently trained (its error will be large and the accuracy will be low). Such a situation will most often occur in the initial phases of the system implementation, or when the databases will have a small number of records that can be used in the training process.

We also propose applying a simple classifier to be able to implement the system even when the databases are not filled with large amounts of data. By simple, we mean the solution which does not require a training process and uses some metric to analyze the sample. An example of such an algorithm is k -nearest neighborhood (k -nn) [27], which computes the distance between two samples using some metric like Minkowski one marked as $\rho(\cdot)$. This algorithm calculates the probability estimator, posteriori belonging of observations x to k class is calculated as:

$$\hat{p}(k|x) = \frac{1}{K} \sum_{i=1}^n I(\rho(x, x_i) \leq \rho(x, x^{(k)})) I(y_i = k), \quad k = 1, \dots, L, \quad (4)$$

where $x^{(k)}$ is k -th sample nearest samples from x using Minkowski metric. Then, the classification process is defined as:

$$\hat{d}_{KNN}(x) = \arg \max_k \hat{p}(k|x). \quad (5)$$

This simple algorithm can be used for numerical data and images. However, an important aspect is having a mechanism that allows to automatically disable the classification option using k -nn, and switch to deep classifiers such as neural networks trained by FL. It can be implemented as the analyzes of accuracy. If the trained model in the FL process will obtain better accuracy than other applied classifiers, then a system uses the currently trained model and updates it when it is such an opportunity. Accuracy is calculated as the ratio of correctly classified samples for the validation database D_{val} using the training database D_{train} . It is done by taking all labeled samples and their random division into these two sets in the ratio of 70% to 30% of all samples ($D_{val} \cup D_{train} = D$). Then the accuracy is calculated using the following formula:

$$acc = \frac{corrected}{|D_{val}|}, \quad (6)$$

where *corrected* means the number of corrected classified samples from D_{val} . In the case of k -nn, the k values are selected by the maximum value of accuracy for a set of k ($k \in \{2, 10\}$). The rejection of a simple classifier in favor of a deep neural network is performed when the general model on the server finished the first FL iteration. Then, the accuracy of the model θ can be determined. After that, the following condition is analyzed to choose the further procedure concerning the applied method (it is repeated after each FL iteration, see Fig. 2):

$$\begin{cases} acc(\theta) < acc \left(\max_K \hat{d}_{KNN} \right) & \text{use } k\text{-nn,} \\ acc(\theta) \geq acc \left(\max_K \hat{d}_{KNN} \right) & \text{use trained model.} \end{cases} \quad (7)$$

Algorithm 1: Server Operation

```

Input: The number of federated learning iterations  $T_{fl}$ ,
all  $N$  databases, the number of workers  $M$ 
1  $k = 0$ ;
2 for each database  $k$  do
3   Create a random general model  $\theta_k$ ;
4    $k++$ ;
5 Assign workers to databases and split database;
6  $k = 0$ ;
7 for  $k < T_{fl}$  do
8   for each database  $k$  do
9     Send a model  $\theta_k$  to workers;
10    for each worker  $i$  do
11      Wait for models from worker  $i$ ;
12      Calculate loss value using Eq. (3);
13    Calculate loss value using Eq. (2);
14    Aggregate all models;
15    Calculate accuracy of new model;
16    Analyze and select the best classifier currently
    based on accuracy using Eq. (7);
17    $k++$ ;
18 if there is a new sample then
19   Check the database;
20   Get the current model for selected database;
21   Calculate results;
22   Send classification result;
23 if there is a request for patient's analysis then
24   Get classification result for patient from all
    databases;
25   Using Takagi-Sugeno model make a decision;
26   Send a decision;
    
```

2) DECISION-MAKING USING FUZZY CONTROLLER

It is important to analyze all medical data from different medical facilities. We assumed that we have N facilities. Each of them can have examination results which could be important to make a final decision. Using machine learning solutions, the classification results mainly is in the range of $(0, 1)$ which indicates the probability of belonging to this class. The best situation is when the probability is equal to 1. It means a 100% of certain. Unfortunately, this situation is hard to get. Moreover, in the case when we have many different classes, it can be difficult to identify one class and not much with less probability. Also, it should be taken into account that certain results are not present, which may be caused by database corruption, data loss, or failure to perform the examination.

Based on this observation, we propose a second operation which a server will perform. Having N results, a final decision should be made, even when some data are not there. For this reason we can define a fuzzy control system which will make the final decision. We propose the use of Takagi-Sugeno

system which is based on rules formulated as:

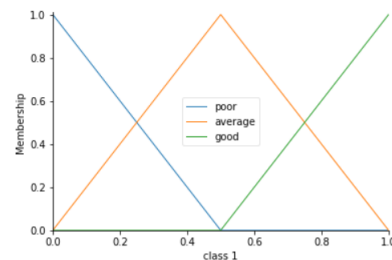
$$\mathcal{R} = \left\{ \mathcal{R}^{(i)} \right\}_{i=1}^l = \left\{ \text{if } \bigwedge_{n=1}^N x_{0n} \text{ is } A_n^{(i)}, \text{ then } y = f_i(x_0) \right\}_{i=1}^l, \quad (8)$$

where i is the number of all rules, x_{0n} is incoming value, $A_n^{(i)}$ is linguistic value, $y = f(x_0)$ is a conclusion of i -th rule.

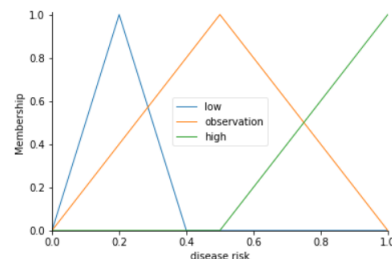
We assume that each classifier will give a value of belonging to a given class (marked as p_i where $i \in \{0, 1, \dots, N-1\}$ and hereinafter understood as the argument x for the given functions in Takagi-Sugeno system) described in the range $(0, 1)$. This value can be visualized as a fuzzy variables, which membership will be understood as poor/average/good and defined by the use of triangular function:

$$\mu(x; a, b, c) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x \leq b, \\ \frac{c-x}{c-b}, & b < x \leq c, \\ 0, & x > c, \end{cases} \quad (9)$$

where a, b, c are a triangular parameter which met the following condition ($a \leq b \leq c$). The mentioned values as poor/average/good will define the value, as can be seen in Fig. 3.



(a) Antecedent



(b) Consequent

FIGURE 3. Visualization of fuzzy sets and linguistic variables on the antecedent and consequent.

Our system will be deciding whether the patient has a low or high risk of disease. Additionally, we distinguish the possibility of uncertainty, which we mark as observation (see Fig. 3b). According to Eq. (8), the rules will be defined as:

- if p_0 is poor and p_1 is poor and \dots and p_{N-1} is poor then risk is low,
- if p_l is good then risk is high,

- if (p_0 is poor or p_0 is average) and (p_1 is poor or p_1 is average) and ... and (p_{N-1} is poor or p_{N-1} is average) then risk is observation.

Inference is done by calculating the degree of belonging of individual rules as:

$$F^{(i)}(x_0) = \mu_{A^{(i)}}(x_{01}) \star_T \mu_{A_2^{(i)}}(x_{02}) \star_T \dots \star_T \mu_{A_N^{(i)}}(x_{0N}), \tag{10}$$

where \star_T means t -norm. And then the final result is calculated by sharpening operation using the following equation:

$$y_0 = \frac{\sum_{i=1}^I F^{(i)}(x_0), \quad y^{(i)}(x_0)}{\sum_{j=1}^I F^{(j)}(x_0)} \tag{11}$$

where $y^{(i)}(x_0) = f_i(x_0)$ is a conclusion of i -th rule, $F^{(i)}(x_0)$ is a membership degree of i -th rule.

Algorithm 2: Worker Operation

Input: An access to samples in database D_I , training iteration T_{train}

- 1 Download a model from the server;
- 2 Set the downloaded model in classifier;
- 3 $o := 0$;
- 4 **for** $o < T_{train}$ **do**
- 5 Train model using D_i ;
- 6 $o ++$;
- 7 **end**
- 8 Upload trained model to server;

C. WORKER'S OPERATION

A worker is a server on which calculations are performed as part of training a given classifier. Its operation begins after obtaining access to samples in a given database D_I and receiving a basic model θ_l^0 . Then, a worker performs a given number of training iterations T_{train} . After reaching the maximum number of iterations, it sends a trained model to server for aggregation. This operations is repeated until the stop condition is not met, for instance reaching the maximum number of FL iterations T_{fl} . A pseudocode of worker's operation is shown in Alg. 2.

IV. EXPERIMENTS

To evaluate our proposition, we used a classic database of skin marks images named MNIST: HAM10000 [28]. This set of images contains 10015 samples which were categorized into seven different categories as dermatofibroma, vascular lesion, actinic keratosis, basal cell carcinoma, benign keratosis, melanoma, and nevi. Experiments were conducted in the form of simulation and analysis of different settings. We simulated $N = 7$ different medical facilities with databases

that contained two classes: diseases and not a disease, for example, dermatofibroma and others, vascular and others, etc. Databases were filled with all samples from this category and 2000 other images. We used $M \in \{1, 2, \dots, 14\}$ workers, federated learning iterations as $T_{fl} \in \{10, 15, 20\}$ with training iterations $T_{train} \in \{25, 50, 100\}$. As a classifier we used learning transfer models like VGG [29], and Inception [30] trained by ADAM algorithm [31].

At first, we analyzed the accuracy level and time of training such architectures using mentioned above parameters. The obtained results of accuracy are presented in Fig. 4, 6, and 8, and the measured time is plotted in Fig. 5, 7, and 9. In Fig. 4, the accuracy level is shown using only 10 rounds of FL and one worker. This parameter setting indicates no additional computing power. In our simulation, the worker was not connected to the server, but this is also a case of their connection, a classic solution without the

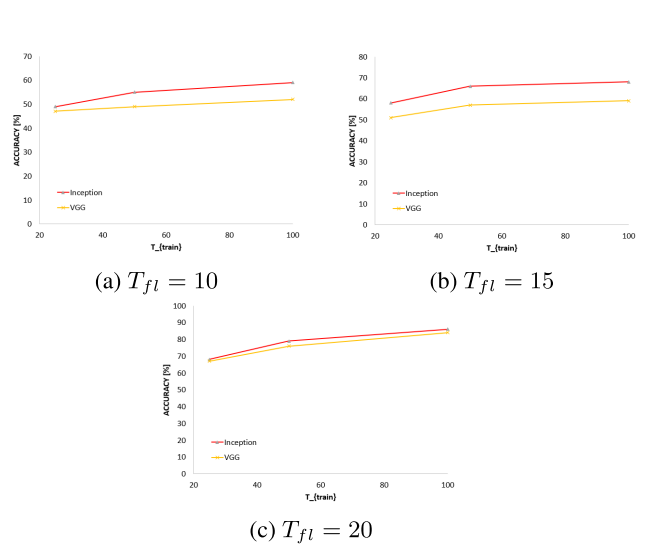


FIGURE 4. Accuracy vs. the number of training iterations T_{train} depending on the number of federated learning rounds T_{fl} using 1 worker.

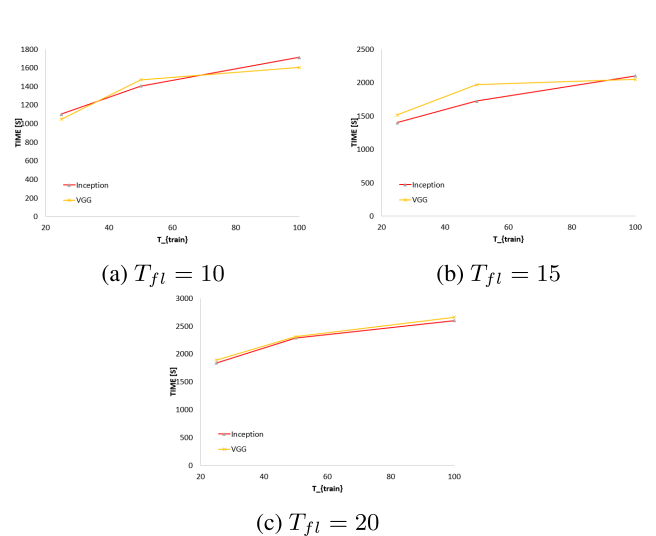


FIGURE 5. Time vs. the number of training iterations T_{train} depending on the number of federated learning rounds T_{fl} using 1 worker.

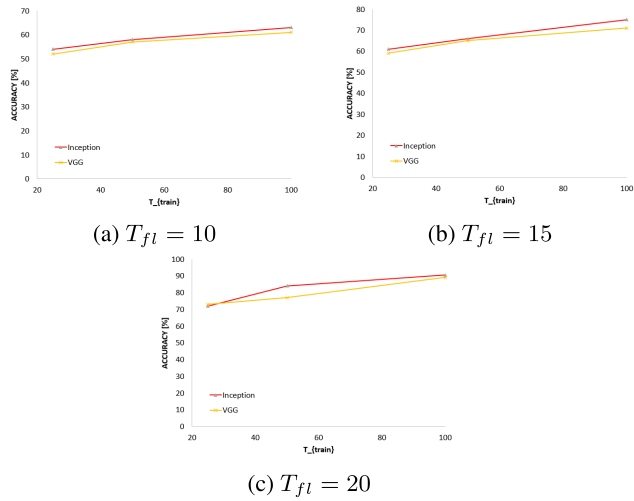


FIGURE 6. Accuracy vs. the number of training iterations T_{train} depending on the number of federated learning rounds T_{fl} using 7 workers.

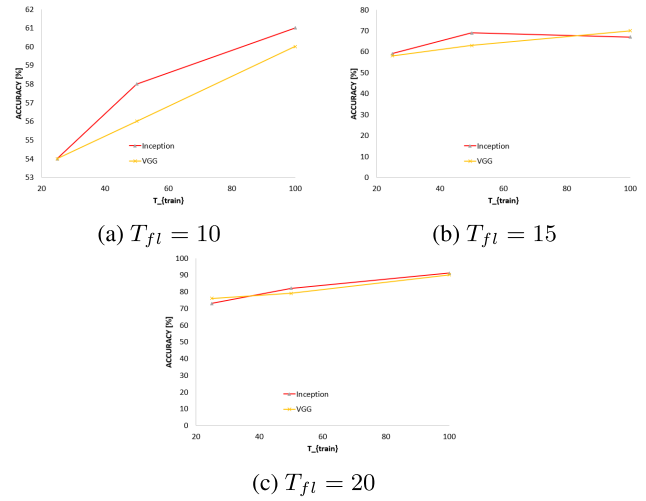


FIGURE 8. Accuracy vs. the number of training iterations T_{train} depending on the number of federated learning rounds T_{fl} using 14 workers.

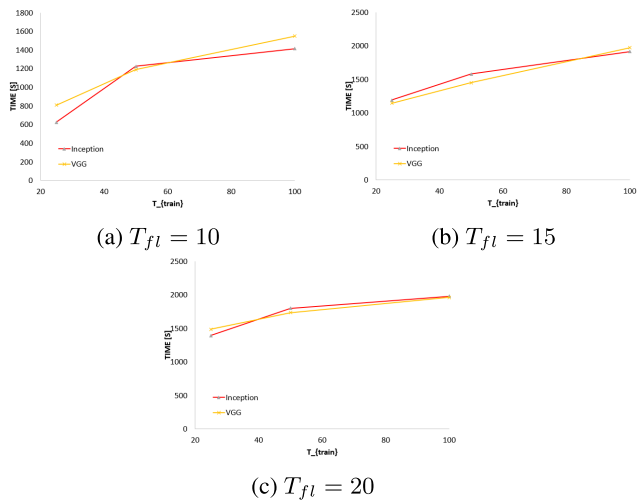


FIGURE 7. Time vs. the number of training iterations T_{train} depending on the number of federated learning rounds T_{fl} using 7 workers.

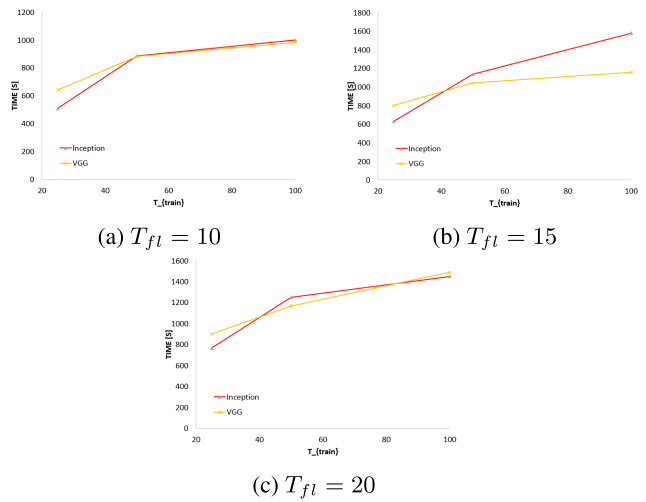
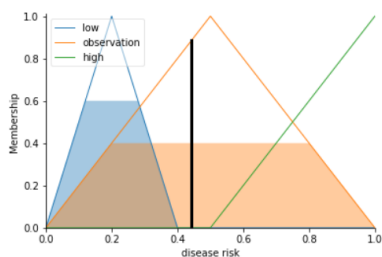


FIGURE 9. Time vs. the number of training iterations T_{train} depending on the number of federated learning rounds T_{fl} using 14 workers.

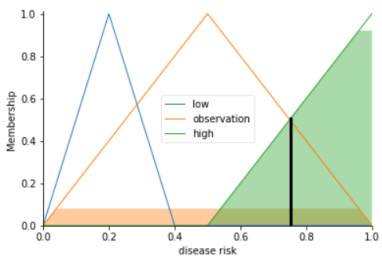
use of federated learning. In this situation, there is a single classifier for each database, so there is no aggregation of models, as there is only one model. Moreover, it should be noted that there is no parallelism, so each classifier is trained after the previous training is completed. The accuracy level started almost at the same point, but Inception architecture reached better results than VGG (see Fig. 4a). In the case of time, both architectures reached a similar time without any difference in the number of T_{train} (Fig. 5a). Accuracy level is on a similar level using both architectures with 1 worker and 15-20 FL rounds, the only difference is visible in Fig. 5b. The small number of T_{train} like 25 or 50 indicates longer training time by nearly 200 seconds. For $T_{train} = 50$, the difference is between using the VGG architecture compared to Inception is 15% time savings. In other cases, there is no difference in time or accuracy measurements. It should be noted, that using only 1 worker allows achieving

maximum accuracy of 86% with 20 federated learning rounds and $T_{train} = 100$.

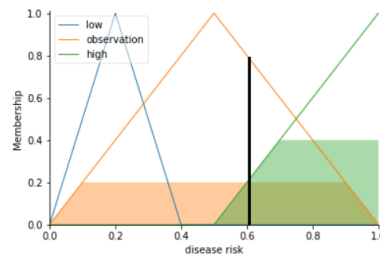
In the next step of the conducted research, we increased the number of workers to 7. In that situation, each worker should be assigned to one database and the results would be similar to the previous tests, and the application of FL still would not make sense. The main difference would be the training time which should be reduced because of parallel calculation on seven different workers. We changed the experimental conditions to merge 3, 2, 2 workers into groups. In this case, when one of the groups would finish the calculation, it is used for another database. And for this settings, the obtained results are shown in Fig. 6 and 7. In the case of measurement of accuracy level, the results were similar by both learning transfer models. The best results were reached by Inception on 88,72%, a VGG with the same settings ($T_{fl} = 20$ and $T_{train} = 100$) reached 88,88%. Time needed to train such



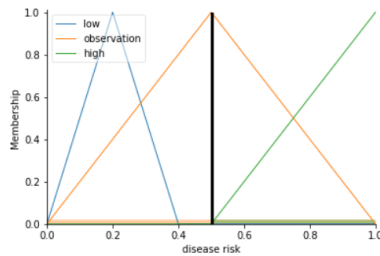
(a) All values were smaller or equal to 0.3.



(c) Two of seven values were higher than 0.9, the rest below 0.2.



(b) One value was at level 0.7, the rest below 0.4.



(d) Sample results using fuzzy consensus for all classifiers.

FIGURE 10. Sample results using fuzzy consensus for all classifiers.

models was smaller than using 1 worker by average value of 24% (using VGG or Inception model) (see Fig. 5 and 7). For all cases, the VGG model shows much more stable results in terms of time as well as accuracy. It is visible by the linearity of the charts of VGG concerning the second learning transfer model.

Increasing the number of workers to 14 means that for each database there will be two workers. For such tests, the results are shown in Fig. 8 and 9. The accuracy chart shows similar dependencies as during previous tests. The plots are very close to each other, and the Inception model is slightly better for almost all conditions in conducted experiments. The best results were achieved, as before, for the maximum number of iterations and rounds, which resulted in an accuracy of 89,12%. The use of 14 workers in the FL approach reduced the training time by nearly 40% to the use of 7 workers combined into groups and by 56% compared to the classic cascade approach. The increasing number of workers shows potential mainly in the reducing time of training process. In the case of accuracy level, the obtained values were also different, but mainly in the case of use and not use FL approach. In the case of a classic approach, the best result was 86%, and in the case of second and third experiments, where more than one worker calculated the same model was reached at a minimal value of 88%. In Tab. 1, we compare our reached accuracy of 89,12% to other solutions in the literature. The most common tool for this kind of classification problem was CNNs like [35]. We reached better results than the deep learning solution by over 0.5%. It is a good result taking into consideration that this difference is at the level of almost 90% of a given method. In [26], [34], the learning transfer was used, and proposed CNN was trained for all classes in one instance of classifier. Again in [38] used learning transfer but with masked recurrent CNN for not only the classification purpose

TABLE 1. Comparison with state-of-art.

Method	Accuracy [%]
CNN with learning transfer [26]	81,4
MobileNet [32]	83,93
CNN with softmax [33]	86,5
DenseNet-161 [34]	86,6
Cascade knowledge diffusion network [35]	88,0
AutoML method [36]	88,0
Deep learning [37]	88,2
ELM [38]	88,5
Deep learning [39]	88,57
Proposed method	89,12

but also for detection of skin marks. The authors achieved the accuracy on the level of 88,5%. Such an approach shows a great potential application because of its high accuracy. It is worth notice, that we used learning transfer solutions without freezing any layers and trained it for classification only two classes for obtaining many classifiers that focuses on a smaller number of classes. It shows that it is a better solution than using one classifier and train it for all classes at once.

After analyzing the application of FL for medical purposes, we tested also the proposed fuzzy consensus described in Sec. III-B2. We used all classification results from each database which led to gain 7 independent probability of belonging to these categories. Such values were entered into the Takagi-Sugeno system to perform fuzzy inference and decide on the risk degree of a given disease. Three decisions were considered – low, high risk of the disease and observation. Observation means that there is no clear possibility of making a risk decision and the patient should observe the skin marks. For all data in the databases, the results were checked and example results are shown in Fig. 10. The presented results allow noticing that in the case of a high classification

result (above 60%), the system does not decide on the disease, and proposes an observation. Again, in the case of values higher than 0.85, the returned decision is high risk. This Takagi-Sugeno system does not require a lot of computing power due to simple calculations. Moreover, the lack of any decision from a particular database will only result in the lack of its analysis. Although it will not significantly affect the results. If different data would be used, such a decision will be made based on model rules that can be targeted at particular information/decisions, which could be crucial for making a decision. However, in cases such as automatic classification using a fuzzy controller, especially taking into account many different classifiers, allows for quick decision making. In practical use, such a solution has a great advantage due to the possibility of quick implementation and classification of the sample even during the training process (FL advantage). The quick multi-decision analysis module also allows to quickly make decisions and even modify rules while the system is running.

V. CONCLUSION

In this paper, we proposed a solution for medical purposes that can be treated as a medical expert system. Each medical facility has its database, where the records, examination results are kept. We added a general server for managing classifiers based on artificial intelligence methods. A doctor from a facility can add new data to the database and request a fast classification. A server can do it by neural architecture. Moreover, a doctor can request for deciding the disease risk which is made based on the patient's data from all available databases. The main advantages of this proposal are the use of federated learning and fuzzy consensus. Federated learning allows for training intelligence methods much faster and during the training, the process makes a classification task. A fuzzy consensus allows for making a decision based on all data. For this purpose, we used the Takagi-Sugeno controller and modeled rules.

For analyzing the proposal, we used a database with skin marks images called MNIST:HAM10000. During conducting experiments, we reached an accuracy level of 89,12%. This result shows the huge potential of our solution to other existing methods, which results in obtaining a better average accuracy when using many smaller classifiers than one by almost 0.5%. A particularly important element is the applied fuzzy controller, thanks to which the results from many databases can be used in making an overall decision. Our tests showed the correct analysis of the results based on applied data. However, it should be mentioned here that in practice there will be different data and format, not only focused on one health problem. To adapt our proposal to other data, it is enough to remodel the rules in the fuzzy system and possibly implement additional classification techniques like classic artificial neural networks. The proposed model of tasks division into many centers showed enormous possibilities in terms of reducing training time. Also, our proposal enables the implementation of the solution even at

the initial stage of systems and databases. Depending on the accuracy of the classifiers for the current state of the database, a classifier is selected. A reasonable replacement of simple classifiers with deep architectures in case of exceeding a certain accuracy condition between them makes our model of operation possible for practical applications already at the initial stages of operation. However, the best results are obtained after selecting deep architectures.

REFERENCES

- [1] T. Hirasawa, M. Yoshida, and S. Obara, "Battery control for leveling the amount of electricity purchase in smart-energy houses," *Int. J. Energy Res.*, vol. 45, no. 1, pp. 807–823, Jan. 2021.
- [2] Z. Lv, D. Chen, R. Lou, and Q. Wang, "Intelligent edge computing based on machine learning for smart city," *Future Gener. Comput. Syst.*, vol. 115, pp. 90–99, Feb. 2021.
- [3] D. Dąbrowska, R. Kucharski, and A. J. Witkowski, "The representativity index of a simple monitoring network with regular theoretical shapes and its practical application for the existing groundwater monitoring network of the Tychy-Urbanowice landfills, Poland," *Environ. Earth Sci.*, vol. 75, no. 9, p. 749, May 2016.
- [4] T. Hyla and N. Wawrzyniak, "Identification of vessels on inland waters using low-quality video streams," in *Proc. 54th Hawaii Int. Conf. Syst. Sci.*, 2021, p. 7269.
- [5] A. Mardani, R. E. Hooker, S. Ozkul, S. Yifan, M. Nilashi, H. Z. Sabzi, and G. C. Fei, "Application of decision making and fuzzy sets theory to evaluate the healthcare and medical problems: A review of three decades of research with recent developments," *Expert Syst. Appl.*, vol. 137, pp. 202–231, Dec. 2019.
- [6] X. Gou and Z. Xu, "Managing noncooperative behaviors in large-scale group decision-making with linguistic preference orderings: The application in internet venture capital," *Inf. Fusion*, vol. 69, pp. 142–155, May 2021.
- [7] Q. Liu, H. Wu, and Z. Xu, "Consensus model based on probability k-means clustering algorithm for large scale group decision making," *Int. J. Mach. Learn. Cybern.*, vol. 12, pp. 1–18, Jan. 2021.
- [8] Y. Lu, Y. Xu, E. Herrera-Viedma, and Y. Han, "Consensus of large-scale group decision making in social network: The minimum cost model based on robust optimization," *Inf. Sci.*, vol. 547, pp. 910–930, Feb. 2021.
- [9] Z. Liu, X. He, and Y. Deng, "Network-based evidential three-way theoretic model for large-scale group decision analysis," *Inf. Sci.*, vol. 547, pp. 689–709, Feb. 2021.
- [10] R. M. Rodríguez, Á. Labella, M. Sesma-Sara, H. Bustince, and L. Martínez, "A cohesion-driven consensus reaching process for large scale group decision making under a hesitant fuzzy linguistic term sets environment," *Comput. Ind. Eng.*, vol. 155, May 2021, Art. no. 107158.
- [11] S. Narayanamoorthy, L. Ramya, and D. Kang, "Normal wiggly hesitant fuzzy set with multi-criteria decision making problem," *AIP Conf.*, vol. 2261, no. 1, 2020, Art. no. 030023.
- [12] A. T. Thorgeirsson and F. Gauterin, "Probabilistic predictions with federated learning," *Entropy*, vol. 23, no. 1, p. 41, Dec. 2020.
- [13] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Federated learning for machinery fault diagnosis with dynamic validation and self-supervision," *Knowl.-Based Syst.*, vol. 213, Feb. 2021, Art. no. 106679.
- [14] S. Park, Y. Suh, and J. Lee, "FedPSO: Federated learning using particle swarm optimization to reduce communication costs," *Sensors*, vol. 21, no. 2, p. 600, Jan. 2021.
- [15] Y. Qi, M. S. Hossain, J. Nie, and X. Li, "Privacy-preserving blockchain-based federated learning for traffic flow prediction," *Future Gener. Comput. Syst.*, vol. 117, pp. 328–337, Apr. 2021.
- [16] B. Zhao, K. Fan, K. Yang, Z. Wang, H. Li, and Y. Yang, "Anonymous and privacy-preserving federated learning with industrial big data," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 6314–6323, Sep. 2021.
- [17] Y. Wei, S. Zhou, S. Leng, S. Maharjan, and Y. Zhang, "Federated learning empowered end-edge-cloud cooperation for 5G HetNet security," *IEEE Netw.*, vol. 35, no. 2, pp. 88–94, Mar. 2021.
- [18] P. D. S. L. Alexandrino, G. F. Gomes, and S. S. Cunha, "A robust optimization for damage detection using multiobjective genetic algorithm, neural network and fuzzy decision making," *Inverse Problems Sci. Eng.*, vol. 28, no. 1, pp. 21–46, Jan. 2020.

- [19] A. Suresh, R. Udendhran, and M. Balamurgan, "Hybridized neural network and decision tree based classifier for prognostic decision making in breast cancers," *Soft Comput.*, vol. 24, no. 11, pp. 7947–7953, Jun. 2020.
- [20] H. Khaloufi, K. Abouelmehdi, and A. Beni-Hssane, "Fog computing in the age of big healthcare data: Powering the medical Internet of Things," in *Advances on Smart and Soft Computing*. Singapore: Springer, 2020, pp. 211–221.
- [21] C. Moore, "Medical Internet of Things-based healthcare systems: Wearable sensor-based devices, patient-generated big data, and real-time clinical monitoring," *Amer. J. Med. Res.*, vol. 7, no. 1, pp. 41–47, 2020.
- [22] K. Pradhan and P. Chawla, "Medical Internet of Things using machine learning algorithms for lung cancer detection," *J. Manage. Anal.*, vol. 7, no. 4, pp. 591–623, Oct. 2020.
- [23] M. Rashid, H. Singh, V. Goyal, S. A. Parah, and A. R. Wani, "Big data based hybrid machine learning model for improving performance of medical Internet of Things data in healthcare systems," in *Healthcare Paradigms in the Internet of Things Ecosystem*. Amsterdam, The Netherlands: Elsevier, 2021, pp. 47–62.
- [24] T. Zhang, A. H. Sodhro, Z. Luo, N. Zahid, M. W. Nawaz, S. Pirbhulal, and M. Muzammal, "A joint deep learning and internet of medical things driven framework for elderly patients," *IEEE Access*, vol. 8, pp. 75822–75832, 2020.
- [25] D. Połap, G. Srivastava, and K. Yu, "Agent architecture of an intelligent medical system based on federated learning and blockchain technology," *J. Inf. Secur. Appl.*, vol. 58, May 2021, Art. no. 102748.
- [26] D. Połap, "Analysis of skin marks through the use of intelligent things," *IEEE Access*, vol. 7, pp. 149355–149363, 2019.
- [27] J. Kaur, M. Singh, and S. K. Singla, "Study of fusion of medical images and classification comparison using different kernels of SVM and K-NN classifiers," in *Proc. 1st IEEE Int. Conf. Meas., Instrum., Control Automat. (ICMICA)*, Jun. 2020, pp. 1–6.
- [28] P. Tschandl, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Tech. Rep., 2018, doi: [10.7910/DVN/DBW86T](https://doi.org/10.7910/DVN/DBW86T).
- [29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015.
- [30] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
- [31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [32] W. Sae-Lim, W. Wettayaprasit, and P. Aiyarak, "Convolutional neural networks using MobileNet for skin lesion classification," in *Proc. 16th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2019, pp. 242–247.
- [33] M. A. Khan, Y.-D. Zhang, M. Sharif, and T. Akram, "Pixels to classes: Intelligent learning framework for multiclass skin lesion localization and classification," *Comput. Electr. Eng.*, vol. 90, Mar. 2021, Art. no. 106956.
- [34] E. Pérez, O. Reyes, and S. Ventura, "Convolutional neural networks for the automatic diagnosis of melanoma: An extensive experimental study," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101858.
- [35] Q. Jin, H. Cui, C. Sun, Z. Meng, and R. Su, "Cascade knowledge diffusion network for skin lesion diagnosis and segmentation," *Appl. Soft Comput.*, vol. 99, Feb. 2021, Art. no. 106881.
- [36] A. Inés, C. Domínguez, J. Heras, E. Mata, and V. Pascual, "Biomedical image classification made easier thanks to transfer and semi-supervised learning," *Comput. Methods Programs Biomed.*, vol. 198, Jan. 2021, Art. no. 105782.
- [37] N. Heller, E. Bussmann, A. Shah, J. Dean, and N. Papanikolopoulos, "Computer aided diagnosis of skin lesions from morphological features," Tech. Rep., 2018.
- [38] M. A. Khan, T. Akram, Y.-D. Zhang, and M. Sharif, "Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework," *Pattern Recognit. Lett.*, vol. 143, pp. 58–66, Mar. 2021.
- [39] S. Nasiri, M. Jung, J. Helsper, and M. Fathi, "Deep-CLASS at ISIC machine learning challenge 2018," 2018, *arXiv:1807.08993*.



DAWID POŁAP (Member, IEEE) received the M.Sc. degree (Hons.) in applied mathematics from the Silesian University of Technology, Gliwice, Poland, in 2017, and the Ph.D. degree in computational intelligence from the Czestochowa University of Technology, Czestochowa, Poland, in 2019. He is currently an Adjunct with the Faculty of Applied Mathematics, Silesian University of Technology. He has authored and coauthored more than 100 research papers in international conferences and journals. His main research interests include image processing and various aspects of machine learning. He was awarded by the Polish Ministry of Science and Higher Education with a "Diamond Grant" for the most talented students and scholarship for the exceptional young scientist. He has participated in various scientific projects (both as a lead investigator and as a scientific investigator) from Polish and Italian universities. He has served as an Editor for *Applied Soft Computing*, *Sensors*, *Evolutionary Intelligence*, and *Algorithms*.

• • •