Use of intuitionistic fuzzy set methods in parallel classification of image data

Alžbeta Michalíková
Department of Computer Science
Matej Bel University
Banská Bystrica, Slovakia
alzbeta.michalikova@umb.sk

Adam Dudáš
Department of Computer Science
Matej Bel University
Banská Bystrica, Slovakia
adam.dudas@umb.sk

***Abstract*—Image classification is process of allocation of image data entities to classes in such a way, that individual elements of a class bear common features. We are focusing on the methods of tire tread image data classification based on intuitionistic fuzzy sets and voting mechanism. The paper presents the design, implementation and experimental evaluation of parallel algorithms for classification of tire tread images based on intuitionistic fuzzy set methods which include similarity measure, distance function and chosen intuitionistic fuzzy negations.**

***Index Terms*—intuitionistic fuzzy sets, classification, negations, parallel computing, data decomposition**

I. Introduction

Fuzzy sets were introduced by professor Lotfi A. Zadeh in 1965 [1]. This mathematical structure was designed as a tool to work with uncertain and vague data. Later as a natural extension of fuzzy sets, intuitionistic fuzzy sets (shortly IFSs) were introduced by Krassimir Atanassov [2]. While fuzzy sets assign the specific membership degree to each data element from a set, IFSs work with membership and with non-membership degree also.

IFSs are being used in various scientific areas and applications. In the past, they were also used for classification of the position of tire tread in the image. For classification purposes the intuitionistic fuzzy similarity measure [3], intuitionistic fuzzy distance measure [4] and various intuitionistic fuzzy negations [5] were used. This identification of the possition of tire tread in a image is important in order to automatically identify (and subsequently recognize) pattern of the tire itself. In this paper we combine these individual methods to one to obtain improved results. Since there is more than 50 IFS negations, we decided to use just those, which gave comparatively better results for the image classification task.

Moreover compared to previous cases we are working with larger dataset of tire treads in combination with parallel computing methods to shorten the time of image processing and classification.

This work is structured as follows:

- Section II is concerned with works related to the classification of image data and intuitionistic fuzzy sets.
- In the Section III we present proposed approach to the classification of images with the use of IFSs.
- Section IV contains description of parallel computation models used for the classification of tire tread images.
- Evaluation of the proposed methods is presented in the Section V.

II. Related Works

This section of the work presents research related to the proposed models and methods from tire tread image classification. In the Subsection A the data preprocessing techniques used in the previous work are briefly introduced. The Subsection B contains brief introduction to the problem of intuitionistic fuzzy sets.

A. Data preprocessing

The concepts and process of data preparation and preprocessing used in this paper are based on the work presented in [3], [4], [5]. Authors of the works preprocessed image data as follows.

In the beginning of the process, the number of classes into which the images are classified is chosen. Then three images from each class are selected as the templates, which represent one of the created classes. The information from these templates is processed and merged into one vector, called pattern. Each unclassified image is also processed to the vector which is compared with the pattern vectors of each class. This comparison is done with the use of specific function. Based on the selected characteristic the image is assigned to the most suitable class.

The research presented in [3] contained the use of the intuitionistic fuzzy functions, the [4] used distance function defined on intuitionistic fuzzy sets and the article [5] was concerned with intuitionsistic fuzzy negations and their use in image classification.

B. Intuitionistic fuzzy set functions

Let X be the universe. An intuitionistic fuzzy set A is a set

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in X\}$$

of the functions $\mu_A : X \to [0, 1], \nu_A : X \to [0, 1]$ such that

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1.$$

Function μ_A is called the membership function and function ν_A is called the non-membership function. Denote by $\mathcal{F}$ the family of all IFSs. In addition define the function π_A as

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x).$$

Then this function is called hesitation margin [2].

In the paper [3] the cosine-based IFS similarity measure (see [6]) was used for classification of the images. It is defined in the following way:

Let $X = \{x_1, x_2, \dots, x_n\}$ be the discrete universe and let $A, B \in \mathcal{F}$ then the function

$$S_C(A,B) = \frac{1}{n}\sum_{i=1}^{n} \frac{\mu_A(x_i)\mu_B(x_i) + \nu_A(x_i)\nu_B(x_i)}{\sqrt{\mu_A^2(x_i) + \nu_A^2(x_i)}\sqrt{\mu_B^2(x_i) + \nu_B^2(x_i)}}$$

is called the cosine-base IFS similarity measure.

The resulting values of this measure are always from the unit interval. The image is classified into the class where the value of S_C (similarity of considered image and pattern image of specific class) is the highest.

Research presented in the paper [4] contained IFS distance functions for classification of images. The best results were obtained with the use of normalized Euclidean distance function defined on IFSs in the following way:

Let $X = \{x_1, x_2, \dots, x_n\}$ be the discrete universe and let $A, B \in \mathcal{F}$ then the function

$$d_E(A,B) = \left(\frac{1}{2n} \sum_{i=1}^{n} \Big(\left(\mu_A(x_i) - \mu_B(x_i)\right)^2 + \right.$$

$$\left. + \left(\nu_A(x_i) - \nu_B(x_i)\right)^2 + \left(\pi_A(x_i) - \pi_B(x_i)\right)^2 \Big) \right)^{\frac{1}{2}}$$

is called normalized Euclidean distance function defined on IFSs.

The resulting values of this measure are always from the unit interval. The image is classified into the class where the value d_E computed between the considered image and pattern image of the specific class is the lowest.

In paper [7] authors defined the function Sim as a tool to measure whether a element A is more similar or more dissimilar to B:

Let $A, B, B^C \in \mathcal{F}$ where B^C is a complement of B. Let d be the distance function defined on $\mathcal{F}$, then

$$\mathrm{Sim}(A,B) = \frac{d(A,B)}{d(A,B^C)}.$$

In the [5], authors used function Sim to classify a set of tire tread images. Since there is more then 50 negations defined on IFSs (see [8]), authors were focused on answering the question of *which negations improve the results of classification of images the most.* In the mentioned paper only normalized Euclidean distance function defined on IFSs was used. We used the results of this study and we used the IFS negations which led to the best results in our research. Specifically, we used the following negations (the numbering is identical to [8]):

$$\neg_1(A) = (\nu_A(x), \mu_A(x)),$$

$$\neg_3(A) = (\nu_A(x), \mu_A(x)\cdot\mu_A(x) + \mu_A(x)\cdot\nu_A(x)),$$

$$\neg_7(A) = (1 - \mathrm{sgn}(1 - \nu_A(x)), \mu_A(x)),$$

$$\neg_9(A) = (1 - \mathrm{sgn}(\mu_A(x)), \mu_A(x)),$$

$$\neg_{52}(A) = (1 - \mu_A(x), \min(1, 1 - \mu_A(x))).$$

The resulting values of these measures were always greater then or equal to zero. While using this approach the closer the value is to zero, the more similar the two IFS elements are. With the use of these approaches each image of dataset was classified into the class where the value Sim measured between the considered image and pattern image of specific class was the lowest.

III. New Approaches to Classification of Tire Images with the Use of Intuitionistic Fuzzy Sets

One of the findings presented in the research papers [3], [4], [5] is the fact, that the obtained results of classification are not always the same when using different methods. Therefore, we decided to combine all seven approaches described in the Section II:

- Step 1: We compute the values of all IFS functions (similarity measure, Euclidian distance and five types of negations) between an input image and all defined patterns (templates of classes).
- Step 2: We compute the value $1 - S_C$ in order to get comparable values for all computed functions.
- Step 3: Computed values for each input image are sorted in an ascending order.
- Step 4: We compute the arithmetic mean from best two results of functions for each class (the closer to zero, the better).
- Step 5: We found the class with the smallest value and we classify the image into this class.

In the studies mentioned in the Section II the set of approximately 350 images was used. This dataset was being classified into seven classes. In the presented paper, we are considering dataset of 4128 images and classify them with the use of the mentioned methods into the five classes (see Fig. 1). With the use of this approach we obtained number of incorrectly classified images. This problem was caused by insufficient or various lighting used in the images, number of tires in the images were with the hubcap and also number of them

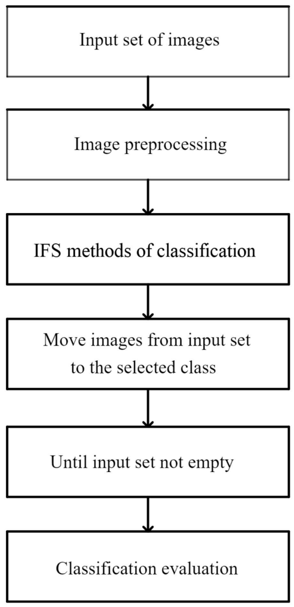

Fig. 1. Scheme of basic model for classification of images

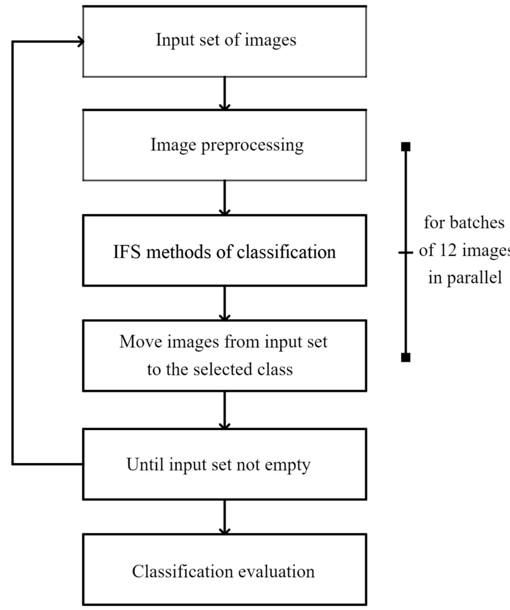

Fig. 2. Scheme of batch-based parallel computation model for classification of images

without hubcap, some of the tires were of motorcycle type, etc. Therefore, we decided to create two phase classification of image data:

- In the first phase, we created 29 image classes and classified all the images into them.
- In the second phase, we merged the classes which represented similar (or the same) position of tire tread.

IV. Parallel Approaches to Classification of Tire Images with the use of Intuitionistic Fuzzy Sets

The presented research includes - in addition to approaches to the classification of image data related to tires which use IFSs - two computational models for the parallel classification of the given images. In this section of the paper, we present a parallel classification method based on batch processing and a parallel classification method that uses data set decomposition.

These methods have been designed and implemented to reduce the time required for the image classification algorithms which use IFSs. Since we use a 12-core system for parallel computing methods, in both cases we are limited by the width of this system [9] - and so we process batches of 12 images and divide the data set into 12 subsets which can be processed in parallel.

A. Batch-based Model

Since the computing architecture we have at our disposal is not wide enough - it does not contain a large enough number of processors - we use batch processing to classify the input set of images. These batches are formed by subsets of the input set, which always contain the first 12 images of tires in the input set. Since, after classification, the images are moved to the dataset of the selected class, the new batch always contains 12 different images. Elements of each batch are processed in parallel, as indicated in the Fig. 2.

Batch-based parallel computation model (presented in the Fig. 2) consists of the following steps:

1) In parallel:
   a) First 12 images (batch) from the input set are preprocessed with the use of preprocessing method described in the Subsection II-A.
   b) Each of preprocessed images from current batch is classified with the use of classification methods described in the Section III.
   c) After the classification image is moved from input data set to the determined class data set. Moved image no longer figures as an element of the input set.
   d) Repeat the steps $a - c$ for new batches until the input set is not empty.
2) After classification itself is done for all of the images from input set, in the last step the model computes correct and incorrect classifications and elapsed time for used for the computation.

B. Data Decomposition Model

The previous computing model is simple from the point of view of implementation but its parallel speedup comes only with large data sets. Therefore, we designed a second model of parallel image processing based on the data decompostion.

In this approach to the problem, the input dataset is divided into n equally sized data parts [11], [12]. Each of these

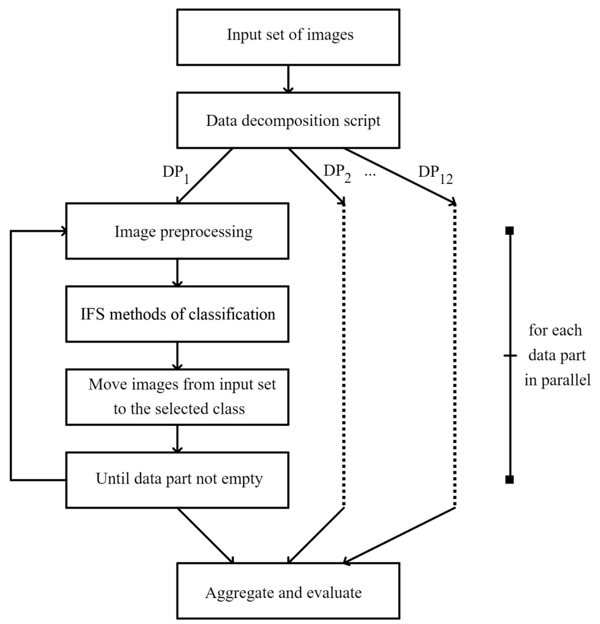

Fig. 3. Scheme of data decompostion parallel computation model for classification of images

parts is subsequently processed in parallel with the others, but the individual steps of these data parts are performed sequentially for each image. Such a method of classifying images requires the aggregation of results from individual data parts and subsequent evaluation of the overall result after the classification is completed.

Data decompostion parallel computation model (Fig. 3) consists of the following steps:

1) Decompostion of the input data set into n equally sized dataparts (DPs). In our case, we decompose the input data set into 12 dataparts.
2) In parallel for each datapart:
    a) Preprocessing of the first image of given datapart with the use of preprocessing method described in the Subsection II-A.
    b) The preprocessed image is classified with the use of classification method described in the Section III.
    c) After the classification image is moved from data part to the determined class data set. Moved image no longer figures as an element of the data part.
    d) Repeat the steps $a - c$ for new images until the data part is not empty.
3) After classification of the images in given dataparts is done, the model aggregates partial evaluations of the classification correctness from the dataparts and evaluates the final value of classification correctness. Also, elapsed time for used for the computation is measured.

V. Evaluation of the Proposed Models

In order to experimentally verify the methods proposed in the Sections III and IV, we used heterogeneous dataset of 4128 tire tread images. These images differed in their size, quality, method of retrieval, lighting used in the image and the possition of the tire in the image. As an addition to the proposed methods, we implemented one auxiliary step to the classification. Since used input set of images contains some images of insufficient size (we set the border for this insufficiency as 64 pixels on any side of an image), we decided to add function, which moves these images into their own folder in order to not take them into account.

From the set of 4128 tire tread images there was only 58 images (1.4% of the dataset) which were not of sufficient size. Therefore, 4070 images were classified in the one of designed classes.

A. Evaluation of Classification with the use of Intuitionistic Fuzzy Sets

For the verification of the proposed IFS methods in tire tread classification the set of 4128 images was used. The results of the classficiation are presented in the confusion matrix in the Table I. We considered five different classes of the tire tread images - left facing whole tires (LW), right facing whole tires (RW), left facing halves of tires (LH), right facing halves of tires (RH) and pattern images (P).

We present comparison of targeted class (columns) and obtained output class (rows) for the classification of the images. For the class cells top part of each cell contains number of tire tread images which were classified in the given class and bottom part of the cell contains percentual part of these images from the whole dataset. For the summation cells (denoted by Σ) we present overal number of the tire tread images for the given class (top of the cell) and percentual correctness of classification of these images (bottom of the cell).

Some of the tire tread images are hard to classify clearly even by humans. For example left facing whole tires and left facing half tires are in some cases hard to discern - there is no hard boundry between these classes, therefore some of the tires can be classified as halves even though they are closer to the whole tires. As seen in the Table I overall correctness of the classification of the given dataset was 78.7%. This result can be considered as satisfactory since the tire tread images are processed further after this classification.

B. Evaluation of Proposed Parallel Models

Since we worked with 4128 instances of image data, on which several operations were performed, the computation of the classification itself took a relatively long time - approximately 93.5 seconds on the system with 3.4 GHz processor and 16 GB of RAM. Therefore, we proposed parallel computational models to classify these data, which were presented in the Section IV.

For the comparison of computational time of the sequential classification, classification using batch-based parallel model and classification using data decomposition parallel model see the Table II and Fig. 4. These measurements were done on the sets of 100, 500, 1000, 2000, 4000 and 4128 images.

It is evident that the time needed for computation of the problem of size 100 and 500 is close to same in the case of sequential and batch-based models. From 1000 input images up, we can see sligh speedup in the computation of the given set of images with the use of batch-based model compared to sequential model, resulting in approximate 25 second reduction of computational time in the case of classification of the whole set (4128 images).

When comparing data decomposition model with any of the previous methods, we can see that no matter what the size of input data set is, this method computes the classification faster. In the case of classifying the whole considered set of images we can see reduction of approximately 80 seconds of computational time compared to sequential model and reduction of approximately 45 seconds compared to the batch-based model.

The Table III contains parallel speedup measured for both proposed parallel methods on the set of 100, 500, 1000, 2000, 4000 and 4128 images.

As can be seen, in both cases parallel speedup grows with the size of input dataset. Yet only in the case of the data decomposition model it is of significance - reaching approximately 7-fold speedup compared to sequential processing of the whole data set.

VI. Conclusion

This paper focuses on the classification of the position of tire tread in the image. This identification of the position of tire tread in a image is important in order to automatically identify (and subsequently recognize) pattern of the tire itself. In our case we used combination of intuitionistic fuzzy methods and parallel computing in order to obtain correct classification of the images in shortened duration.

We present novel method for image classification based on the IFS functions - similarity measure, Euclidean distance and five types of IFS negations. These methods were implement in two parallel computing models - so called Batch-based model and Data Decomposition model. Experimental evaluation of these methods presented in the Section V shows correctness of the classification of heterogeneous set of 4128 tire tread images equal to 78.7%, while proposed parallel models show more than 7-fold speedup of computational time compared to sequential computations.

In the present authors realize the classification of the tire tread images with the use of intuitionistic fuzzy equivalence matrices. After the successful implementation of this concept we are planning to use IF equivalence matrices to improve presented research. Additionally we will create the database of tire tread prints from the obtained cutout sections of the images adhering to the principles presented in [13], [14].

Acknowledgment

This contribution has been processed as part of the grant project Implementation of Blended Learning into Preparation of Future Mathematics Teachers and Future Computer Science Teachers, project no. 001UMB-4/2020.

The research was partially supported by the grant of The Ministry of Education, Science, Research and Sport of Slovak Republic - Implementation of new trends in computer science to teaching of algorithmic thinking and programming in Informatics for secondary education, project number KEGA 018UMB-4/2020.

Computing was performed in the High Performance Computing Center of the Matej Bel University in Banska Bystrica using the HPC infrastructure acquired in project ITMS 26230120002 and 26210120002 (Slovak infrastructure for high-performance computing) supported by the Research & Development Operational Programme funded by the ERDF.

The support of Advtech AirPollution project (Applying some advanced technologies in teaching and research, in relation to air pollution, 2021-1-RO01-KA220-HED-000030286) funded by European Union within the framework of Erasmus+ Program is gratefully acknowledged.

References

[1] Zadeh, L. A. (1965). Fuzzy sets, *Information and Control*, vol. 8., pp. 338-353.

[2] Atanassov, K. T. (1983). Intuitionistic Fuzzy Sets, *VII ITKR Session*, Sofia, 20-23 June 1983 (Deposed in Centr. Sci.-Techn. Library of the Bulg. Acad. of Sci., 1697/84) (in Bulgarian). Reprinted: *Int. J. Bioautomation*, 2016, 20(S1), pp. 1-6.

[3] Michalíková, A. (2019) Intuitionistic Fuzzy Sets and Their Use in Image Classification, in *Notes on Intuitionistic Fuzzy Sets*. Sofia, Bulgaria. Volume 25, Number 2, pp. 60-66.

[4] Michalíková, A. (2019) Classification of Images by Using Distance Functions Defined on Intuitionistic Fuzzy Sets, *Advances in Intelligence Systems and Computing*. Warsaw, Poland. Springer, Cham., pp. 66-74.

[5] Michalíková, A. (2020) Intuitionistic fuzzy negations and their use in image classification, *Notes on Intuitionistic Fuzzy Sets*, 26 (3), pp. 22-32.

[6] Ye, J. (2011). Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Mathematical and Computer Modelling*, 53(1–2), pp. 91-97.

[7] Szmidt, E., Kacprzyk, J.: Analysis of Similarity Measures for Atanassov's Intuitionistic Fuzzy Sets, in *IFSA/EUSFLAT Conference*, pp. 1416-1421. (2009)

[8] Atanassov, K. T. (2017). *Intuitionistic Fuzzy Logics*, Springer International Publishing AG 2017, Switzerland. ISSN 1434-9922, ISBN 978-3-319-48952-0, DOI 10.1007/978-3-319-48953-7.

[9] J. Škrinárová, E. Vesel. Model of education and training strategy for the high performance computing. *14th IEEE International Conference on Emerging eLearning Technologies and Applications*, ISBN 978-1-5090-4701-7

[10] A. Michalíková, M. Vagač. A tire tread pattern detection based on fuzzy logic. In proceedings of Advances in Intelligent Systems and Computing, 2016, Vol. 400, pp. 381-388. DOI: 10.1007/978-3-319-26154-6_29

[11] J. Škrinárová, A. Dudáš. Optimization of the Functional Decomposition of Parallel and Distributed Computations in Graph Coloring With the Use of High-Performance Computing. IEEE Access, 2022, Vol. 10, pp. 34996-35011. DOI: 10.1109/ACCESS.2022.3162215

[12] A. Dudáš, J. Škrinárová, E. Vesel. Optimization design for parallel coloring of a set of graphs in the High-Performance Computing. In proceedings of INFORMATICS 2019 - IEEE 15th International Scientific Conference on Informatics, 2019, art. no. 9119253, pp. 11-17. DOI: 10.1109/Informatics47936.2019.9119253

[13] Kvet, M., Papan, J. The Complexity of the Data Retrieval Process Using the Proposed Index Extension. (2022) IEEE Access, Vol.10, pp. 46187-46213. DOI: 10.1109/$ACCESS$.2022.3170711

[14] Steingartner, W., Novitzká, V. Natural Semantics for Domain-Specific Language. (2021) Communications in Computer and Information Science, 1450 CCIS, pp. 181-192. DOI: $10.1007/978-3-030-85082-1\_17$

TABLE I
CONFUSION MATRIX FOR CLASSIFICATION OF TIRE TREAD IMAGES BASED ON THE IFS METHODS

	LW	**RW**	**LH**	**RH**	**P**	Σ
LW	**1582** **38.87%**	211 5.18%	237 5.82%	87 2.13%	103 2.53%	2220 71.26%
RW	24 0.59%	**187** **4.59%**	0 0%	6 0.15%	8 0.2%	225 83.11%
LH	46 1.13%	12 0.3%	**1129** **27.74%**	8 0.2%	32 0.79%	1227 92.01%
RH	8 0.2%	7 0.17%	5 0.12%	**99** **2.43%**	4 0.1%	123 80.49%
P	32 0.79%	20 0.49%	13 0.32%	4 0.1%	**206** **5.06%**	275 74.91%
Σ	1692 93.5%	437 42.79%	1384 81.58%	204 48.53%	353 58.36%	**4070** **78.70%**

TABLE II
COMPARISON OF COMPUTATIONAL TIME FOR SEQUENTIAL AND PARALLEL APPROACHES TO CLUSTERING

	100 images	500 images	1000 images	2000 images	4000 images	whole set
Sequential	2.69 sec	8.67 sec	18.08 sec	37.74 sec	87.95 sec	93.58 sec
Batch-based Model	3.01 sec	8.11 sec	13.83 sec	30.24 sec	57.59 sec	58.38 sec
Data Decompostion Model	1.6 sec	2.48 sec	3.78 sec	7.25 sec	12.51 sec	13.02 sec

TABLE III
PARALLEL SPEEDUP OF COMPUTATIONAL TIME FOR VARIOUS SIZES OF INPUT DATASETS

	100 images	500 images	1000 images	2000 images	4000 images	whole set
Speedup of Batch-based Model	0.89	1.06	1.3	1.24	1.52	1.6
Speedup of Data Decompostion Model	1.68	3.49	4.78	5.2	7.03	7.18

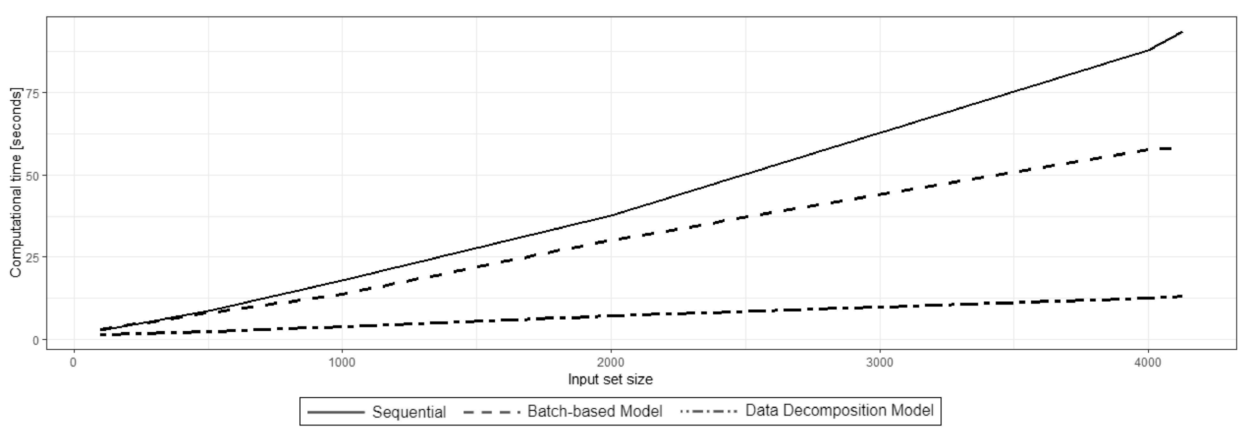

Fig. 4. Comparison of computational time for sequential and parallel approaches to clustering