VOLUME-8, ISSUE-4, APRIL-2019 • PRINT ISSN No 2277 - 8160

Original Research Paper

Engineering

FUZZY BASED INTELLIGENT CLINICAL DECISION SUPPORT SYSTEM

Dr. Anooj P. K.

College Registrar and IT Faculty Al Musanna College of Technology Sultanate of Oman.

ABSTRACT Clinical Decision Support Systems (CDSSs) have been utilizing clinical information for a considerable length of time to offer help to clinicians and enhance medicinal services. With the sharp increment in the appropriation of Electronic Health Record (EHR) frameworks, conceiving a framework fit for confronting ambiguity and vulnerability in EHRs and in the meantime protecting the interpretability of the framework is attractive in the plan of keen CDSSs. This paper introduces an upgraded fuzzy evidential (OFE) framework, which can be utilized to help choices in clinical applications as well as in other setting. We have tried the execution of OFE-CDDS on three surely understood UCI clinical datasets: Heart Disease and Pima Indians Diabetes. Our proposed framework accomplished the most noteworthy arrangement precision among comparative frameworks, with exactness of 3.36% and 8.49% over the best similar papers in tests done on the above datasets, individually. The quantity of tenets and found the middle value of number of characteristics in the precursor part of the principles were sensible. These outcomes demonstrate that OFE-CDDS can be utilized as a choice emotionally supportive network in restorative applications.

KEYWORDS : Fuzzy, CDSS, Intelligent system, EHR

1. INTRODUCTION

Keen clinical choice emotionally supportive networks have increased unique consideration in the previous decade with the expanded accessibility of electronic wellbeing information. Numerous hypothetical and exact inquires about have been attempted in this field; be that as it may, useful execution of these frameworks has been hindered by various difficulties. The intrinsic vulnerability of EHRs is one of the imposing difficulties here, and can be found in the entire procedure of medicinal basic leadership. Patients, doctors, nurture, and even lab results can be viewed as wellsprings of vulnerability. Therefore, settling this vulnerability and helping doctors in their basic leadership is the longstanding desire of specialists around there. Preventable medicinal mistakes have lamentable ramifications for individuals' wellbeing, as well as are in charge of immediate and circuitous expenses in restorative consideration. Then again, with respect to the unfaltering increment in the quantity of patients, numerous doctors guarantee that they can't assign as much time as before for inspecting patients. The requirement for keen CDSSs is, in this manner, irrefutable.

As it very well may be found, clinical side effects, research facility results and other data about a patient are sustained to a CDSS as sources of info. The yield of this framework is a finding about the nearness of an explicit ailment, gave so as to help doctors in their choice. A clever CDSS is involved two areas: computational deduction and information. Numerous calculations have been produced for the initial segment, for example, rule-based techniques, Bayesian conviction systems, heuristic strategies, counterfeit neural systems, and so forth. In any case, most canny CDSSs have a discovery structure and are not interpretable, in this way discovering little acknowledgment among doctors. As of late, frameworks dependent on fuzzy hypothesis have gotten unique enthusiasm for the field of therapeutic basic leadership because of their reasonable interpretability. Fuzzy IF-THEN principles are comparable to doctors' demeanors, and they are fit for displaying the imprecision in side effects, for example, "hypertension". Notwithstanding imprecision, we experience vulnerability during the time spent basic leadership because of an absence of data. Proof hypothesis, similar to doctors, gathers dependent on accessible proof. For example, "high fever" and "persevering hack" are considered as two bits of proof in drug. A doctor allots a conviction to every one of the conceivable analyses (cool, avian flu, and so on.) in light of this and other proof. Eventually, the doctor communicates his last determination dependent on the blend of these convictions. In this manner, a framework dependent on the quality of fuzzy hypothesis and proof hypothesis will demonstrate an extraordinary consistency with the methodology through which a doctor settles on a choice. The Fuzzy-proof methodology has been generally utilized by analysts taking a shot at choice help in drug: a versatile fuzzy evidential thinking technique for division of multi-methodology MR mind pictures [1]; a fuzzy evidential system for displaying the finding procedure for thyroid illness [2]; fuzzy evidential guidelines for hazard evaluation of coronary illness [3]; a structure for fuzzy evidential thinking for analysis of sort II diabetes [4].

In this paper, a three stage streamlined fuzzy evidential framework is proposed. In the plan of each progression, reasonable thought has been given to safeguarding the interpretability of the framework. In the initial step, a lot of fuzzy affiliation decides that have been streamlined and chosen by hereditary calculation (GA) is separated. At that point, these tenets are recast inside the system of fuzzy evidential guidelines with non-streamlined convictions. The last advance enhances the convictions for each standard and gives the last fuzzy evidential guideline base as a vigorous shrewd CDSS.

The rest of the paper is composed as pursues. Area 2, after a concise starter talk of the two speculations of intrigue, depicts each progression of the proposed methodology. Segment 3 displays the test set-up, and the proposed framework is utilized for finding of coronary illness, thyroid ailment and diabetes. Area 4 examines the outcomes. At end, segment 5 finishes up the paper.

2. THE METHOD

Dempster-Shafer theory, was first proposed by Glenn Shafer in 1976 as an extra improvement to Dempster's work. This part covers this theory in a word; an entire dialog can be found in Shafer's book [5]. The theory can be expressed as pursues:

Let H={h_1,h_2,… ,h_n} be the packaging of understanding (FD) of an issue having n thorough and absolutely random theories h_i. The course of action of all subsets of the FD is known as the power set of H, which is showed up 2H. In this hypothesis, a basic conviction errand (bba), connoted by m2^H→[0,1], is characterized on every individual from the power set of H.The function *m(A)* is nonnegative for any $A\epsilon 2^H$ and $m(\emptyset) = 0$. In Shafer's work, the sum of all basic beliefs equals one, $\sum_{A\epsilon 2^H} m(A) = 1$. Shafer also defined belief and plausibility functions on basic beliefs—Eqs. (1)-(2).

$$\forall\, A\epsilon\ 2^H : Bel(A) = \sum_{B\in 2^\Theta,\ B\subseteq A} m(B) \quad (1)$$

$$\forall\, A\epsilon\ 2^H : Pl(A) = \sum_{B\in 2^\Theta,\ B\cap A\neq\emptyset} m(B) \quad (2)$$

Those individuals from the power set H that have positive conviction esteems are called kernels and are signified by K(m). Presently, accept that two conviction capacities, Bel_1 (.) and Bel_2 (.), are given by two proof sources, B_1 and B_2, and we have the essential conviction assignments, m_1 (.) and m_2 (.), individually.

Dempster proposed Eq. (3) for consolidating the essential conviction assignments. $m(.) = [m_1 \oplus m_2](.) =$

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \dfrac{\sum_{\substack{X,Y \in 2^\Theta \\ X \cap Y = A}} m_1(X) m_2(Y)}{1 - \sum_{\substack{X,Y \in 2^\Theta \\ X \cap Y = \emptyset}} m_1(X) m_2(Y)} \quad \forall (A \end{cases} \quad (4)$$

We define the degree of conflict between two sources of evidence as shown in Eq. (4).

$$k_{12} \triangleq \sum_{\substack{X,Y \in 2^\Theta \\ X \cap Y = \emptyset}} m_1(X) m_2(Y) \quad (4)$$

As can be seen from Eqs. (3)-(4), when the sources of evidence are in complete conflict with each other, which means that $k_{12} = 1$ Eq. (3) does not provide a reliable belief. This conflict is seen when for $A \subseteq 2^{\mathrm{H}}$, we have $Bel_1(A) = 1$ and $Bel_2(\bar{A}) = 1$. To address the weaknesses and limitations of the Dempster-Shafer combination rule, many alternative combination rules have been proposed by several researchers[6-8]. In this paper, we employ Murphy's combination rule[7], which is the arithmetic average of beliefs as shown in Eq. (5).

$$Bel_M(A) = \frac{1}{2}[Bel_1(A) + Bel_2(A)] \quad (5)$$

2.1 Fuzzy affiliation rules for classification

Fuzzy affiliation rules are generally utilized for exhibiting the connection between highlights in a database. These guidelines are communicated as "X is A→Y is B", where An and B are the fuzzy sets identified with X and Y. In this documentation, "X is An" is known as the forerunner and "Y is B" is the consequent piece of the standard. Affiliation rules give the upside of finding the connection between physiological factors in clinical applications. For example, the negative connection between's high glucose levels and Glasgow trance state score (GCS) [9] can be accounted for as affiliation rules, "On the off chance that Deep extreme lethargies, high glucose level." It gives the idea that cooperative principles can promptly be deciphered by a doctor. The fuzzy affiliation rules are typically evaluated by two measures, support and certainty [10], and are characterized as pursues: $Support(< X, A > \rightarrow < Y, B >) = \frac{\sum_{x_i \in D} \mu_{AB}(x_i)}{|N|}$

$$Confidence(< X, A > \rightarrow < Y, B >) = \frac{\sum_{x_i \in D} \mu_{AB}(x_i)}{\sum_{x_i \in D} \mu_A(x_i)}$$

where the parameters N, N, $\mu_A(x_i)$ and $\mu_{AB}(x_i)$, separately mean the quantity of occasions in database D, how much occurrence x_i matches the precursor part of the standard, and how much case x_i matches the standard.

2.1 The proposed system

In this subsection, we examine the proposed framework in detail. A schematic of the proposed enhanced half breed fuzzy evidential master framework is illustrated below.

The proposed framework comprises of three noteworthy points:

1. Extracting essential fuzzy affiliation rules.
2. Adapting the essential fuzzy standards to a fuzzy evidential system.
3. Optimizing the parameters of the fuzzy evidential guideline base.

The proposed framework makes utilization of the preparation information and gives an enhanced principle base. In what pursues, we explain the three referenced strides in detail and talk about the qualities of each part.

Step 1 : Extracting essential fuzzy affiliation rules

In this stage, we will separate fuzzy affiliation rules from our information framework. To set up our essential standard base, we utilized an adjusted variant of the pursuit tree. Without loss of simplification, we arrange our traits in any subjective way. The base of the tree will be the primary property, which partitions into all characterized participation works on it. In the following dimension, we compose all the conceivable twofold mixes of the fuzzy sets on the first and different traitsWe will proceed with our branches yet remember the standard that we are not permitted to extend the tree in backward request of properties. The benefit of this pursuit tree is its capacity to bounce from the primary credit to any after quality without the commitment of considering the ones in the middle. This legitimacy prompts building the best short principles

Step 2: Adapting the essential fuzzy standards to a fuzzy evidential system

This step is the major stage in the design of the proposed system. In this step, we adapt the rule base to a fuzzy evidential framework and inject beliefs into our rules.

We can expand our primary fuzzy rule base to the fuzzy evidential rule base by considering a belief value for each consequent class. Thus, we will rewrite our rules in the form of

$$\begin{aligned} &R_k: IF\ x_1 is\ A_{j1} and \ldots and\ x_m is\ A_{jm} \\ &THEN\ \{(C_1, \bar{\beta}_{1k}), (C_2, \bar{\beta}_{2k}), \ldots, (C_N, \bar{\beta}_{Nk})\} \\ &(\textstyle\sum_{i=1}^{N} \bar{\beta}_{ik} \leq 1), \\ &with\ a\ rule\ weight\ \theta_k\ k \epsilon \{1, \ldots, N\} \end{aligned} \quad (6)$$

where L is the number of rules in the rule base. As can be seen in Eq. (6), a belief value, $\bar{\beta}_{ik}$, is assigned to the possibility that instance x belongs to the ith class. We assume that the sum of beliefs can be equal or less than one. When we restrict our rules to have $\sum_{i=1}^{N} \beta_{i,k} = 1$, we are implicitly expressing that we have all the possible hypotheses in the consequent part of our rules. However, without this restriction and allowing the summation to be equal to one or less, we are assigning an invisible belief to uncertainty (any other hypotheses that we are not aware of). This degree of freedom is very useful in clinical applications. Consider a patient with persistent cough. There are several possible causes for this symptom, for instance asthma or postnasal drip. However, we cannot definitely reject other hypotheses like lung cancer, bronchiectasis, etc. In this case, although we may have just the most common hypotheses in the consequent part of our rules, by assigning $\sum_{i=1}^{N} \bar{\beta}_{ik} \leq 1$ we are indicating that we are aware of other possible unknown hypotheses. In other words, we are considering our uncertainty due to lack of information.

The activation weight of a rule is calculated as

$$w_k = \frac{\theta_k \prod_{i=1}^{T_k} \alpha_{i,j}^k}{\sum_{l=1}^{L} (\theta_l \prod_{i=1}^{T_l} \alpha_{i,j}^l)}$$

where T_k is the number of attributes in the antecedent part of the $k$tu rule, and $\alpha_{i,j}^k$ is the membership degree of the input in the jth fuzzy set defined on the ith attribute. In the next step, we can define the basic belief assigned to each possible class as $m_{n,k} = w_k \beta_{n,k}$ $n = 1, \ldots, N$, where N denotes the number of hypotheses.

Thus far, we have defined a basic belief for each consequent class in each rule. The final step is combining these beliefs. In this paper, we used the Murphy combination rule, shown in Eq. (5), to combine beliefs from various sources of evidence (rules).

In order to preserve the capability of the system in encountering missing features, we propose that the belief values can be updated with respect to the rate of missingness. Thus, if features are missing the beliefs will be updated as

$$\beta_{ik} = \bar{\beta}_{ik} \frac{\sum_{t=1}^{T_k} (M_k(t) \sum_{j=1}^{S_t} \alpha_{tj})}{\sum_{t=1}^{T_k} M_k(t)} \quad (7)$$

where α_{tj} indicates the membership degree of the input in the jth fuzzy set defined on the tth attribute, and S_t denotes the number of fuzzy membership functions defined on the tth attribute. $M_k(t)$ shows the missing status of the tth attribute in the kth rule, and is defined as

$$M_k(t) = \begin{cases} 1, & tth\ atribute\ is\ not\ missed \\ 0, & other\ wise \end{cases} \quad (8).$$

We reduce the belief values according to the portion of missing data.

Step 3: Optimizing the parameters of the fuzzy evidential guideline base.

In the past advances, we got an essential fuzzy evidential standard base. In this progression, we will acquire the varied standard base by utilizing GA to choose the best guidelines. Besides, as can be found in Eq. (6), there are a few parameters to be set in each standard. Other than the parameters, we ought to characterize the exact state of fuzzy participations on each trait.

3.Experiments and results

3.1Data

In this investigation, we dissect the execution of our proposed CDSS on three understood clinical informational index from the UCI machine learning archive (http://mlearn.ics.uci.edu /ML Repository.html) [19]. In what pursues, we research the execution of the proposed CDSS on the accompanying datasets: the coronary illness dataset [20], PIMA Indians diabetes dataset [21]. A short portrayal of each dataset is given in Table 1.

Table 1 **Dataset description.**

Name	# instances	# features	# classes
Heart Disease (Cleveland)	297	13	5
PIMA	336	8	2

3.2Experiment design and results

So as to examine the execution of the proposed framework, a 10-overlay cross-approval technique was utilized, part the dataset into 10 sections, preparing the framework on nine of these parts, and utilizing the staying one for the test. The correctnesses on the 10 runs were arrived at the midpoint of and announced with standard deviation.

The Cleveland dataset was utilized as an outstanding testing lopsidedness dataset with five classes. The dispersion of occurrences in each class is appeared Table 2.

Table 2 Number of instances in each class of the Cleveland dataset.

Class Name	Number of instances
Risk 0	164
Risk 1	55
Risk 2	36
Risk 3	35
Risk 4	13

Despite the fact that in numerous examinations on this dataset, experimentation has been constrained to researching two classes (occasions with risk 0 and occurrences with a danger of more than zero), we explore the execution of our framework in the two situations (2-classes and 5-classes). To conquer the difficulties of imbalanced information, we utilize a various leveled structure in removing rules in stage 1 and 2. In this development, at the zero dimension, we extricate rules for the 2-class issue. At that point, we simply consider cases of patients experiencing coronary illness and concentrate runs so as to arrange the power of the ailment.

The quantity of principles and normal number of traits in the precursor part of the standards for each trial are appeared Table 3.

Table 3 **Number of rules and attributes for each experiment.**

Name	Number of rules	Average number of attributes in antecedence
Heart Disease (2-classes)	27	2.57
Heart Disease (5-classes)	85	2.57
PIMA	47	3.72

Table 4 compares the performance of our proposed system with those of similar studies done on each dataset. The result of the best ethod for each dataset is in bold format.

The quantity of principles and the normal number of characteristics in the antecedences in investigations done on Heart sickness and the Pima dataset are appeared in Figure 1 and contrasted and FARC-HD . The order precision is likewise indicated.

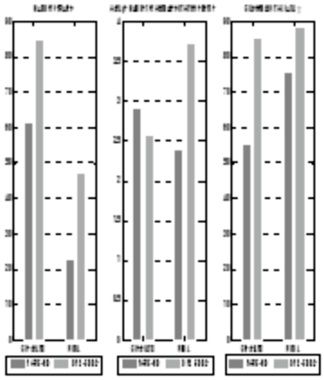

Figure 1 **Comparison of FARC-HD and OFE-CDDS on Cleveland dataset and Pima dataset.**

As can be found in Figure 1, despite the fact that our proposed framework has a bigger number of guidelines than FARC-HD, it exhibits better arrangement precision on both datasets.

The outcomes uncover that our proposed strategy performs well on clinical information. Along these lines, this framework can be utilized as a choice emotionally supportive network in restorative applications. Moreover, the low estimation of standard deviation demonstrates that this framework can give vigorous outcomes.

3.Discussion and Conclusion

This paper has presented an advanced fuzzy evidential framework. Nonlinear conduct of the human body and inborn vulnerability in clinical information are two noteworthy difficulties in the plan of savvy CDSSs. Therefore, the structure of an interpretable wise CDSS, equipped for beating this vulnerability was our principle objective. Amid advancement, we endeavored to save the interpretability of our framework and in the meantime accomplish the best exactness in basic leadership. In the initial step, we made utilization of fuzzy affiliation rules, which are fit for communicating the connection between the side effects and indications of a patient. This methodology is comparable to the essential advance of basic leadership by a doctor, as talked about in the past areas. In this way, in this progression, we thought about the relationship and dubiousness of manifestations. In the second step, we endeavored to show the vulnerability in basic leadership by infusing conviction esteems into our standard base. This demonstrating is like genuine circumstances in which doctors express their level of assurance about an illness conclusion; for example, a doctor communicates "I immovably accept … ". In the last advance, we upgraded our standard base by utilizing a GA. In advancing the conviction esteems, we accepted that there are other obscure speculations, and by permitting the summation of the conviction esteems to be equivalent to one or less, we added a level of opportunity to our framework. We at that point executed the framework on three surely understood datasets to research the two its precision and interpretability. By accomplishing 93.36% precision in arranging the Heart Disease dataset in a 2-class situation, our framework performs superior to those in practically identical research. To keep away from difficulties because of imbalanced information, we proposed a various leveled technique to remove rules. In this usage, the method was practically equivalent to the conduct of a doctor; a doctor initially looks at the likelihood of an infection and afterward makes utilization of other data to explore its seriousness. This various leveled approach is likewise in accordance with exactness prescription objectives: giving redid and exact human services to patients. Getting 85.56% precision demonstrates the solid execution of our framework in experiencing imbalanced clinical information. In examinations done on the diabetes dataset and thyroid datasets, the precision of our framework (88.09% and 100%, separately) surpasses that of others. The sensible number of tenets in each investigation and the low number of properties in the predecessor part of the standards are different points of interest of the proposed framework. In spite of the fact that our framework exhibits a decent execution on a few clinical datasets, its execution on different EHRs presently can't seem to be build up. An overview should be possible to consider the effect of utilizing other mix runs in execution of the framework.

REFERENCES

1. Shachter, R. D., and Heckerman, D. E., Thinking backward for knowledge acquisition, AI Mag. 8, 55-63, 1987.
2. Bayes, T., An essay towards solving a problem in the doctrine of changes, Biometrika 46, 293-298, 1958; reprint of original work of 1763.
3. Von Neumann, J., and Morgenstern, 0., Theory of Games and Economic Behavior, Princeton University Press, Princeton, N.J., 1947.
4. Laplace, P. S., A Philosophical Essay on Probabilities, Wiley, New York, 1902; Translated from the 6th French edition, 1812.
5. Lehman, R. S., On confirmation and rational betting, J. Symb.Logic 20, 251-262, 1955.
6. Shimony, A., Coherence and the axioms of confirmation, J. Symb. Logic 20, l- 28, 1955.
7. Cox, R., Probability, frequency and reasonable expectation, Am. J. Phys. 14(1);1-13, 1946.
8. Tribus, M., Rational Descriptions, Decisions, and Designs, Pergamon Press, New York, 2017.
9. Lindley, D. V., Scoring rules and the inevitability of probability.Int. Stat. Rev.50, 1-26, 1982.
10. Horvitz, E. J., Heckerman, D. E., and Langlotz, C. P., A framework for comparing alternative formalisms for plausible reasoning, in Proceedings AAA/- 86 Fifth National Conference on AI, AAAI, Philadelphia, Penn., 210-214, August 1986.
11. Zadeh, L. A., The role of fuzzy logic in the management of uncertainty in expert systems, Fuzzy Sets Syst. 11, 199-227, 2017.
12. Cohen, P. R., Heuristic Reasoning About Uncertainty:An AI Approach,Pitman, London, 1985.
13. Shafer, G., Probability judgment in artificial intelligence, in Uncertainty in Artificial Intelligence. L. N. Kana!and J. F. Lemmer, (Eds.), in North-Holland, New York, 2008.
14. Duda, R., Gaschnig, J., and Hart, P. Model design in the PROSPECTOR consultant system for mineral exploration, in Expert Systems in the Microelec tronic Age (D. Michie (Ed.), Edinburgh University Press, Edinburgh, Scotland,153-167, 1979.
15. Savage, L. J., The Foundations of Statistics, 2nd ed., Dover, New York, 1972. (First edition 1954.)
16. De Finetti, B., Theory of Probability, Wiley, New York, 1970.
17. Fishburn, P. C., Subjective expected utility: a review of normative theories.Theory Decision 13, 139-199,1981.
18. Patrick, E. A., Review of pattern recognition in medicine, IEEE Trans. Syst. Man Cybern. 6, 2009.
19. Hammond, K. R., Towards a unified approach to the study of expert judgment, in Expert Judgment and Expert Systems, (J. L. Mumpower, Ed.), Springer-Verlag, Berlin, Heidelberg, 1-16, 1987.
20. Carrol, N., Expert systems for clinical diagnosis: are they worth the effort? Behav.Sci. 32, 274-292, 1987.
21. Howard, R. A., Decision analysis: perspectives on inference, decision, and experimentation, Proc. IEEE 58(5), 632-643, 1970. Also in [75] Chapter 43.
22. Kahneman, D., Slovic, P., and Tversky, A. (Eds.), Judgement under Uncertainty: Heuristics and Biases, Cambridge University Press, 2016.
23. Raiffa, H., and Schlaifer, R., Applied Statistical Decision Theory, Harvard University Press, Cambridge, Mass., 1961.
24. Howard, R. A., Decision analysis: applied decision theory, in Proceedings of the 4th International Conference on Operational Research, (D. B. Hertz and J. Melese, Eds.), 55-77, 1966. Also in Chapter 4.
25. Raiffa, H., Decision Analysis:Introductory Lectures on Choice Under Uncertainty, Addison-Wesley,Reading, Mass., 1968.
26. Howard, R. A., and Matheson, J. E. (Eds.), Readings on the Principles and Applications of Decision Analysis, Strategic Decisions Group, Menlo Park, Cal.,1984.
27. Von Winterfeldt, D. and Edwards, W., Decision Analysis and Behavioral Research, Cambridge University Press, New York, 1986.
28. Keeney, R. L., and Raiffa, H.,Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Wiley, New York, 1976.
29. McDermott, J., R1: a rule-based configurer of computer systems, AI 19, 39-88, 2011.
30. De Kleer, J. and Williams, B.,Diagnosing multiple faults, AI 32(1), 97-130, 1987.
31. Genesereth, M., The use of design descriptions in automated diagnosis, AI 24,311-319, 1984.
32. Reggia, J. A., Diagnostic expert systems based on a set covering model, Int. J. Man-Mach. Stud. 19, 437-460, 1983.
33. Szolovits, P., and Pauker, S. G., Categorical and probabilistic reasoning in medical diagnosis, AI 11, 115-144, 2017.
34. Warner, H. R., Toronto, A. F., Veasy, L. G., and Stephenson, R., A mathematical approach to medical diagnosis: application to congenital heart disease, JAMA 177,177-183, 1961.
35. De Dombal, F. T., Leaper, D. J., Staniland, J. R., McCann, A. P., and Horrocks, J. C., Computer-aided diagnosis of acute abdominal pain, Br. Med. J. 2, 9-13,2018.