Soft Computing
https://doi.org/10.1007/s00500-022-07407-y

METHODOLOGIES AND APPLICATION

A novel distance between single valued neutrosophic sets and its application in pattern recognition

Minxia Luo[1] · Guofeng Zhang[1] · Lixian Wu[1]

Accepted: 3 January 2022

Abstract

Among the most classic measures in single valued neutrosophic sets ($SVNS$) theory, the distance is an important tool to compare and calculate degree of difference between $SVNS$. Although there exist some types of distance for single valued neutrosophic sets, most of them lack of strictly axiomatic definition and exist counter-intuitive cases. In this paper, a novel distance between single valued neutrosophic sets based on matrix norm is given. Then we proved that the new distance satisfies the axiomatic definition of the metric. Finally, the distance is applicated to pattern recognition and medical diagnoses.

Keywords Single valued neutrosophic sets · Distance measure · Pattern recognition · Medical diagnoses

1 Introduction

In the real world, uncertain, vague, inaccurate information can be found in many domains. The theory of fuzzy set had been proposed by Zadeh (1965), which had been a great success in dealing with uncertainty. Later, some extension of the fuzzy set has been studied by many scholars. For example, Atanassov (1986) presented intuitionistic fuzzy sets, which are characterized by a membership function and a non-membership function. Intuitionistic fuzzy sets are considered to be more effective method to handle uncertain, vague information. However, the intuitionistic fuzzy sets cannot better depict vague, indeterminate and inconsistent information. For example, three groups of experts evaluate the benefits of the fund, the first groups of experts believes that the probability of the fund will be profitable is 0.6, the second groups of experts believes that the probability of the fund will be loss is 0.3, the third groups of experts not sure whether the fund will be profitable is 0.3. In order to more comprehensive handle indeterminate and inconsistent information, Smarandache (1999, 2010) introduced the concept of neutrosophic set, which characterized by a true membership function $t_A(x)$, a false membership function $f_A(x)$ and an indeterminate function $i_A(x)$, where $t_A(x), i_A(x), f_A(x) \in]0^-, 1^+[$. Ma et al. (2019) studied generalized neutrosophic extended triplet group. Since the nonstandard interval $]0^-, 1^+[$ may result in neutrosophic set is not easy to apply in practical problems. Smarandache (1998) and Wang et al. (2010) proposed single valued neutrosophic set ($SVNS$), which characterized by a true membership function $t_A(x)$, a false membership function $f_A(x)$ and an indeterminate function $i_A(x)$, where $t_A(x), i_A(x), f_A(x) \in [0, 1]$. In recent years, single-valued neutrosophic set has been applied in many fields including deductive filters (Borzooei et al. 2017), clustering analysis (Karaaslan 2017), decision-making problems and control theory (Liu and Li 2017; Liu 2016; Huang 2016; Luo et al. 2019; Zhang et al. 2018).

Distance measure is effective mathematical tool to measure the degree of difference between two objects. In recent years, the study of distance has attracted considerable attention of many researchers. As a pioneering research, (Broumi and Smarandache 2013b) presented Hausdorff distance between neutrosophic sets (NS). Later, the same author (Broumi and Smarandache 2015) extended Hausdorff distance for single valued neutrosophic sets and applied in medical diagnosis. The Hamming distance, Euclidean distance, normalized Hamming distance and normalized Euclidean distance between neutrosophic sets were studied by Majumdar (2015). The generalized weighted distance between neutrosophic sets was given by Ye (2013). Broumi et al. (2014) gave generalized weighted distance between interval neutrosophic sets. Huang (2016) defined a new

✉ Minxia Luo
mxluo@cjlu.edu.cn

[1] Department of Information and Computing Science, China Jiliang University, Hangzhou 310018, People's Republic of China

Published online: 07 September 2022

distance measure, which consider the relationship of truth membership function, indeterminacy-membership function and falsity-membership function. In order to more precision, Garg Nancy (2017) proposed some new distance measures along with two parameters and applied them to pattern recognition and medical diagnoses. Besides, Samuel et al. introduced cosine logarithmic distance (Samuel and Narmadhagnanam 2018) between single valued neutrosophic sets. Ren et al. (2019) studied a Chi-square distance-based similarity measure between single valued neutrosophic sets and applied it to pattern recognition. Similarity measure can be used express the degree of similarity between two objects. Majumdar and Samanta (2014) studied some similarity measures between single valued neutrosophic sets. Broumi and Smarandache (2013b) given several similarity measures. Hamming distance, Euclidean distance and similarity measures between interval neutrosophic sets were studied by Ye (2014a). Ye (2014b) studied vector similarity measure for single valued neutrosophic sets and applied it to multi-criteria decision-making problems. Recently, some similarity measures between single valued neutrosophic sets have been put forward, such as similarity measures based on tangent function, cotangent function and generalized Dice measures (Ye and Fu 2016; Ye 2017b, 2015).

Although the existing distance takes into account the interrelationship between elements, there are often unreasonable results in some cases (see Example 1). In order to deal with these cases, motivated by Luo and Zhao (2018) proposed distance measure between intuitionistic fuzzy sets, we give a novel distance between single valued neutrosophic sets based on matrix norm and a strictly monotonic real function.

The rest of this paper is organized as follows. In Sect. 2, we review some basic concepts for single valued neutrosophic sets and main results related to distance measure. In Sect. 3, a new distance between single valued neutrosophic sets is proposed. In Sect. 4, the effectiveness and feasibility of the novel distance measures are illustrated with some numerical examples of pattern recognition as well as medical diagnoses. The final section includes conclusion and further research.

2 Preliminaries

In this section, we review some basic concepts for single valued neutrosophic sets and main results related to distance measure, which will be applied to other sections of the paper. In this paper, $SVNS(X)$ stands for all single valued neutrosophic subsets on universal $X = \{x_1, x_2, \ldots, x_n\}$.

Definition 1 (Wang et al. 2010; Smarandache 1998) A single valued neutrosophic set A on universal $X = \{x_1, x_2, \ldots, x_n\}$ can be defined as follows

$$A = \{\langle t_A(x_j), i_A(x_j), f_A(x_j)\rangle | x_j \in X\},$$

where $t_A(x_j), i_A(x_j), f_A(x_j) \in [0, 1]$ for each x_j in X, a truth-membership function $t_A(x_j)$, an indeterminacy-membership function $i_A(x_j)$, and a falsity-membership function $f_A(x_j)$, the sum of $t_A(x_j)$, $i_A(x_j)$ and $f_A(x_j)$ satisfies the condition $0 \leq t_A(x_j) + i_A(x_j) + f_A(x_j) \leq 3$.

Definition 2 (Wang et al. 2010) Let A, B be two single valued neutrosophic sets on universal $X = \{x_1, x_2, \ldots, x_n\}$, the following relations are defined as follows:

(1) $A \subseteq B$ if and only $t_A(x_j) \leq t_B(x_j)$, $i_A(x_j) \geq i_B(x_j)$ and $f_A(x_j) \geq f_B(x_j)$;
(2) $A = B$ if and only $A \subseteq B$ and $B \subseteq A$;
(3) $A \cap B = \langle \min(t_A(x_j), t_B(x_j)), \max(i_A(x_j), i_B(x_j)), \max(f_A(x_j), f_B(x_j))\rangle$;
(4) $A \cup B = \langle \max(t_A(x_j), t_B(x_j)), \min(i_A(x_j), i_B(x_j)), \min(f_A(x_j), f_B(x_j))\rangle$;
(5) $A^c = \{\langle f_A(x_j), 1 - i_A(x_j), t_A(x_j)\rangle | x_j \in X\}$.

Definition 3 (Liu 2005) Let $M^{n\times n}$ be a matrix space, a mapping $\|U\|$: $M^{n\times n} \to R$ is called a matrix norm of U, if the following properties holds:

(1) $\|U\| \geq 0$, for any $U \in M^{n\times n}$, $\|U\| = 0$ if and only if $U = 0$;
(2) $\|aU\| = |a|\|U\|$, $a \in M$, for any $U \in M^{n\times n}$;
(3) $\|U + V\| \leq \|U\| + \|V\|$, for any $U, V \in M^{n\times n}$;
(4) $\|UV\| \leq \|U\|\|V\|$, for any $U, V \in M^{n\times n}$.

Lemma 1 (Xu and Zhang 2005) *If* $\|U\|$ *is a matrix norm of* U, *then* $\|-U\| = \|U\|$.

Definition 4 (Choudhary 1992) A metric space is an order pair (X, d) where X is a set and d is a metric on X, i.e., a real function d: $X \times X \to [0, +\infty)$ such that for any $x, y, z \in X$, the following holds:

$(d1)$ $d(x, y) \geq 0$;
$(d2)$ $d(x, y) = 0$, if and only $x = y$;
$(d3)$ $d(x, y) \leq d(x, z) + d(y, z)$.

The function d is called a distance.

We recall some existing distance between single valued neutrosophic sets on universal $X = \{x_1, x_2, \ldots, x_n\}$, which are given as follows: The extended Hausdorff distance (Broumi and Smarandache 2013a):

$$d_H(A, B) = \frac{1}{n}\sum_{j=1}^{n} \max\left\{|t_A(x_j) - t_B(x_j)|, |i_A(x_j) - i_B(x_j)|, |f_A(x_j) - f_B(x_j)|\right\}$$

The parametric normalized Hamming distance (Garg Nancy 2017):

$$d_{TH}(A,B) = \frac{1}{3n(2+t)}\sum_{j=1}^{n}\Big\{\big|-t|t_A(x_j)-t_B(x_j)| + |i_A(x_j)-i_B(x_j)| + |f_A(x_j)-f_B(x_j)|\big| + \big||t_A(x_j)-t_B(x_j)| - t|i_A(x_j)-i_B(x_j)| - |f_A(x_j)-f_B(x_j)|\big| + \big||t_A(x_j)-t_B(x_j)| - |i_A(x_j)-i_B(x_j)| - t|f_A(x_j)-f_B(x_j)|\big|\Big\}.$$

The normalized Hamming distance (Majumdar 2015):

$$d_{NH}(A,B) = \frac{1}{3n}\sum_{j=1}^{n}\big\{|t_A(x_j)-t_B(x_j)| + |i_A(x_j) - i_B(x_j)| + |f_A(x_j)-f_B(x_j)|\big\}.$$

The normalized Euclidean distance (Majumdar 2015):

$$d_{NE}(A,B) = \left(\frac{1}{3n}\sum_{j=1}^{n}\big\{(t_A(x_j)-t_B(x_j))^2 + (i_A(x_j)-i_B(x_j))^2 + (f_A(x_j)-f_B(x_j))^2\big\}\right)^{\frac{1}{2}}.$$

The Gulfam Shahzadi's distance measure (Shahzadi et al. 2017):

$$d_{S_1}(A,B) = 1-\frac{1}{n}\sum_{j=1}^{n}\{\min(t_A(x_j),t_B(x_j)) + \min(i_A(x_j),i_B(x_j)) + \min(f_A(x_j),f_B(x_j))\} / \{\max(t_A(x_j),t_B(x_j)) + \max(i_A(x_j),i_B(x_j)) + \max(f_A(x_j),f_B(x_j))\}.$$

$$d_{S_2}(A,B) = 1-\left(\sum_{j=1}^{n}\{\min(t_A(x_j),t_B(x_j)) + \min(i_A(x_j),i_B(x_j)) + \min(f_A(x_j),f_B(x_j))\}\right) / \left(\sum_{j=1}^{n}\{\max(t_A(x_j),t_B(x_j)) + \max(i_A(x_j),i_B(x_j)) + \max(f_A(x_j),f_B(x_j))\}\right).$$

$$d_{S_3}(A,B) = 1-\frac{1}{n}\left[\sum_{j=1}^{n}\Big(1-\frac{1}{3}\{|t_A(x_j)-t_B(x_j)| + |i_A(x_j)-i_B(x_j)| + |f_A(x_j)-f_B(x_j)|\}\Big)\right].$$

$$d_{S_4}(A,B) = \sum_{j=1}^{n}\{|t_A(x_j)-t_B(x_j)| + |i_A(x_j)-i_B(x_j)| + |f_A(x_j)-f_B(x_j)|\} / \sum_{j=1}^{n}\{|t_A(x_j)+t_B(x_j)| + |i_A(x_j)+i_B(x_j)| + |f_A(x_j)+f_B(x_j)|\}.$$

The cosine distance measure (Ye 2015a):

$$d_{CS_1}(A,B) = 1-\frac{1}{n}\sum_{j=1}^{n}\cos\left[\frac{\pi}{2}\max\big\{|t_A(x_j)-t_B(x_j)|, |i_A(x_j)-i_B(x_j)|, |f_A(x_j)-f_B(x_j)|\big\}\right].$$

$$d_{CS_2}(A,B) = 1-\frac{1}{n}\sum_{j=1}^{n}\cos\left[\frac{\pi}{6}\big\{|t_A(x_j)-t_B(x_j)| + |i_A(x_j)-i_B(x_j)| + |f_A(x_j)-f_B(x_j)|\big\}\right].$$

The tangent distance measure (Ye and Fu 2016):

$$d_{T_1}(A,B) = \frac{1}{n}\sum_{j=1}^{n}\tan\left[\frac{\pi}{4}\max\big\{|t_A(x_j)-t_B(x_j)|, |i_A(x_j)-i_B(x_j)|, |f_A(x_j)-f_B(x_j)|\big\}\right]$$

$$d_{T_2}(A,B) = \frac{1}{n}\sum_{j=1}^{n}\tan\left[\frac{\pi}{12}\big\{|t_A(x_j)-t_B(x_j)| + |i_A(x_j)-i_B(x_j)| + |f_A(x_j)-f_B(x_j)|\big\}\right]$$

The vector distance measure (Ye 2014b):

$$d_V(A,B) = 1-\frac{1}{n}\sum_{j=1}^{n}\Big[(t_A(x_j)t_B(x_j) + i_A(x_j)i_B(x_j) + f_A(x_j)f_B(x_j)) / \sqrt{t_A^2(x_j)+i_A^2(x_j)+f_A^2(x_j)} \cdot \sqrt{t_B^2(x_j)+i_B^2(x_j)+f_B^2(x_j)}\Big].$$

3 A new distance between single valued neutrosophic sets

In this section, we present a new distance between single valued neutrosophic sets on universal $X = \{x_1, x_2, \ldots, x_n\}$.

Theorem 1 *Let A, B be two single valued neutrosophic sets on universal* $X = \{x_1, x_2, \ldots, x_n\}$. *A function* g: $[0, 1] \times [0, 1] \to [0, 1]$ *is a strictly monotonic binary function. The mapping* d_g: $SVNS(X) \times SVNS(X) \to [0, 1]$ *can be defined as follows*

$$d_g(A, B; g) = \frac{1}{3n}(\|\Pi(t_A) - \Pi(t_B)\| + \|\Pi(i_A) - \Pi(i_B)\| + \|\Pi(f_A) - \Pi(f_B)\|)$$

where

$$\Pi(t_A) = \begin{bmatrix} g(t_A(x_1), t_A(x_1)) & g(t_A(x_1), t_A(x_2)) & \cdots & g(t_A(x_1), t_A(x_n)) \\ g(t_A(x_2), t_A(x_1)) & g(t_A(x_2), t_A(x_2)) & \cdots & g(t_A(x_2), t_A(x_n)) \\ \vdots & \vdots & \cdots & \vdots \\ g(t_A(x_n), t_A(x_1)) & g(t_A(x_n), t_A(x_2)) & \cdots & g(t_A(x_n), t_A(x_n)) \end{bmatrix},$$

$$\Pi(i_A) = \begin{bmatrix} g(i_A(x_1), i_A(x_1)) & g(i_A(x_1), i_A(x_2)) & \cdots & g(i_A(x_1), i_A(x_n)) \\ g(i_A(x_2), i_A(x_1)) & g(i_A(x_2), i_A(x_2)) & \cdots & g(i_A(x_2), i_A(x_n)) \\ \vdots & \vdots & \cdots & \vdots \\ g(i_A(x_n), i_A(x_1)) & g(i_A(x_n), i_A(x_2)) & \cdots & g(i_A(x_n), i_A(x_n)) \end{bmatrix},$$

$$\Pi(f_A) = \begin{bmatrix} g(f_A(x_1), f_A(x_1)) & g(f_A(x_1), f_A(x_2)) & \cdots & g(f_A(x_1), f_A(x_n)) \\ g(f_A(x_2), f_A(x_1)) & g(f_A(x_2), f_A(x_2)) & \cdots & g(f_A(x_2), f_A(x_n)) \\ \vdots & \vdots & \cdots & \vdots \\ g(f_A(x_n), f_A(x_1)) & g(f_A(x_n), f_A(x_2)) & \cdots & g(f_A(x_n), f_A(x_n)) \end{bmatrix},$$

$\|\Pi\| = \sqrt{\lambda_{max}}$, λ *is the largest non-negative eigenvalue of positive definite matrix* $\Pi^T\Pi$ *(*Π^T *is the transpose of matrix* Π*) (Baker 2001). Then,* d_g *is a metric on* $SVNS(X)$. $(SVNS(X), d_g)$ *is called a metric space and* d_g *is called a distance on* $SVNS(X)$.

Proof ($d1$) Obviously, $d_g \geq 0$, we obtain $0 \leq g(t_A(x_i), t_A(x_j)) \leq 1$ and $0 \leq g(t_B(x_i), t_B(x_j)) \leq 1$, thus $\|\Pi(t_A) - \Pi(t_B)\| \leq n$, $\|\Pi(i_A) - \Pi(i_B)\| \leq n$ and $\|\Pi(f_A) - \Pi(f_B)\| \leq n$, then $\frac{\|\Pi(t_A)-\Pi(t_B)\|}{3n} + \frac{\|\Pi(f_A)-\Pi(f_B)\|}{3n} + \frac{\|\Pi(f_A)-\Pi(f_B)\|}{3n} \leq 1$. Therefore, $0 \leq d_g \leq 1$.

($d2$) If $d_g(A, B; g) = 0$, then $\|\Pi(t_A) - \Pi(t_B)\| + \|\Pi(i_A) - \Pi(i_B)\| + \|\Pi(f_A) - \Pi(f_B)\| = 0$, by Definition 3 (1), we have $\|\Pi(t_A) - \Pi(t_B)\|$=0, $\|\Pi(i_A) - \Pi(i_B)\| = 0$ and $\|\Pi(f_A) - \Pi(f_B)\| = 0$, then $\|\Pi(i_A)\| = \|\Pi(i_B)\|$, $\|\Pi(i_A)\| = \|\Pi(i_B)\|$ and $\|\Pi(f_A)\| = \|\Pi(f_B)\|$. Thus, we obtain $g(t_A(x_i), t_A(x_j)) = g(t_B(x_i), t_B(x_j))$, $g(i_A(x_i), i_A(x_j)) = g(i_B(x_i), i_B(x_j))$ and $g(f_A(x_i), f_A(x_j)) = g(f_B(x_i), f_B(x_j))$. Because g is a strictly monotonic binary function for each argument, so we obtain $t_A(x_i) = t_B(x_i)$, $i_A(x_i) = i_B(x_i)$ and $f_A(x_i) = f_B(x_i)$, i.e., $A = B$.

If $A = B$, then we obtain $t_A(x_i) = t_B(x_i)$, $i_A(x_i) = i_B(x_i)$ and $f_A(x_i) = f_B(x_i)$, because g is a strictly monotonic binary function for each argument, thus $g(t_A(x_i), t_A(x_j)) = g(t_B(x_i), t_B(x_j))$, $g(i_A(x_i), i_A(x_j)) = g(i_B(x_i), i_B(x_j))$ and $g(f_A(x_i), f_A(x_j)) = g(f_B(x_i), f_B(x_j))$, then $\|\Pi(t_A) - \Pi(t_B)\|$=0, $\|\Pi(i_A) - \Pi(i_B)\| = 0$ and $\|\Pi(f_A) - \Pi(f_B)\| = 0$, so $d_g(A, B; g) = 0$.

Therefore, $d_g(A, B; g) = 0$ if and only if $A = B$.

($d3$)

$$\begin{aligned} d_g&(A, B; g) \\ &= \frac{\|\Pi(t_A) - \Pi(t_B)\| + \|\Pi(i_A) - \Pi(i_B)\|}{3n} + \frac{\|\Pi(f_A) - \Pi(f_B)\|}{3n} \\ &= \frac{\|\Pi(t_A) - \Pi(t_C) + \Pi(t_C) - \Pi(t_B)\|}{3n} + \frac{\|\Pi(i_A) - \Pi(i_C) + \Pi(i_C) - \Pi(i_B)\|}{3n} + \frac{\|\Pi(f_A) - \Pi(f_C) + \Pi(f_C) - \Pi(f_B)\|}{3n} \\ &\leq \frac{\|\Pi(t_A) - \Pi(t_C)\| + \|\Pi(t_C) - \Pi(t_B)\|}{3n} + \frac{\|\Pi(i_A) - \Pi(i_C)\| + \|\Pi(i_C) - \Pi(i_B)\|}{3n} + \frac{\|\Pi(f_A) - \Pi(f_C)\| + \|\Pi(f_C) - \Pi(f_B)\|}{3n} \\ &= \frac{\|\Pi(t_A) - \Pi(t_C)\| + \|\Pi(i_A) - \Pi(i_C)\| + \|}{3n} + \frac{\Pi(f_A) - \Pi(f_C)\|}{3n} + \frac{\|\Pi(t_B) - \Pi(t_C)\| + \|\Pi(i_B) - \Pi(i_C)\|}{3n} + \frac{\|\Pi(f_B) - \Pi(f_C)\|}{3n} \\ &= d_g(A, C; g) + d_g(B, C; g) \quad \text{(by Definition 3 (3))} \end{aligned}$$

In short, $d_g(A, B; g) \leq d_g(A, C; g) + d_g(B, C; g)$. Therefore, d_g is a metric on $SVNS(X)$. □

4 Application in pattern recognition

In this section, let $g(a, b) = 1-a-b$. In order to demonstrate the effectiveness and feasibility of the proposed distance, we give a numerical example to compare the proposed distance measure with existing distance measure in Example 1. Moreover, we present a algorithm for pattern classification and medical diagnosis in Example 2–5 to illustrate the advantage of the proposed distance.

Table 1 Comparison of different distance measure in terms of counter-intuitive case (Counter-intuitive case denoted by bold)

	1	2	3	4
A_i	$\langle 0.9, 0.2, 0.3\rangle$	$\langle 0.9, 0.2, 0.3\rangle$	$\langle 0.8, 0.3, 0.2\rangle$	$\langle 0.8, 0.3, 0.2\rangle$
	$\langle 0.7, 0.1, 0.2\rangle$	$\langle 0.7, 0.1, 0.2\rangle$	$\langle 0.7, 0.1, 0.2\rangle$	$\langle 0.7, 0.1, 0.2\rangle$
B_i	$\langle 0.7, 0.2, 0.5\rangle$	$\langle 1.0, 0.3, 0.1\rangle$	$\langle 0.7, 0.3, 0.2\rangle$	$\langle 0.9, 0.2, 0.1\rangle$
	$\langle 0.6, 0.2, 0.4\rangle$	$\langle 0.6, 0.3, 0.3\rangle$	$\langle 0.6, 0.2, 0.3\rangle$	$\langle 0.8, 0.2, 0.3\rangle$
d_H (Broumi and Smarandache 2013a)	**0.200**	**0.200**	**0.100**	**0.100**
d_{NH} (Majumdar 2015)	**0.133**	**0.133**	0.067	0.100
d_{NE} (Majumdar 2015)	0.153	0.141	**0.092**	**0.092**
d_{S_1} (Shahzadi et al. 2017)	**0.279**	**0.279**	0.192	0.223
d_{S_2} (Shahzadi et al. 2017)	**0.205**	**0.205**	0.192	0.223
d_{S_3} (Shahzadi et al. 2017)	**0.133**	**0.133**	0.067	0.100
d_{S_4} (Shahzadi et al. 2017)	**0.160**	**0.160**	0.087	0.125
d_{CS_1} (Ye 2015a)	**0.049**	**0.049**	**0.012**	**0.012**
d_{CS_2} (Ye 2015a)	**0.022**	**0.022**	0.007	0.012
d_{T_1} (Ye and Fu 2016)	**0.158**	**0.158**	**0.079**	**0.079**
d_{T_2} (Ye and Fu 2016)	**0.105**	**0.105**	0.159	0.240
d_g	0.252	0.283	0.164	0.254

4.1 Numerical comparisons

Example 1 Let A, B be single valued neutrosophic sets on X, the results of the proposed distance measure compared with some existing distance measures are shown in Table 1.

For the results of Table 1, by comparing A_1, A_2 and B_1, B_2, when $A_1 = A_2$, $B_1 \neq B_2$, we can see that the distance d_H (Broumi and Smarandache 2013a), d_{NH} (Majumdar 2015), d_{S_1}, d_{S_2}, d_{S_3} and d_{S_4} (Shahzadi et al. 2017), d_{CS_1} and d_{CS_2} (Ye 2015a), d_{T_1} and d_{T_2} (Ye and Fu 2016) inconsistent with our intuition. Similarity, by comparing A_3, A_4 and B_3, B_4, we can see that the distance measures d_H (Broumi and Smarandache 2013a), d_{NE} (Majumdar 2015), d_{S_1} (Shahzadi et al. 2017), d_{CS_1} (Ye 2015a) and d_{T_1} (Ye and Fu 2016) inconsistent with our intuition. Obviously, the proposed distance is not exist counter-intuitive cases. The results show that the most effective method is our proposed the new distance.

4.2 Algorithm and applications

4.2.1 Algorithm for pattern recognition

Given universe $X = \{x_1, x_2, \ldots, x_n\}$, m patterns $P_j = \{\langle t_{P_j}(x_i), i_{P_j}(x_i), f_{P_j}(x_i)\rangle | x_i \in X\}$ $(j = 1, 2, \ldots, m)$ and a text sample $S = \{\langle t_S(x_i), i_S(x_i), f_S(x_i)\rangle | x_i \in X\}$. Which pattern does S belong to? The recognition process is as follows:

Step 1. Compute the distance measure $d(P_j, S)$ $(j = 1, 2, \ldots, m)$, between P_j and S.

Step 2. Select the smallest $d(P_{j_0}, S)$ from $d(P_j, S)$ $(j = 1, 2, \ldots, m)$, i.e., $d(P_{j_0}, S) = \min_{1\leq j\leq m} d(P_j, S)$ and then classify the test sample S to pattern P_{j_0}.

Step 3. Compute the degree of confidence (DoC), $DoC^{(j_0)} = \sum_{j=1, j\neq j_0}^{m} \left|d(P_j, S) - d(P_{j_0}, S)\right|$ (Luo and Zhao 2018). If $DoC^{(j_0)}$ is greater, then the recognition result using this distance is more believable.

4.2.2 Applications in pattern recognition

Example 2 (Garg Nancy 2017) Given patterns P_1, P_2, P_3 and the test sample S as presented in Table 2. Which pattern does S belong to?

We obtain the classification results of the distances and the confidence degree of each distance in Table 3. According to Table 3, the results show that $d_g(P_2, S) \leq d_g(P_3, S) \leq d_g(P_1, S)$, i.e., the minimum distance is $d_g(P_2, S)$. Thus, it is shown that the test sample S belongs to P_2, where the result is the same as Garg Nancy (2017). In addition, it is easy to find that d_g with the highest degree of confidence than the remaining distances in Table 3. For the results of Table 3, the distance d_g not only can classify the test sample S to pattern P_2 but also obtain a much higher degree of confidence. Therefore, the distance d_g can be seen as a reasonable application to pattern recognition.

Example 3 Given patterns P_1, P_2, P_3 and the test sample S as presented in Table 4. Which pattern does S belong to?

We obtain the classification results of the distances and the confidence degree of each distance in Table 5. According

Table 2 Patterns and test sample

	x_1	x_2	x_3	x_4
Pattern P_1	$\langle 0.7, 0.0, 0.1\rangle$	$\langle 0.6, 0.1, 0.2\rangle$	$\langle 0.8, 0.7, 0.6\rangle$	$\langle 0.5, 0.2, 0.3\rangle$
Pattern P_2	$\langle 0.4, 0.2, 0.3\rangle$	$\langle 0.7, 0.1, 0.0\rangle$	$\langle 0.1, 0.1, 0.6\rangle$	$\langle 0.5, 0.3, 0.6\rangle$
Pattern P_3	$\langle 0.5, 0.2, 0.2\rangle$	$\langle 0.4, 0.1, 0.2\rangle$	$\langle 0.1, 0.1, 0.4\rangle$	$\langle 0.4, 0.1, 0.2\rangle$
Test sample S	$\langle 0.4, 0.1, 0.4\rangle$	$\langle 0.6, 0.1, 0.1\rangle$	$\langle 0.1, 0.0, 0.4\rangle$	$\langle 0.4, 0.4, 0.7\rangle$

Table 3 The results of distance measure application in pattern recognition

Distances	$d(P_1, S)$	$d(P_2, S)$	$d(P_3, S)$	Classification results	DoC
d_H (Broumi and Smarandache 2013a)	0.375	0.250	0.125	P_2	0.375
d_{NH} (Majumdar 2015)	0.258	0.083	0.133	P_2	0.225
d_{NE} (Majumdar 2015)	0.345	0.100	0.196	P_2	0.341
d_{CS_1} (Ye 2015a)	0.215	0.021	0.101	P_2	0.274
d_{CS_2} (Ye 2015a)	0.116	0.007	0.030	P_2	0.132
d_{T_1} (Ye and Fu 2016)	0.314	0.099	0.202	P_2	0.318
d_{T_2} (Ye and Fu 2016)	0.146	0.066	0.106	P_2	0.120
d_{S_1} (Shahzadi et al. 2017)	0.473	0.246	0.349	P_2	0.330
d_{S_2} (Shahzadi et al. 2017)	0.415	0.233	0.390	P_2	0.339
d_{S_3} (Shahzadi et al. 2017)	0.258	0.083	0.133	P_2	0.225
d_{S_4} (Shahzadi et al. 2017)	0.365	0.132	0.246	P_2	0.347
d_V (Ye 2014b)	0.176	0.019	0.085	P_2	0.223
d_g	0.243	0.080	0.321	P_2	0.404

Table 4 Patterns and test sample

	x_1	x_2	x_3
Pattern P_1	$\langle 0.7, 0.5, 0.2\rangle$	$\langle 0.8, 0.3, 0.5\rangle$	$\langle 0.9, 0.1, 0.1\rangle$
Pattern P_2	$\langle 0.8, 0.4, 0.1\rangle$	$\langle 0.5, 0.3, 0.5\rangle$	$\langle 0.7, 0.3, 0.4\rangle$
Pattern P_3	$\langle 0.5, 0.4, 0.3\rangle$	$\langle 0.6, 0.4, 0.6\rangle$	$\langle 0.6, 0.2, 0.1\rangle$
Test sample S	$\langle 0.6, 0.3, 0.2\rangle$	$\langle 0.7, 0.2, 0.6\rangle$	$\langle 0.8, 0.3, 0.1\rangle$

to Table 5, d_H (Broumi and Smarandache 2013a), d_{NH} and d_{NE} (Majumdar 2015), d_{S_3} (Shahzadi et al. 2017), d_{CS_1} and d_{CS_2} (Ye 2015a), d_{T_1} and d_{T_2} (Ye and Fu 2016) cannot be determined the classification results. And the results show that $d_g(P_1, S) \le d_g(P_3, S) \le d_g(P_2, S)$, i.e., the minimum distance is $d_g(P_1, S)$. Thus, it is shown that the test sample S belongs to P_1. In addition, it is easy to find that d_g with the highest degree of confidence than d_{S_1}, d_{S_2} and d_{S_4} (Shahzadi et al. 2017), d_V (Ye 2014b). For the results of Table 5, the distance d_g not only can give correct classification results but also obtain a much higher degree of confidence. Therefore, the distance d_g can be seen as a reasonable application to pattern recognition.

4.2.3 Applications for medical diagnosis

Example 4 (Samuel and Narmadhagnanam 2018; Shahzadi et al. 2017) Given three patients $P = \{$Ali, Hamza, Imran$\}$, five symptoms $S = \{$Temperature, Insulin, Blood pressure, Blood plates, Cough$\}$ and three diseases $D = \{$Diabetes, Dengue, Tuberculosis$\}$. The relation $P \rightarrow S$ and $D \rightarrow S$ are shown in Tables 6 and 7, respectively. What kind of disease could each patient suffer from?

The medical diagnostic results by the minimum distance d_g are shown in Table 8. Obviously, the first line shows Ali suffering from Dengue, the second line shows Hamza suffering from Diabetes , the third line shows Imran suffering from Tuberculosis. By comparing, our results are the same as Samuel and Narmadhagnanam (2018) and Shahzadi et al. (2017) in Table 9. Further, the diagnostic results based on the proposed distance are more believable than these diagnostic results in Samuel and Narmadhagnanam (2018) and Shahzadi et al. (2017).

Example 5 (Garg Nancy 2017; Ye and Fu 2016) Given four patients $P = \{p_1, p_2, p_3, p_4\}$, five symptoms $S = \{$Temperature, Headache, Stomach pain, Cough, Chest pain$\}$, five possible diseases $D = \{$Viral fever, Stenocardia, Typhoid, Gastritis, Malaria$\}$. The relation $P \rightarrow S$ and $D \rightarrow S$ are shown in Table 10 and Table 11, respectively. What kind of disease could each patient suffer from?

The medical diagnostic results by the minimum distance d_g are given in Table 12. Obviously, the first line shows p_1, p_3 suffering from Viral fever, the second line shows p_2 suffering from Malaria, the fourth line shows p_4 suffering from Stenocardia. In order to illustrate our results, we give a com-

Table 5 The results of distance measure application in pattern recognition

Distances	$d(P_1, S)$	$d(P_2, S)$	$d(P_3, S)$	Classification results	DoC
d_H(Broumi and Smarandache 2013a)	0.167	0.233	0.167	Cannot be determined	No
d_{NH} (Majumdar 2015)	0.150	0.200	0.150	Cannot be determined	No
d_{NE} (Majumdar 2015)	0.147	0.191	0.147	Cannot be determined	No
d_{S_1} (Shahzadi et al. 2017)	0.207	0.267	0.219	P_1	0.072
d_{S_2} (Shahzadi et al. 2017)	0.205	0.267	0.214	P_1	0.071
d_{S_3} (Shahzadi et al. 2017)	0.100	0.133	0.100	Cannot be determined	No
d_{S_4} (Shahzadi et al. 2017)	0.114	0.154	0.122	P_1	0.048
d_{CS_1} (Ye 2015a)	0.037	0.069	0.037	Cannot be determined	No
d_{CS_2} (Ye 2015a)	0.012	0.022	0.012	Cannot be determined	No
d_{T_1} (Ye and Fu 2016)	0.132	0.186	0.132	Cannot be determined	No
d_{T_2} (Ye and Fu 2016)	0.079	0.105	0.079	Cannot be determined	No
d_V (Ye 2014b)	0.030	0.020	0.035	P_1	0.015
d_g	0.090	0.175	0.142	P_1	0.137

Table 6 Characteristic valued between patients and symptoms

	Temperature	Insulin	Blood pressure	Blood plates	Cough
Ali	⟨0.8, 0.1, 0.1⟩	⟨0.2, 0.2, 0.6⟩	⟨0.4, 0.2, 0.4⟩	⟨0.8, 0.1, 0.1⟩	⟨0.3, 0.3, 0.4⟩
Hamza	⟨0.6, 0.2, 0.2⟩	⟨0.9, 0.0, 0.1⟩	⟨0.1, 0.1, 0.8⟩	⟨0.2, 0.1, 0.7⟩	⟨0.5, 0.1, 0.4⟩
Imran	⟨0.4, 0.2, 0.4⟩	⟨0.2, 0.1, 0.7⟩	⟨0.1, 0.2, 0.7⟩	⟨0.3, 0.1, 0.6⟩	⟨0.8, 0.0, 0.2⟩

Table 7 Characteristic valued between diseases and symptoms

	Temperature	Insulin	Blood pressure	Blood plates	Cough
Diabetes	⟨0.2, 0.0, 0.8⟩	⟨0.9, 0.0, 0.1⟩	⟨0.1, 0.1, 0.8⟩	⟨0.1, 0.1, 0.8⟩	⟨0.1, 0.1, 0.8⟩
Dengue	⟨0.9, 0.0, 0.1⟩	⟨0.0, 0.2, 0.8⟩	⟨0.8, 0.1, 0.1⟩	⟨0.9, 0.0, 0.1⟩	⟨0.1, 0.1, 0.8⟩
Tuberculosis	⟨0.6, 0.2, 0.2⟩	⟨0.0, 0.1, 0.9⟩	⟨0.4, 0.2, 0.4⟩	⟨0.0, 0.2, 0.8⟩	⟨0.9, 0.0, 0.1⟩

Table 8 Diagnosis results based on the proposed distance

	Diabetes	Dengue	Tuberculosis	Diagnosis result	DoC
Ali	0.296	0.179	0.246	Dengue	0.184
Hamza	0.185	0.281	0.238	Diabetes	0.149
Imran	0.244	0.258	0.174	Tuberculosis	0.154

Table 9 Comparative results

	Ali (DoC)	Hamza (DoC)	Imran (DoC)
The result in Samuel and Narmadhagnanam (2018)	Dengue (0.174)	Diabetes (0.145)	Tuberculosis (0.137)
The result in Shahzadi et al. (2017)	Dengue (0.167)	Diabetes (0.133)	Tuberculosis (0.146)
Our result	Dengue (0.184)	Diabetes (0.149)	Tuberculosis (0.154)

Table 10 Characteristic valued between patients and symptoms

	Temperature	Headache	Stomach pain	Cough	Chest pain
p_1	⟨0.8, 0.2, 0.1⟩	⟨0.6, 0.3, 0.1⟩	⟨0.2, 0.1, 0.8⟩	⟨0.6, 0.5, 0.1⟩	⟨0.1, 0.4, 0.6⟩
p_2	⟨0.6, 0.6, 0.1⟩	⟨0.1, 0.2, 0.6⟩	⟨0.3, 0.2, 0.8⟩	⟨0.6, 0.2, 0.3⟩	⟨0.2, 0.3, 0.7⟩
p_3	⟨0.3, 0.1, 0.2⟩	⟨0.3, 0.2, 0.2⟩	⟨0.7, 0.6, 0.7⟩	⟨0.3, 0.2, 0.2⟩	⟨0.4, 0.4, 0.3⟩
p_4	⟨0.2, 0.1, 0.7⟩	⟨0.2, 0.3, 0.7⟩	⟨0.2, 0.2, 0.7⟩	⟨0.2, 0.1, 0.8⟩	⟨0.8, 0.2, 0.1⟩

Table 11 Characteristic valued between diseases and symptoms

	Temperature	Headache	Stomach pain	Cough	Chest pain
Viral fever	$\langle 0.4, 0.6, 0.0\rangle$	$\langle 0.3, 0.2, 0.5\rangle$	$\langle 0.1, 0.3, 0.7\rangle$	$\langle 0.4, 0.3, 0.3\rangle$	$\langle 0.1, 0.2, 0.7\rangle$
Malaria	$\langle 0.7, 0.3, 0.0\rangle$	$\langle 0.2, 0.2, 0.6\rangle$	$\langle 0.0, 0.1, 0.9\rangle$	$\langle 0.7, 0.3, 0.0\rangle$	$\langle 0.1, 0.1, 0.8\rangle$
Typhoid	$\langle 0.3, 0.4, 0.3\rangle$	$\langle 0.6, 0.3, 0.1\rangle$	$\langle 0.2, 0.1, 0.7\rangle$	$\langle 0.2, 0.2, 0.6\rangle$	$\langle 0.1, 0.0, 0.9\rangle$
Gastritis	$\langle 0.1, 0.2, 0.7\rangle$	$\langle 0.2, 0.4, 0.4\rangle$	$\langle 0.8, 0.2, 0.0\rangle$	$\langle 0.2, 0.1, 0.7\rangle$	$\langle 0.2, 0.1, 0.7\rangle$
Stenocardia	$\langle 0.1, 0.1, 0.8\rangle$	$\langle 0.0, 0.2, 0.8\rangle$	$\langle 0.2, 0.0, 0.8\rangle$	$\langle 0.2, 0.0, 0.8\rangle$	$\langle 0.8, 0.1, 0.1\rangle$

Table 12 Diagnosis results based on the proposed distance

	Viral fever	Malaria	Typhoid	Gastritis	Stenocardia	Diagnosis results	DoC
p_1	0.177	0.191	0.206	0.262	0.264	Viral fever	0.215
p_2	0.150	0.135	0.174	0.239	0.238	Malaria	0.261
p_3	0.212	0.256	0.241	0.232	0.245	Viral fever	0.126
p_4	0.256	0.263	0.246	0.226	0.205	Stenocardia	0.171

Table 13 Comparative results

	$p_1(DoC)$	$p_2(DoC)$	$p_3(DoC)$	$p_4(DoC)$
The result in Garg Nancy (2017) ($t = 3$)	Viral fever (0.187)	Malaria (0.224)	Viral fever (0.107)	Stenocardia (0.153)
The result in Ye and Fu (2016)	Viral fever (0.169)	Malaria (0.194)	Viral fever (0.116)	Stenocardia (0.147)
Our result	Viral fever (0.215)	Malaria (0.267)	Viral fever (0.126)	Stenocardia (0.171)

parison with the exist study in Table 13. By comparing with Table 13, we can easily see that p_1, p_3 suffering from Viral fever, p_2 suffers from Malaria and p_4 suffering from Stenocardia the same as Garg Nancy (2017) and Ye and Fu (2016). Further, the diagnostic results based on the proposed distance are more believable than these diagnostic results in Garg Nancy (2017) and Ye and Fu (2016).

5 Conclusions

Although many distance between $SVNSs$ have been presented to handle uncertainty and inconsistency in various fields, most of them exist counter-intuitive cases. In this paper, we present a new distance between single valued neutrosophic sets, which overcomes the counter-intuitive cases of the existing distances. Furthermore, we applied the proposed distance in pattern recognition and medical diagnosis. Obviously, experimental results show, the proposed distance not only can be accurately classified test sample to pattern but also obtain a much higher degree of confidence. In the future, we will consider the distance based on different binary functions and matrix norms. And we will apply the distance measure to other fields, such as the multi-attribute decision making, clustering analysis.

Funding The funding was provided by the National Natural Science Foundation of China (Grant Nos. 12171445, 61773019).

Data availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Atanassov K (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20:87–96

Baker A (2001) Matrix groups: an introduction to Lie group theory. Springer, London

Borzooei RA, Farahani H, Moniri M (2017) Neutrosophic deductive filters on BL-algebras. J Intell Fuzzy Syst 26:2993–3004

Broumi S, Smarandache F (2013) Correlation coefficient of interval neutrosophic set. Appl Mech Mater 43:6511–517

Broumi S, Smarandache F (2013) Several similarity measures of neutrosophic sets. Neutrosoph Sets Syst 1:54–62

Broumi S, Smarandache F (2015) Extended Hausdorff distance and similarity measures for neutrosophic refined sets and their applications in medical diagnosis. J New Theory 7:64–78

Broumi S, Deli I, Smarandache F (2014) Distance and similarity measures of interval neutrosophic soft sets. In: Critical review, center for mathematics of uncertainty, vol 8. Creighton University, USA, pp 14–31

Choudhary B (1992) The elements of complex analysis. New Age International Ltd

Garg Nancy HD (2017) Some new biparametric distance measures on single valued neutrosophic sets with application to pattern recognition and medical diagnosis. Information 8:2078–2489

Huang H (2016) New distance measure of single-valued neutrosophic sets and its application. Int J Intell Syst 31:1021–1032

Karaaslan F (2017) Correlation coefficients of single valued neutrosophic refined soft sets and their applications in clustering analysis. Neural Comput Appl 28:2781–2793

Liu XZ, Yang M (2005) Matrix theory. Huazhong University of Science and Technoology Press

Liu PD, Li HG (2017) Multiple attribute decision-making method based on some normal neutrosophic Bonferroni mean operators. Neural Comput Appl 28:179–194

Liu PD (2016) The aggregation operators based on archimedean t-conorm and t-norm for single-valued neutrosophic numbers and their application to decision making. Int J Fuzzy Syst 18:1–15

Luo MX, Zhao RR (2018) A distance measure between intuitionistic fuzzy sets and its application in medical diagnosis. Artif Intell Med 89:34–39

Luo MX, Wu LX, Zhou KY, Zhang HR (2019) Multi-criteria decision making method based on the single valued neutrosophic sets. J Intell Fuzzy Syst 37:2403–2417

Ma YC, Zhang XH, Yang XF, Zhou X (2019) Generalized neutrosophic extended triplet group. Symmetry 11:327–337

Majumdar P (2015) Neutrosophic sets and its applications to decision making. Comput Intell Big Data Anal 19:97–115

Majumdar P, Samanta SK (2014) On similarity and entropy of netrosophic sets. J Intell Fuzzy Syst 26:1245–1252

Ren HP, Xiao SX, Zhou H (2019) A Chi-square distance-based similarity measure of single valued neutrosophic set and applications. Int J Comput Commun control 14:78–89

Samuel AE, Narmadhagnanam R (2018) cosine logarithmic distance of single valued neutrosophic sets in medical diagnoses. International Journal of Engineering Research and Applications 8:26–29

Samuel AE, Narmadhagnanam R (2018) Sine logarithmic distance of neutrosophic refined sets in medical diagnoses. J Glob Res Math Arch

Shahzadi G, Akram M, Saeid AB (2017) An application of single valued neutrosophic sets in medical diagnosis. Neutrosoph Sets Syst 18:80–88

Smarandache F (1999) A unifying field in logics. Neutrosophy: neutrosophic probability, set and logic. American Research Press, Rehoboth

Smarandache F. (1998) Neutrosophy, Neutrosophic probability, set, and logic. Analytic synthesis and synthetic analysis, Philosophy

Smarandache F (2010) Neutrosophic set-a generalization of the intuitionistic fuzzy set. J Defense Resour Manag 1:38–42

Wang H, Smarandache F, Zhang Y, Sunderraman R (2010) Single valued neutrosophic sets. Multispace Multistruct 4:410–413

Xu ZS, Zhang KY (2005) A brief tutorial on matrix theory. Science Press

Ye J (2013) Multicriteria decision-making method using the correlation coefficient under single valued neutrosophic environment. Int J Gen Syst 42:386–394

Ye J (2014) Similarity measure between interval neutrosophic sets and their applications in multiciteria decision making. J Intell Fuzzy Syst 26:165–172

Ye J (2015) Improved cosine similarity measures of simplified neutrosophic sets for medical diagnoses. Artif Intell Med 63:171–179

Ye J, Fu J (2016) Multi-period medical diagnosis method using a single valued neutrosophic similarity measure based on tangent function. Comput Methods Programs Biomed 123:142–149

Ye J (2014) Vector similarity measures of simplified neutrosophic sets and their application in multicriteria decision making. Int J Fuzzy Syst 16:204–211

Ye J (2017) Single-valued neutrosophic similarity measures based on cotangent function and their application in the fault diagnosis of steam turbine. Soft Comput 21:1–9

Ye J (2015) The generalized Dice measures for multiple attribute decision making under simplified neutrosophic environments. Intell Fuzzy Syst 31:663–671

Zhang XH, Bo CX, Smarandache F, Dai JH (2018) New inclusion relation of neutrosophic sets with applications and related lattice structure. Int J Mach Learn Cybern 9:1753–1763

Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353