Journal of Neutrosophic and Fuzzy Systems

♦ ISSN: 2771-6449 Vol: 11, No. 01, 2026 ♦ PP. 17–24

♦ Article DOI: https://doi.org/10.54216/JNFS.110103

FULL LENGTH ARTICLE

Hesitation-Gated Fuzzy–Neutrosophic Prototype Learning under Asymmetric Label Noise

Necati Olgun[1,*] **Ahmed Hatip**[1]

[1] Department of Mathematics, Gaziantep University, Gaziantep, Turkey

Emails: Olgun@gantep.edu.tr · kollnaar5@gmail.com

Received: September 30, 2025 Revised: November 27, 2025 Accepted: January 30, 2026 ⋆ Corresponding author

ABSTRACT

Label corruption is difficult for prototype classifiers because a mislabeled observation does two things at once: it perturbs the prototype v_{y_i} associated with the supplied label and obscures whether the observation is genuinely ambiguous or simply inconsistent with its assigned class. This paper develops an adaptive fuzzy–neutrosophic prototype learning algorithm that separates these effects. For each training observation, the membership vector $u_i = (u_{i1}, \ldots, u_{iK}) \in \Delta^{K-1}$ induced by the current prototypes is converted into the evidence state $z_i = (T_i, I_i, F_i) \in [0,1]^3$: truth is the membership assigned to the observed class, falsity is the strongest competing membership, and indeterminacy is the normalized membership entropy. These quantities drive three coupled mechanisms: a contradiction margin $c_i = [F_i - T_i]_+$ that attenuates unreliable labels, an entropy-dependent fuzzy exponent $m_i \in [m_{\min}, m_{\max}]$ that adapts membership weighting near class overlap, and a conservative soft-label correction activated only when $F_i > T_i$ and the hesitation I_i is sufficiently small. The resulting Adaptive Fuzzy–Neutrosophic Prototype Learning (AFNPL) algorithm remains a lightweight prototype method with linear cost in the number of observations, classes and features per iteration. A reproducible three-class study evaluates 0–40% cyclic asymmetric label corruption under low, medium and high class overlap. At medium overlap and 40% corruption, AFNPL obtains 91.54% test accuracy and 91.55% macro-F1, compared with 72.52%/72.37% for noisy class means and 80.23%/80.17% for trimmed class means. Its internal contradiction score also detects corrupted labels with mean AUC between 0.959 and 0.971 across the contaminated settings. The contribution is therefore not only a robust prototype update, but a fuzzy learning mechanism in which neutrosophic truth, indeterminacy and falsity have explicit algorithmic roles in label reliability and adaptive fuzzy weighting.

Keywords: Fuzzy prototype learning ▪ Single-valued neutrosophic information ▪ Label noise ▪ Adaptive fuzzy exponent ▪ Uncertainty ▪ Robust classification

1. INTRODUCTION

Fuzzy learning was introduced to avoid the brittle assumption that an observation must belong entirely to one class or cluster, replacing a hard indicator $\mathbf{1}\{y_i = k\}$ by graded memberships $u_{ik} \in [0,1]$. In fuzzy c-means (FCM), membership degrees distribute an observation over several prototypes rather than forcing a hard assignment [1]. Possibilistic clustering later relaxed the unit-sum membership constraint to reduce sensitivity to atypical points [2], while type-2 fuzzy sets provided an additional layer for representing uncertainty in the membership function itself [3]. These ideas remain relevant whenever the data geometry is uncertain, boundaries overlap, or supervision cannot be treated as completely reliable; in such cases, a membership vector with $\max_k u_{ik} < 1$ conveys information that a hard label suppresses.

Label noise introduces a different uncertainty source. A training point can be geometrically typical of a latent class $y_i^\star$ while carrying an observed label $y_i \neq y_i^\star$. In high-capacity learning systems, corrupted labels can be memorized, motivating loss correction, sample selection, co-teaching and noise-aware objectives [4, 5, 6, 7, 8, 9]. Prototype methods do not memorize in the same way as deep networks, but they suffer a direct geometric distortion: mislabeled observations displace class representatives, and the resulting prototypes then produce distorted memberships for otherwise clean points. A method that merely assigns fuzzy memberships is therefore not automatically a method for reasoning about label reliability: the condition $u_{iy_i} < \max_{k\neq y_i} u_{ik}$ must be interpreted differently from ordinary boundary ambiguity.

Neutrosophic representations offer a useful vocabulary for this problem. Single-valued neutrosophic information describes truth, indeterminacy and falsity independently on $[0,1]$ [10]. Recent research has proposed similarity, entropy and distance measures [11, 12, 13, 14], new aggregation and score functions [15, 16, 17], dynamic formulations [18], hybrid decision models [19, 20], and neutrosophic cognitive-map applications [21]. Most of this literature treats a neutrosophic triple $z = (T,I,F)$ as the input to a decision or aggregation procedure. The present work uses it differently: $z_i = (T_i,I_i,F_i)$ is generated internally by the fuzzy learner and controls how strongly each training label is trusted.

The central idea is to distinguish *ambiguity* from *contradiction*. Suppose the current fuzzy memberships of observation i are $u_{i1},\dots,u_{iK}$ and the observed label is y_i. A point near a class boundary may have high normalized membership entropy $I_i \approx 1$, and this should be interpreted as indeterminacy rather than as evidence that its label is wrong. By contrast, if $F_i - T_i > 0$ while the membership vector is sharp ($I_i \ll 1$), the observation provides stronger evidence of contradiction. The proposed method encodes these two situations separately and uses them to regulate both the fuzzy membership-power exponent and the prototype update.

The paper makes four contributions, all expressed through the coupled state $(u_i, z_i, r_i, m_i, \widetilde{y}_i)$ updated at each iteration. First, it introduces a sample-wise fuzzy–neutrosophic evidence state derived directly from prototype memberships, without requiring externally supplied neutrosophic labels. Second, it proposes a hesitation-adaptive membership-power exponent and a contradiction-dependent reliability weight, creating a fuzzy learning rule whose prototype weighting varies with local ambiguity. Third, it develops a conservative soft-label correction that moves mass away from a contradicted observed label only when the competing fuzzy evidence is sufficiently decisive. Fourth, it evaluates classification robustness, prototype displacement, noise-detection ability and convergence behavior under controlled asymmetric corruption. The emphasis is therefore algorithmic learning and classification: neutrosophic evidence is generated within the learning loop rather than supplied as a fixed decision descriptor.

The remainder of the paper is organized as follows, with the mathematical development proceeding from u_{ik} to $z_i = (T_i,I_i,F_i)$ and finally to the prototype update v_k^{new}. Section 2 reviews fuzzy clustering, neutrosophic uncertainty and noisy-label learning. Section 3 defines the fuzzy membership and evidence quantities used by the method. Section 4 presents AFNPL and its main properties. Section 5 describes the experimental protocol. Section 6 reports classification, prototype and diagnostic results with graphical analysis. Section 7 discusses the scope and limitations, and Section 8 concludes the study.

2. RELATED WORK

2.1 Fuzzy and possibilistic prototype learning

The FCM objective assigns memberships u_{ik} and prototypes v_k by minimizing a membership-weighted within-cluster distortion, conventionally of the form $\sum_i \sum_k u_{ik}^m \|x_i - v_k\|_2^2$ [1]. The exponent $m > 1$ controls fuzziness: values close to one approach hard assignments, whereas larger values flatten membership differences. The same global m is normally applied to all observations, even though geometric uncertainty can vary considerably across the feature space. Possibilistic c-means removes the probabilistic partition constraint and interprets membership more as typicality [2]. Intuitionistic fuzzy clustering adds a hesitation component and has motivated algorithms that distinguish membership from non-membership [22, 23]. Type-2 fuzzy sets similarly emphasize uncertainty about membership itself [3].

The proposed method shares the prototype efficiency of FCM but differs in two respects: the global scalar m is replaced by m_i, and the update weight becomes a function of $z_i = (T_i,I_i,F_i)$. The membership-power exponent is observation-specific rather than global, and its value is not tuned from distance alone: it is driven by a normalized uncertainty measure of the full membership vector. More importantly, the observed class label participates in a neutrosophic evidence calculation, so contradictory supervision can be attenuated instead of being treated as unquestioned ground truth.

2.2 Neutrosophic information in learning and decision systems

Fuzzy sets [24] and intuitionistic fuzzy sets [22] progressively expanded the representation of partial membership and hesitation. Single-valued neutrosophic sets provide three independent coordinates $(T,I,F) \in [0,1]^3$ for truth, indeterminacy and falsity [10]. Subsequent work has investigated entropy and similarity [11, 12], distance measures for pattern recognition [13, 14], algebraic operators and scoring mechanisms [15, 16, 17], dynamic formulations and cognitive-map applications [18, 21]. Hybrid MCDM methods continue to extend the use of neutrosophic information in structured decision problems [19, 20].

A common pattern is to begin with a neutrosophic description $z = (T,I,F)$ supplied by a decision maker or derived from an application-specific mapping, then define operations on z. AFNPL instead constructs (T_i,I_i,F_i) from fuzzy prototype memberships during learning. Thus the neutrosophic state is transient and endogenous: $z_i^{(t+1)} \neq z_i^{(t)}$ whenever the prototype update changes the induced membership geometry. This allows truth, falsity and indeterminacy to act as control variables in the learning algorithm rather than only as values to be aggregated after elicitation.

2.3 Learning with corrupted labels

Noisy-label learning has produced several broad strategies. Loss-correction approaches estimate or use corruption structure [4]; robust losses modify how difficult or contradictory observations contribute to the objective [5, 8]; co-teaching uses peer networks to exchange small-loss examples [6]; probabilistic methods infer clean/noisy states [7]; and sample-selection frameworks analyze when filtering improves generalization [9]. These methods demonstrate that confidence alone is not sufficient: uncertainty must be interpreted relative to the observed label y_i and the evolving model state, rather than through $\max_k u_{ik}$ alone.

AFNPL transfers this principle to a lightweight fuzzy prototype setting by mapping $(u_i, y_i) \mapsto (T_i, I_i, F_i) \mapsto (r_i, m_i, \widetilde{y}_i)$. It does not estimate a full transition matrix and does not discard observations. Instead, it continuously decomposes the compatibility of an observed label into truth, falsity and indeterminacy, then uses those values to set the effective fuzziness and contribution of the sample.

3. MATHEMATICAL PRELIMINARIES

3.1 Supervised prototype setting

Let the training set be

$$\mathscr{D} = \{(x_i, y_i)\}_{i=1}^{n}, \qquad x_i \in \mathbb{R}^d, \quad y_i \in \{1, \dots, K\}, \tag{1}$$

where y_i is the observed label and may differ from the unobserved true class. Let $V = (v_1, \dots, v_K)$ be class prototypes with $v_k \in \mathbb{R}^d$. Define squared Euclidean dissimilarity

$$d_{ik} = \|x_i - v_k\|_2^2. \tag{2}$$

Rather than the classical inverse-distance FCM membership, we use a temperature-scaled soft prototype membership

$$u_{ik} = \frac{\exp[-d_{ik}/(\lambda s)]}{\sum_{j=1}^{K} \exp[-d_{ij}/(\lambda s)]}, \tag{3}$$

where $s > 0$ is the median nearest-prototype squared distance in the current iteration and $\lambda > 0$ is fixed. Hence $u_{ik} \in (0,1)$ and $\sum_k u_{ik} = 1$.

3.2 Fuzzy–neutrosophic evidence state

For the observed class y_i, define

$$T_i = u_{iy_i}, \tag{4}$$

$$F_i = \max_{k \neq y_i} u_{ik}, \tag{5}$$

$$I_i = -\frac{1}{\log K} \sum_{k=1}^{K} u_{ik} \log u_{ik}, \tag{6}$$

The triple $z_i = (T_i, I_i, F_i)$ lies in $[0,1]^3$; as usual, $0 \log 0$ is interpreted as 0 by continuity. No condition $T_i + I_i + F_i = 1$ is imposed, consistent with the single-valued neutrosophic representation [10]. Here the three components have operational meanings: T_i is geometric support for the supplied class, F_i is the strongest geometric opposition, and I_i is membership hesitation.

Define the positive contradiction margin

$$c_i = [F_i - T_i]_+ = \max(0, F_i - T_i). \tag{7}$$

A high c_i is not sufficient by itself to conclude that a label is corrupted, because the membership vector may be diffuse, i.e., I_i may simultaneously be close to 1. We therefore define the diagnostic contradiction score

$$q_i = c_i(1 - I_i), \tag{8}$$

which is large only when an alternative class dominates and the fuzzy assignment is comparatively decisive.

Proposition 1 (Bounded evidence)**.** *For every observation,* $T_i, F_i, I_i, c_i, q_i \in [0,1]$.

Proof. Memberships are in $[0,1]$, so $T_i, F_i \in [0,1]$ and therefore $c_i \in [0,1]$. The Shannon entropy of a K-category probability vector lies in $[0, \log K]$, giving $I_i \in [0,1]$. Finally, q_i is the product of two values in $[0,1]$. □

4. PROPOSED METHOD

4.1 Hesitation-adaptive fuzzy exponent

A fixed fuzzy exponent m applies the same transformation $u_{ik} \mapsto u_{ik}^m$ to a central point and a boundary point. AFNPL instead sets

$$m_i = m_{\min} + (m_{\max} - m_{\min}) I_i, \tag{9}$$

with $1 < m_{\min} \leq m_{\max}$. Low-entropy observations therefore use a smaller membership-power exponent in the prototype update, while uncertain boundary observations use a larger one. The preliminary memberships in (3) are not recomputed with m_i; instead, m_i controls how strongly those memberships enter the effective weights in (13). This is the principal fuzzy contribution of the method: the strength of fuzzy membership weighting becomes a sample-specific response to the current hesitation state rather than a single global exponent.

4.2 Contradiction gate and soft-label repair

The reliability gate is

$$r_i = e^{-\kappa c_i}\{0.55 + 0.45(1 - I_i)\}, \qquad \kappa \geq 0. \tag{10}$$

Thus contradiction decreases reliability exponentially, $\partial r_i / \partial c_i = -\kappa r_i < 0$, while high indeterminacy also reduces the gate but does not force it to zero. This distinction prevents ambiguous boundary points from being treated exactly like confidently contradicted labels.

Let e_{y_i} denote the one-hot vector for the observed label. The correction strength is

$$\alpha_i = \min\{0.95,\ 2.2\gamma c_i(1 - I_i)\}, \qquad 0 \leq \gamma \leq 1, \tag{11}$$

and the repaired soft target is

$$\widetilde{y}_i = (1 - \alpha_i) e_{y_i} + \alpha_i u_i. \tag{12}$$

When the observed class remains at least as plausible as every alternative, $F_i \leq T_i$ implies $c_i = \alpha_i = 0$ and hence no

Algorithm 1 Adaptive Fuzzy–Neutrosophic Prototype Learning (AFNPL)

Require: X, observed labels y, K, $m_{\min}, m_{\max}, \kappa, \gamma, \delta, \lambda$
Initialize each v_k by the mean of observations carrying label k
repeat
  Compute d_{ik} and scale s; obtain memberships u_{ik} from (3)
  Compute T_i, F_i, I_i from (4)–(6)
  Set $c_i = [F_i - T_i]_+$ and m_i from (9)
  Compute r_i from (10) and α_i from (11)
  Form repaired targets $\widetilde{y}_i$ using (12)
  Compute w_{ik} from (13)
  Update all prototypes with (14)
until prototype change is below tolerance or the iteration limit is reached
return prototypes V and diagnostic scores $q_i = c_i(1 - I_i)$

correction occurs. A correction is activated only when $F_i > T_i$, and it is weakened when membership entropy is large.

Proposition 2 (Conservative correction)**.** *If $F_i \le T_i$, then $\widetilde{y}_i = e_{y_i}$. If $F_i > T_i$, the total mass moved away from the observed one-hot label is at most* 0.95 *and is nonincreasing in I_i for fixed T_i, F_i.*

Proof. The first statement follows from $c_i = 0$. For the second, $\alpha_i \le 0.95$ by construction. Before clipping, α_i is proportional to $(1 - I_i)$; clipping preserves nonincrease with respect to I_i. □

4.3 Prototype update

For class k, AFNPL assigns observation i the effective weight

$$w_{ik} = r_i \widetilde{y}_{ik} u_{ik}^{m_i} + \delta(1 - r_i) u_{ik}^2, \tag{13}$$

where $0 < \delta \ll 1$ retains a small geometry-only contribution when the supplied label is strongly distrusted. The prototype update is

$$v_k^{\text{new}} = \frac{\sum_{i=1}^n w_{ik} x_i}{\sum_{i=1}^n w_{ik} + \varepsilon}. \tag{14}$$

The first term in (13), $r_i \widetilde{y}_{ik} u_{ik}^{m_i}$, preserves supervised information when it is compatible with the prototype geometry. The second term avoids complete sample deletion and permits a contradicted point to support the class toward which its fuzzy membership already points.

4.4 Interpretation and computational cost

The method can be read as a closed feedback loop, $V^{(t)} \to U^{(t)} \to Z^{(t)} \to W^{(t)} \to V^{(t+1)}$. Prototypes generate fuzzy memberships; memberships generate neutrosophic evidence; evidence controls fuzziness, label reliability and correction; and those quantities generate the next prototypes. Unlike hard filtering, all observations remain available. Unlike an ordinary fuzzy prototype learner, however, an observed label is not granted the same authority at every iteration because its effective contribution is modulated by $r_i \in (0, 1]$.

For n observations, K classes and d features, distance calculation and prototype accumulation both require $O(nKd)$ arithmetic. All evidence and gating operations are $O(nK)$ or less. Hence one AFNPL iteration is $O(nKd)$ in time and $O(nK + Kd)$ in working memory if the membership matrix is retained. Thus the method retains linear per-iteration scaling in n, K and d, while its learned state consists only of the K prototypes in addition to the sample-wise quantities computed during each iteration.

5. EXPERIMENTAL SETUP

5.1 Controlled data model

A controlled experiment is used because the true labels $y_i^\star$ and uncontaminated class centers μ_k must be known in order to measure both classification and prototype displacement. Three bivariate Gaussian classes have means

$$\mu_1 = (-1.65, -0.55), \quad \mu_2 = (1.55, -0.50), \quad \mu_3 = (0, 1.70). \tag{15}$$

and common base covariance

$$\Sigma = \begin{bmatrix} 0.78 & 0.22 \\ 0.22 & 0.68 \end{bmatrix}. \tag{16}$$

The covariance is multiplied by a^2 with overlap multiplier $a \in \{0.78, 1.00, 1.25\}$, denoted low, medium and high overlap. Each training panel contains 90 observations per class; each independent test panel contains 500 per class.

Training labels are corrupted asymmetrically. Independently for each training observation, corruption occurs with probability $\eta \in \{0, 0.1, 0.2, 0.3, 0.4\}$. A corrupted class k is changed cyclically to $k + 1$ modulo three. This construction is intentionally not symmetric: with the corruption map $k \mapsto (k \bmod 3) + 1$, cyclic contamination induces directional prototype bias and therefore tests whether reliability weighting can resist systematic displacement.

For every overlap–corruption combination, 120 independent replications are generated. All pseudo-random seeds are fixed in the supplied code. Reported metrics are test accuracy, macro-F1 and mean prototype displacement

$$D_V = \frac{1}{K} \sum_{k=1}^K \|v_k - \mu_k\|_2. \tag{17}$$

5.2 Comparators and parameters

Four prototype estimators $\widehat{V}^{(M)}$, $M \in \{\text{NCM}, \text{Trimmed}, \text{Fuzzy-PL}, \text{AFNPL}\}$, are compared. **NCM** uses the ordinary mean of all observations carrying each noisy label. **Trimmed-NCM** removes the farthest 20% of observations within each noisy label group before recomputing the class mean. **Fuzzy-PL** is a supervised fuzzy prototype learner using a fixed exponent $m = 2$ and the observed one-hot label in its update. **AFNPL** uses Algorithm 1. The fixed AFNPL settings are $m_{\min} = 1.55$, $m_{\max} = 2.75$, $\kappa = 5$, $\gamma = 0.90$, $\delta = 0.08$ and $\lambda = 0.8$. These values are held constant across all corruption and overlap conditions.

Noise-detection ability is evaluated only for AFNPL, because its contradiction score $q_i = c_i(1 - I_i)$ in (8) is an explicit diagnostic output whose ranking can be compared with the binary corruption mask. The area under the ROC curve (AUC)

Table 1. Medium-overlap classification results. Values are means over 120 independent replications.

Noise	Method	Accuracy	Macro-F1
0.0	NCM	0.9184	0.9185
	Trimmed-NCM	0.9178	0.9179
	Fuzzy-PL	0.9176	0.9177
	AFNPL	0.9176	0.9177
0.2	NCM	0.8980	0.8980
	Trimmed-NCM	0.9141	0.9142
	Fuzzy-PL	0.9165	0.9166
	AFNPL	**0.9167**	**0.9167**
0.3	NCM	0.8441	0.8439
	Trimmed-NCM	0.8931	0.8931
	Fuzzy-PL	0.9156	0.9157
	AFNPL	**0.9160**	**0.9161**
0.4	NCM	0.7252	0.7237
	Trimmed-NCM	0.8023	0.8017
	Fuzzy-PL	0.9150	0.9151
	AFNPL	**0.9154**	**0.9155**

is calculated against the known corruption mask. An additional ablation study uses medium overlap and $\eta = 0.3$ with 200 independent replications. The complete source code, replicate-level CSV files and figure-generation procedure are included in the submission package.

6. RESULTS

6.1 Classification under asymmetric corruption

Table 1 summarizes the medium-overlap experiment. In the clean setting, all methods are within about one tenth of a percentage point: NCM gives 91.84% accuracy and AFNPL gives 91.76%. The robust mechanism therefore does not create a meaningful clean-data penalty: the absolute accuracy difference is $|0.9184 - 0.9176| = 8 \times 10^{-4}$. The behavior changes as corruption becomes directional. At 30% corruption, NCM falls to 84.41% and Trimmed-NCM to 89.31%, whereas Fuzzy-PL and AFNPL remain at 91.56% and 91.60%. At 40%, AFNPL reaches 91.54% accuracy and 91.55% macro-F1, while NCM reaches 72.52%/72.37% and Trimmed-NCM 80.23%/80.17%.

Figure 1 shows the complete corruption trajectory as a function of $\eta \in [0, 0.4]$. The principal separation is between methods that continue to trust the noisy label group geometrically and methods that use fuzzy compatibility. AFNPL is not designed to manufacture a large advantage over a strong fuzzy prototype baseline when that baseline is already robust; rather, its contribution is to retain comparable predictive performance while producing the interpretable sample-level quantities (T_i, I_i, F_i, q_i).

6.2 Effect of class overlap

Figure 2 reports AFNPL macro-F1 across the three overlap regimes. Increasing geometric overlap lowers the attainable classification rate, as expected, but the curves remain nearly flat with respect to label corruption. At 40% corruption, AFNPL obtains macro-F1 values 0.9692, 0.9155 and 0.8476

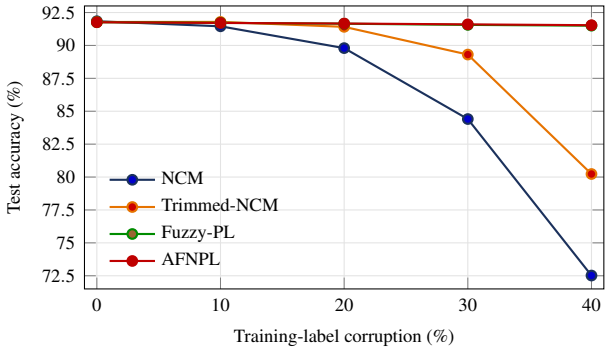

Figure 1. Test accuracy versus asymmetric training-label corruption at medium class overlap.

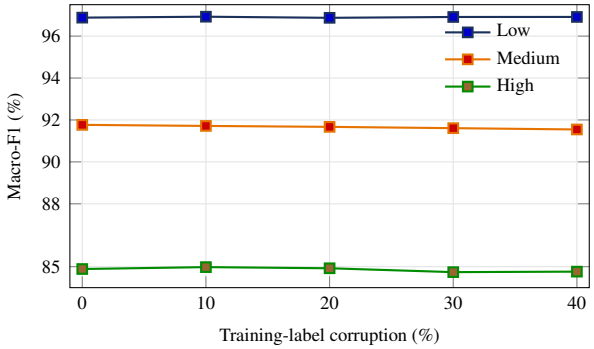

Figure 2. AFNPL macro-F1 under low, medium and high geometric overlap.

for low, medium and high overlap, respectively. The corresponding NCM values are 0.7689, 0.7237 and 0.6774. This indicates that the dominant performance loss for AFNPL comes from intrinsic class overlap a rather than from the imposed label corruption level η over the investigated grid.

Prototype displacement $D_V = K^{-1} \sum_k \|v_k - \mu_k\|_2$ provides a complementary view. Under medium overlap and 40% corruption, mean displacement is 1.1543 for NCM, 0.9760 for Trimmed-NCM, 0.2351 for Fuzzy-PL and 0.2239 for AFNPL. At low overlap the same corruption level gives 1.1578, 0.9317, 0.1503 and 0.1220, respectively. In the high-overlap regime, AFNPL does not uniformly minimize displacement: its mean value 0.3743 is above the Fuzzy-PL value 0.3445 even though its classification accuracy is slightly higher (0.8475 versus 0.8463). This is useful evidence against overinterpreting prototype distance as a universal surrogate for classification quality: in this setting, $D_V^{\text{AFNPL}} > D_V^{\text{Fuzzy-PL}}$ can coexist with $\text{Acc}_{\text{AFNPL}} > \text{Acc}_{\text{Fuzzy-PL}}$.

6.3 Corrupted-label detection from neutrosophic evidence

The contradiction score $q_i = c_i(1 - I_i)$ is not used as a hard rejection rule, but it can be evaluated as a corruption indicator. Figure 3 shows mean AUC across the contaminated medium-overlap settings. Mean AUC is 0.9672, 0.9709, 0.9682 and 0.9592 at corruption levels 10%, 20%, 30% and 40%, respectively. Thus the same fuzzy–neutrosophic quantities used to control the update also provide a useful diagnostic ranking q_i of suspicious labels. Variability is modest at 10–30% corruption (standard deviations 0.0263, 0.0178 and 0.0151), whereas the 40% setting has a larger standard deviation of 0.0853 because of one rare low-AUC replication; its median AUC remains 0.9697.

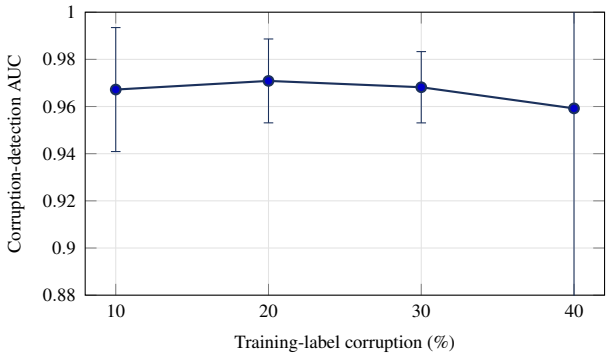

Figure 3. AUC of the AFNPL contradiction score for identifying corrupted training labels. Error bars show one standard deviation over 120 replications.

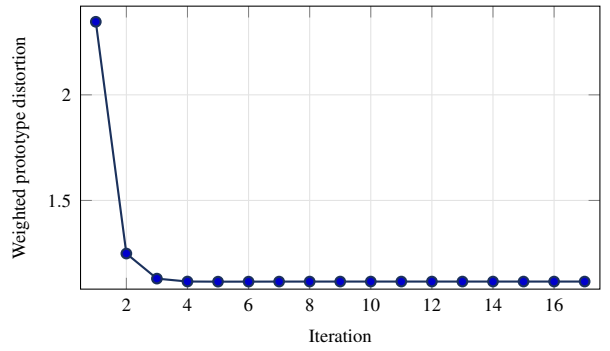

Figure 4. Representative AFNPL weighted-distortion trace at medium overlap and 30% asymmetric label corruption.

This diagnostic property illustrates why I_i is separated from F_i: for fixed c_i, $q_i = c_i(1-I_i)$ decreases linearly as hesitation increases. A boundary point can have moderate or even high competing membership, but its high entropy reduces $1-I_i$ and therefore suppresses the corruption score. A sharply classified point whose supplied label disagrees with the strongest prototype receives a much larger score.

6.4 Ablation and iterative behavior

The ablation study at medium overlap and 30% corruption is intentionally reported even though differences are small. Full AFNPL obtains mean accuracy 0.91694. Fixing the fuzzy exponent at $m = 2$ gives 0.91687, removing the contradiction gate gives 0.91687, and disabling label correction gives 0.91693. These results show that the classification gain is not attributable to one dominant switch in this controlled setting, since all ablated accuracies differ from the full value by less than 10^{-4}. The components mainly affect how evidence is interpreted and how suspicious observations are weighted; the underlying fuzzy prototype geometry is already a strong source of robustness.

Figure 4 shows the weighted prototype distortion for a representative 30%-corrupted panel. It drops from 2.3460 at the first recorded iteration to 1.1161 by iteration four and stabilizes near 1.1159. A very small rebound appears after the fifth iteration because $w_{ik}^{(t)}$ and $m_i^{(t)}$ change with $V^{(t)}$; the plotted quantity is therefore a diagnostic distortion, not a fixed objective $J(V)$ with a claimed monotonicity theorem.

7. DISCUSSION

The experiments support three conclusions across the investigated grid $(a,\eta) \in \{0.78, 1.00, 1.25\} \times \{0, 0.1, 0.2, 0.3, 0.4\}$. First, structured asymmetric label corruption can substantially damage ordinary prototypes. When corrupted classes preferentially move in one direction, class means are systematically displaced rather than receiving only nondirectional perturbations. Trimming helps but is insufficient when the contaminated observations are not extreme in feature space. Fuzzy compatibility is particularly effective because a point carrying the wrong label can still satisfy $u_{iy_i} \ll \max_{k \neq y_i} u_{ik}$ even when it is not a Euclidean outlier.

Second, the main value of the neutrosophic layer is not a dramatic increase over every fuzzy baseline. Fuzzy-PL is already highly robust in the present geometry, and AFNPL improves its classification results only slightly. The added contribution is representational and algorithmic: AFNPL distinguishes observed-label support (T_i), boundary hesitation (I_i) and opposing-class support (F_i), then gives each quantity a direct role in the learning rule. The resulting contradiction score $q_i = c_i(1-I_i)$ achieves high corruption-detection AUC without training a separate detector. This distinction is important because an incremental numerical gain alone would not justify introducing additional uncertainty machinery.

Third, the adaptive fuzzy exponent should be interpreted as local regularization of prototype weighting. Equation (9) increases the power applied to preliminary memberships when the prototype geometry is uncertain and reduces it when one prototype dominates. This differs from selecting a single global FCM exponent. It also differs from type-2 fuzzy modeling, where uncertainty is represented in the membership function itself [3]. AFNPL remains a type-1 membership model, but the membership-power exponent becomes the endogenous function $m_i = m_{\min} + (m_{\max} - m_{\min})I_i$.

Several limitations define the appropriate scope of the claims, particularly because the present design fixes $d = 2$, $K = 3$, and $\eta \leq 0.4$. The study uses Gaussian synthetic data in two dimensions so that corruption, overlap and true prototypes are exactly controlled. The resulting evidence is strong for mechanism validation but is not a substitute for benchmark studies on high-dimensional empirical datasets. The cyclic noise model is deliberately structured; instance-dependent and open-set label noise may behave differently. Euclidean prototypes also assume that classes can be represented reasonably by centers in the chosen feature space. Kernelized, metric-learning or multi-prototype extensions would be needed for strongly nonconvex classes. Finally, the current parameter constants were fixed across experiments but not estimated from theory. A future version could connect κ, $m_{\min}$ and $m_{\max}$ to measurable separation statistics such as inter-prototype distances $\|v_k - v_\ell\|_2$ or membership entropy summaries.

These limitations suggest several extensions within the scope of fuzzy and neutrosophic systems, replacing the present map $u_i \mapsto (T_i, I_i, F_i)$ by richer uncertainty constructions where appropriate. The evidence triple could be built from interval type-2 memberships, graph-based local memberships, or multiple prototypes per class. Dynamic evidence could be introduced for streaming settings, concept drift or changing

annotator reliability, paralleling recent interest in dynamic neutrosophic models [18]. The diagnostic score could also trigger relabeling when $q_i > \tau_q$ for an application-defined threshold τ_q, rather than automatically changing labels, which would be attractive in human-in-the-loop classification.

8. CONCLUSION

This paper introduced AFNPL, a fuzzy–neutrosophic prototype learning algorithm for classification when observed labels y_i may differ systematically from latent classes $y_i^\star$. The method generates a neutrosophic evidence state from current fuzzy memberships: membership in the observed class is treated as truth support, the strongest competing membership as falsity support, and normalized membership entropy as indeterminacy. These three values are then used to compute m_i, r_i and $\widetilde{y}_i$, respectively governing adaptive membership power, contradictory supervision and conservative soft-label correction.

The controlled experiments show that the method preserves the clean-data behavior of ordinary prototype learners while resisting cyclic asymmetric label corruption. At medium overlap and 40% corruption, AFNPL maintains 91.54% accuracy compared with 72.52% for noisy class means and 80.23% for a trimmed estimator. Its internal contradiction score also provides strong discrimination between corrupted and clean labels, with mean AUC close to or above 0.96 across all contaminated settings. The results therefore support the use of the endogenous state $z_i = (T_i, I_i, F_i)$ as an active component of fuzzy learning rather than only as a representation passed to a downstream decision rule. Future work should test the mechanism on empirical high-dimensional benchmarks, extend it to non-Euclidean prototypes and develop parameter-selection rules tied to observed class geometry.

REFERENCES

[1] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

[2] R. Krishnapuram and J. M. Keller, "A possibilistic approach to clustering," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 2, pp. 98–110, 1993.

[3] J. M. Mendel and R. I. B. John, "Type-2 fuzzy sets made simple," *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 117–127, 2002.

[4] G. Patrini, A. Rozza, A. K. Menon, R. Nock, and L. Qu, "Making deep neural networks robust to label noise: A loss correction approach," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2233–2241, 2017.

[5] Z. Zhang and M. R. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," *Advances in Neural Information Processing Systems*, vol. 31, pp. 8778–8788, 2018.

[6] B. Han, Q. Yao, X. Yu, G. Niu, M. Xu, W. Hu, I. Tsang, and M. Sugiyama, "Co-teaching: Robust training of deep neural networks with extremely noisy labels," *Advances in Neural Information Processing Systems*, vol. 31, pp. 8527–8537, 2018.

[7] E. Arazo, D. Ortego, P. Albert, N. O'Connor, and K. McGuinness, "Unsupervised label noise modeling and loss correction," *Proceedings of the 36th International Conference on Machine Learning, Proceedings of Machine Learning Research*, vol. 97, pp. 312–321, 2019.

[8] Y. Wang, X. Ma, Z. Chen, Y. Luo, J. Yi, and J. Bailey, "Symmetric cross entropy for robust learning with noisy labels," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 322–330, 2019.

[9] G. Jiang, W. Wang, Y. Qian, and J. Liang, "A unified sample selection framework for output noise filtering: An error-bound perspective," *Journal of Machine Learning Research*, vol. 22, no. 18, pp. 1–66, 2021.

[10] H. Wang, F. Smarandache, Y. Zhang, and R. Sunderraman, "Single valued neutrosophic sets," *Multispace and Multistructure*, vol. 4, pp. 410–413, 2010.

[11] K. Qin and L. Wang, "New similarity and entropy measures of single-valued neutrosophic sets with applications in multi-attribute decision making," *Soft Computing*, vol. 24, pp. 16 165–16 176, 2020.

[12] J. S. Chai, G. Selvachandran, F. Smarandache, V. C. Gerogiannis, L. H. Son, Q.-T. Bui, and B. Vo, "New similarity measures for single-valued neutrosophic sets with applications in pattern recognition and medical diagnosis problems," *Complex & Intelligent Systems*, vol. 7, pp. 703–723, 2021.

[13] M. Luo, G. Zhang, and L. Wu, "A novel distance between single valued neutrosophic sets and its application in pattern recognition," *Soft Computing*, vol. 26, pp. 11 129–11 137, 2022.

[14] Y. Zeng, H. Ren, T. Yang, S. Xiao, and N. Xiong, "A novel similarity measure of single-valued neutrosophic sets based on modified Manhattan distance and its applications," *Electronics*, vol. 11, no. 6, p. 941, 2022.

[15] H. Garg, "A new exponential-logarithm-based single-valued neutrosophic set and their applications," *Expert Systems with Applications*, vol. 238, p. 121854, 2024.

[16] F. Wang, "Novel score function and standard coefficient-based single-valued neutrosophic MCDM for live streaming sales," *Information Sciences*, vol. 654, p. 119836, 2024.

[17] T. Senapati, "An Aczel-Alsina aggregation-based outranking method for multiple attribute decision-making using single-valued neutrosophic numbers," *Complex & Intelligent Systems*, vol. 10, pp. 1185–1199, 2024.

[18] J. Qiu, L. Jiang, H. Fan, P. Li, and C. You, "Dynamic nonlinear simplified neutrosophic sets for multiple-attribute group decision making," *Heliyon*, vol. 10, no. 5, p. e27493, 2024.

[19] A. Paraskevas and M. Madas, "Selection of academic staff based on a hybrid multi-criteria decision method under neutrosophic environment," *Operations Research Forum*, vol. 5, p. 23, 2024.

[20] M. A. Dasan, E. Bementa, M. Aslam, and V. F. Little Flower, "Multi-attribute decision-making problem in career determination using single-valued neutrosophic distance measure," *Complex & Intelligent Systems*, vol. 10, pp. 5411–5425, 2024.

[21] J. L. R. Villafuerte, S. B. E. Pijal, M. I. M. Verdezoto, L. Cevallos-Torres, and S. Mirzaliev, "Enhancing decision-making in uncertain environments: The role of neutrosophic cognitive maps in analyzing complex systems," *Journal of Intelligent Systems and Internet of Things*, vol. 13, no. 1, pp. 287–292, 2024.

[22] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.

[23] Z. Xu and J. Wu, "Intuitionistic fuzzy c-means clustering algorithms," *Journal of Systems Engineering and Electronics*, vol. 21, no. 4, pp. 580–590, 2010.

[24] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.