October 2023

Guiding Image Classifier Using a Neuro-fuzzy Controller

Lenord Melvix Joseph Stephen Max

Guiding Image Classifier Using a Neuro-fuzzy Controller

ABSTRACT

This disclosure describes a neuro-fuzzy controller that can be utilized to guide image classifier networks for classification of subjective attributes. Per techniques of this disclosure, linguistic expert rules for memberships of an image to various output categories of the subjective attribute(s) are framed and the classification is analyzed as a fuzzy system. Fuzzy rules and fuzzy inference output from this system are used to guide a neural network to effectively incorporate the expert rules. Specific loss functions are utilized to guide the image classifier. A *fuzzy-rule contradiction loss* is utilized to capture a weighted deviation of image classifier prediction from expert rules. A *fuzzy inference loss* is utilized to capture overall deviation from fuzzy inference output. Utilization of the neuro-fuzzy controller can enable image classifier models to classify images according to subjective attributes, e.g., to provide accurate labels for family friendliness of a restaurant based on images of the restaurant.

KEYWORDS

- Neuro-fuzzy controller
- Image classifier
- Crisp signal
- Fuzzy inference
- Fuzzy logic
- Subject attribute
- Image understanding
- Model bias

BACKGROUND

Image classifiers are usually trained to learn attributes that can be unambiguously associated with one or more real world entities. However, subjective attributes may not strictly fit into these criteria since such attributes are inferred from the presence of relevant objects and their inter-relationships. Classification based on subjective attributes can pose challenges to existing image understanding techniques that are based on the detection of consistently occurring spatially coherent objects in the image domain. Additionally, building unbiased models for subjective attributes may require large, carefully sampled datasets and can incur a corresponding labeling cost.

DESCRIPTION

This disclosure describes a neuro-fuzzy controller that can be utilized to guide image classifier networks for classification of subjective attributes. Per techniques of this disclosure, linguistic expert rules for memberships of an image to various output categories of the subjective attribute(s) are framed and the classification is analyzed as a fuzzy system. Fuzzy rules and fuzzy inference output from this system are used to guide a neural network to effectively incorporate the expert rules. Specific loss functions are utilized to guide the image classifier. A *fuzzy-rule contradiction loss* is utilized to capture a weighted deviation of image classifier prediction from expert rules. A *fuzzy inference loss* is utilized to capture overall deviation from fuzzy inference output.

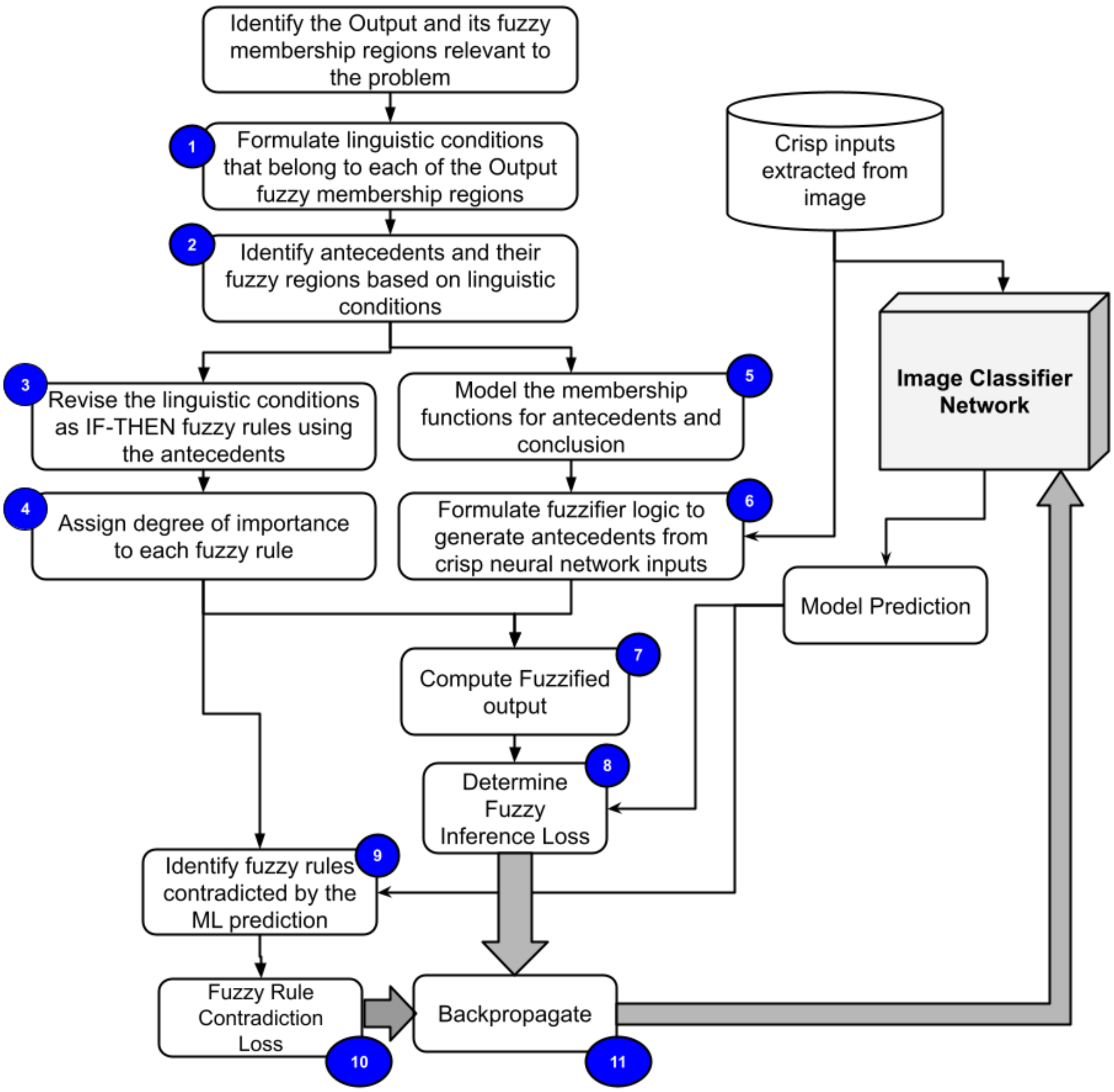

Fig. 1: Training of image classifier network using a neuro-fuzzy controller

Fig. 1 depicts an example model training workflow to train an image classifier network using a neuro-fuzzy controller, per techniques of this disclosure. Key features in the workflow are numbered in blue bubbles and are described below.

1. Conditions under which an image is considered a member of one of the output fuzzy regions are solicited from multiple experts of various backgrounds and denoted as “expert rules.” Contradictory expert rules can be modified or excluded for consistency.

2. The most frequently recurring antecedents employed in the expert rules are determined along with fuzzy regions for each of the antecedents that are relevant to the subject of the images. The antecedents can be denoted as $\sum_{d=1}^{D} \quad x_d$, where D is the number of antecedents identified for the system.

3. Using these antecedents, the expert rules are revised into standard fuzzy IF-THEN rules:

$$If\ x_1\ is\ A_{1k}\ and\ \ x_2\ is\ A_{2k}\ and\ \ \ldots\ \ x_d\ is\ A_{dk}\ and\ \ then\ \ y_k = \ f_k(x)$$

where k = 1, 2, ..., K and K is the number of fuzzy rules, x_i is the ith input variable, A_{ik} is a fuzzy set for ith input of kth rule, and is a fuzzy conjunction operator, y_k is the output, and $f_k(.)$ is output of kth rule.

4. A weight is assigned to each fuzzy rule based on the expert understanding of the problem statement and significance of the rule: $w_i \in [0,1]$ where w_i is the weight assigned to ith fuzzy rule. Relatively high weights can be assigned to certain rules that are not to be broken.

5. The membership functions for each antecedent are modeled based on the expert understanding of the input and overlap between the fuzzy regions. A trapezoidal membership function can be utilized for the antecedents.

6. Crisp-signals are extracted from the images using foundational models to generate inputs to the fuzzy system.

7. During the model training phase, at each iteration, the fuzzy output is computed using the model inputs and fuzzy rules e.g., by a Mamdani Fuzzy Inference System:

$$\sum_{i=1}^{I} \quad \mu_i\,(y)/y\ \ where\ \mu_i\,(y) \in [0,1],$$

where μ_i is the membership association of the output y to ith fuzzy region.

8. Deviation between the predicted model output and fuzzy membership association of the fuzzy output is determined and mathematically modeled. The deviation is denoted as a Fuzzy Inference Loss, $Loss_{fuzzy\ inference\ loss}$.

It is important to note that the prediction from a neural network is a probability of the image belonging to a particular class while the fuzzy inference system provides a degree of membership to fuzzy sets. Although these values lie within [0,1] they are fundamentally different and cannot be directly correlated. To suitably handle this difference, the region with highest membership association is identified:

$$t = iloc(max(\mu_i(y)/y)) \ \ where\ \mu_i(y)/y \in [0,1]$$

where $\mu_i(y)$ is the degree of membership association of the output y to ith fuzzy region, and t is the index of the fuzzy region with the highest membership association. A one hot encoded vector of length T is constructed to represent the total number of fuzzy regions in the output y. This must also be equal to the number of classes predicted by the neural network:

$$f_i = 0 \ where\ i \neq t$$

$$f_i = 1 \ where\ i = t$$

where f is the one hot encoded vector with only the index corresponding to the highest membership set. Cross entropy loss can be applied to this vector and the neural network predictions to compute the Fuzzy Inference loss:

$$Loss_{fuzzy\ inference\ loss} = -\sum_i^T \ f_i \log(s_i),$$

where s is the model prediction score of ith class.

9. Fuzzy rules that are contradicted by the predicted output are identified and the deviation is mathematically modeled as a loss function taking into account the relative weights of the fuzzy

rules. This loss function can be denoted as a Fuzzy Rule Contradiction Loss,$Loss_{fuzzy\ rule\ contradiction\ loss}$.

Similar to the fuzzy inference loss, in this case both the fuzzified antecedents and output are assigned to the region with highest membership association. Rules are differentiated based on whether they confirm, contradict, or escape.

- A prediction ***confirms*** a fuzzy rule if all the antecedent memberships match with that specified in the rule and the neural network predicted class is the same as the concluded output specified in the rule.
- A prediction ***contradicts*** a fuzzy rule if all the antecedent memberships match with that specified in the rule and the neural network predicted class is different from the concluded output specified in the rule.
- A prediction ***escapes*** a fuzzy rule if any of the antecedent memberships do not match with that specified in the rule.

For loss computation, confirmed and escaped rules are not considered and only the contradicted rules are considered. For a contradicted rule $A_k^{contradict}$ where A_k is the kth fuzzy rule in the database identified to be contradicted by the model prediction and a corresponding weight w_k, the fuzzy rule contradiction loss can be formulated as:

$$Loss_{fuzzy\ rule\ contradiction\ loss} = -\sum_{k=1}^{K} \left(w_k \cdot \left(\sum_{i}^{T} f_i \log(s_i)\right) \cdot l\right)$$

$$where\ l = 1\ if\ A_k = A_k^{contradict}\ ;\ else\ l = 0$$

In a typical scenario where membership of all antecedents are defined for all fuzzy rules, there is only one contracting rule. In some scenarios, there may also be partial rules being triggered which can lead to multiple rules being contradicted. The weights, w_k can be utilized to handle the loss values in this case.

10. The loss functions $Loss_{fuzzy\ inference\ loss}$ and $Loss_{fuzzy\ rule\ contradiction\ loss}$ are back propagated along with other loss functions used by the image classifier (e.g., cross entropy loss) during the training phase.

In an example, the workflow can be applied to a specific example use case of classifying restaurants based on received restaurant images as to a degree of their "family-friendliness."

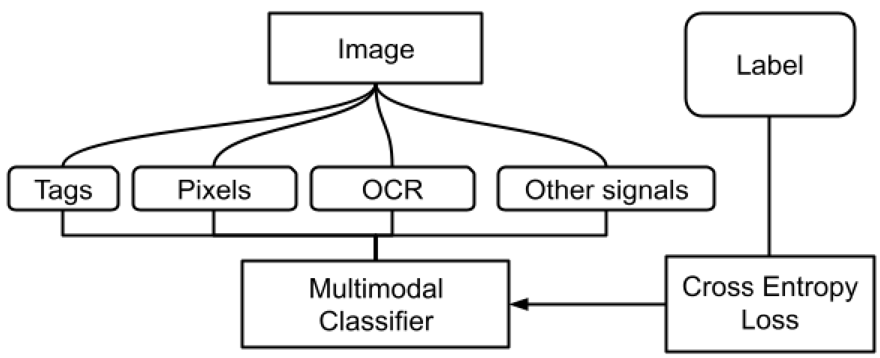

Fig. 2: Image classifier model

Fig. 2 depicts an example image classifier model that supports factual attributes and that does not include neuro-fuzzy controller inputs. As depicted in Fig. 2, standard and readily available signals extracted from images are utilized to build a classifier. Raw pixel analysis may be utilized to handle the complexity of labels. However, insufficient data can result in such an image classifier model anchoring its belief to irrelevant objects, e.g., predicting family friendliness of a restaurant based on the presence of a clock on the wall in images. Additionally, unconscious data biases can also result in models enforcing cultural stereotypes.

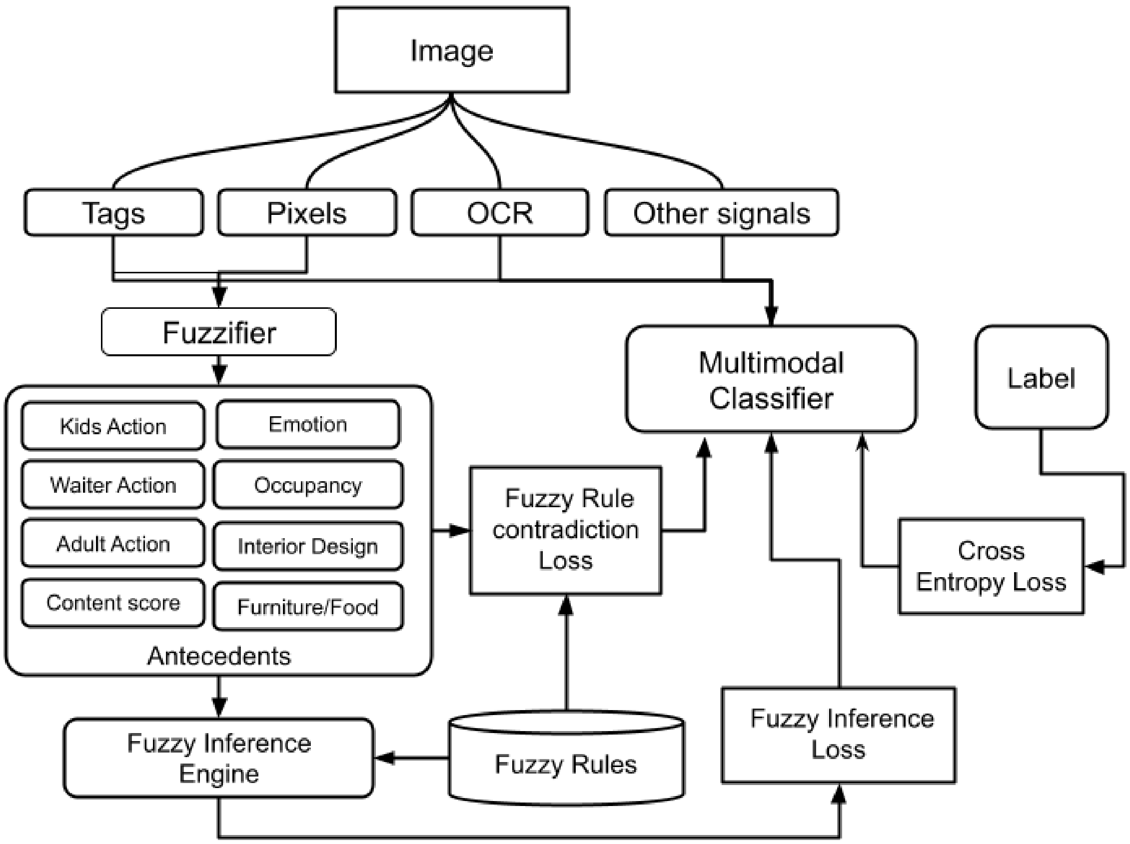

Fig. 3: Image Classifier Model with neuro-fuzzy controller inputs

Fig. 3 depicts application of an image classifier model with neuro-fuzzy controller inputs to classify restaurants based on their family friendliness, per techniques of this disclosure.

Step 0: Formulating the Fuzzy System

The fuzzy system is formulated by defining membership functions associated with the different degrees of family friendliness.

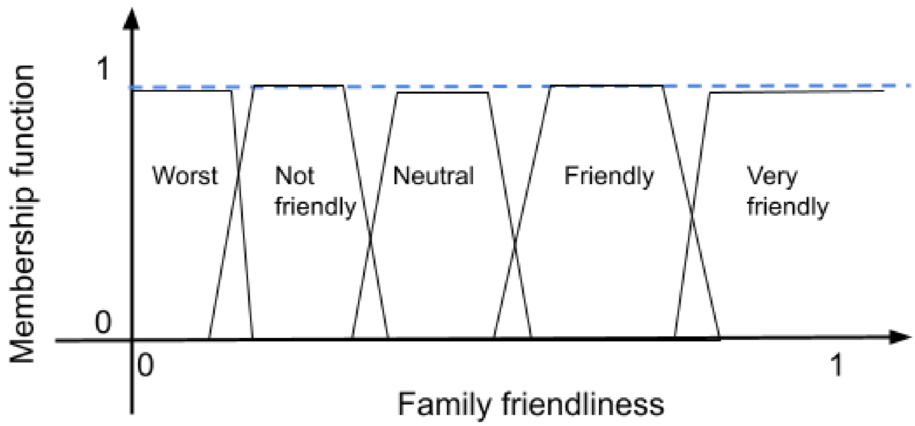

Fig. 4: Membership functions for degrees of family friendliness

Fig. 4 depicts example levels of family friendliness and corresponding membership functions. The membership function can be treated as an output of a fuzzy system with 5 partially overlapping fuzzy regions. The membership functions are modeled as partially overlapping trapezoidal functions for not-friendly, neutral, friendly regions. As shown in Fig. 4, worst and very friendly extremes are modeled with left open and right open functions.

Step 1: Expert Rule Generation

Linguistic conditions are obtained from experts across multiple backgrounds. For example, these can be responses to the question, “What characteristics in an image would you expect in order to classify it in a particular output category?”

Step 2: Identify the Antecedents

Based on the linguistic conditions, antecedents are identified that can be utilized to differentiate an image across categories. In some implementations, the antecedents can be identified by language models.

Antecedent	Fuzzy Regions
Kids Action	Missing, Active, Hyperactive
Waiter Action	Missing, Can't tell, Engaged
Adult Action	Missing, Active, Hyperactive
Dominant Emotion	Sad, Neutral, Happy, Ecstatic
Occupancy	Empty, Sparse, Occupied, Crowded
Interior Design	Can't tell, Offensive, Neutral, Aesthetic
Furniture/Food	Missing, Adequate
Inappropriate Content	Not Applicable, Neutral, Explicit

Table 1: Antecedents and corresponding fuzzy regions

Table 1 depicts example antecedents and their fuzzy regions.

Step 3: Revise Linguistic Conditions to Fuzzy Rules

The linguistic rules are revised into a fuzzy IF-THEN format using the antecedents.

Step 4: Assign degree of importance (weights) to fuzzy rules

Weights are utilized to enforce some rules more strictly than others. For example, rules related to inappropriate content scores can be assigned very high weights and hence override any other antecedent related rule in the system and/or other loss to the neural network.

Step 5: Model the Antecedent Membership Functions

The membership functions for each of the fuzzy regions in the antecedents is modeled for each antecedent.

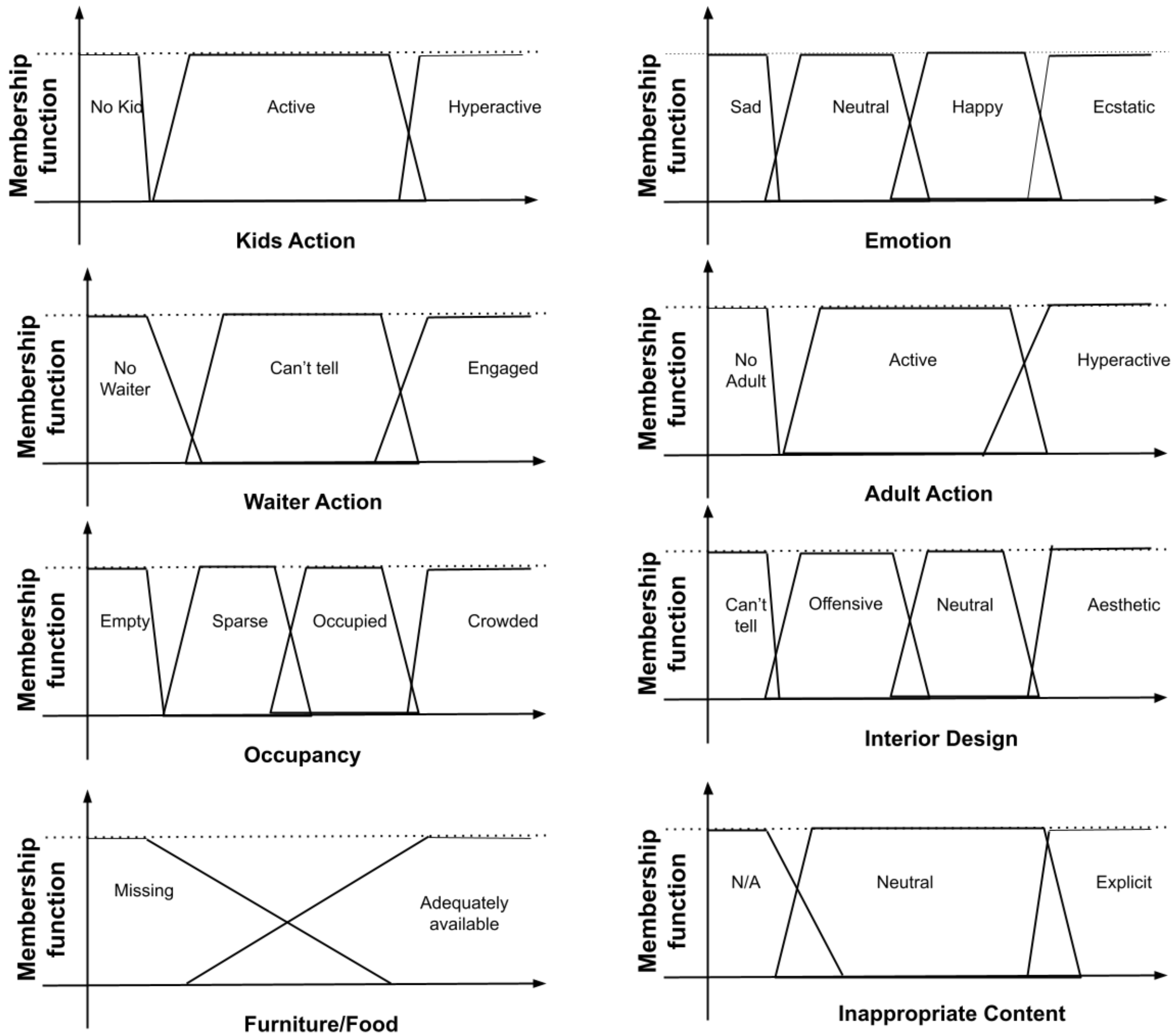

Fig. 5: Antecedent membership functions

Fig. 5 depicts example membership functions modeled for this example. As depicted in Fig. 5, the overlap in some fuzzy regions and exclusions in others can be deliberately set up based on common sense understanding of the fuzzy boundaries for the antecedents. For example, in the kid action antecedent, there is no overlap between “No Kid” and “Active”; while in case of emotion antecedent, there is a significant overlap of “Happy” fuzzy region with both “Neutral” and “Ecstatic.”

<u>Step 6: Fuzzify Image Metadata into Antecedent fuzzy sets</u>

Image metadata is fuzzified to generate additional fuzzy inferences. Subsequent to performing steps 0-6, the fuzzy output and loss values are back propagated into the neural network, e.g., similar to steps described with reference to Fig. 1.

Upon completion of training, the trained image classifier with neuro-fuzzy controller inputs can be utilized to analyze images of restaurants to provide accurate classification and labels for a degree of family friendliness. Using the described methods, image classifiers can be trained for any other purpose.

CONCLUSION

This disclosure describes a novel neuro-fuzzy controller that can be utilized to guide image classifier networks for classification of subjective attributes. Per techniques of this disclosure, linguistic expert rules for memberships of an image to various output categories of the subjective attribute are framed and the classification is analyzed as a fuzzy system. Fuzzy rules and fuzzy inference output from this system are used to guide a neural network to effectively incorporate the expert rules. Specific loss functions are utilized to guide the image classifier; a *fuzzy-rule contradiction loss* to capture a weighted deviation of image classifier prediction from expert rules, and a *fuzzy inference loss* to capture overall deviation from fuzzy inference output. Utilization of the neuro-fuzzy controller can enable image classifier models to classify images according to subjective attributes, e.g., to provide accurate labels for family friendliness of a restaurant based on images of the restaurant.

REFERENCES

1. N. Talpur, S. J. Abdulkadir and M. H. Hasan, "A deep learning based neuro-fuzzy approach for solving classification problems," 2020 International Conference on Computational Intelligence (ICCI), Bandar Seri Iskandar, Malaysia, 2020, pp. 167-172, doi: 10.1109/ICCI51257.2020.9247639.

2. Vaishnavi Sonwalkar, “Neuro Fuzzy Controller,” https://vsonwalkar3.medium.com/neuro-fuzzy-controller-ab6246d0cb90, accessed on 25 September 2023.

3. Francesco Gelli, Tiberio Uricchio, Xiangnan He, Alberto Del Bimbo, and TatSeng Chua, “Learning Subjective Attributes of Images from Auxiliary Sources,” In Proceedings of the 27th ACM International Conference on Multimedia (MM ’19), October 21–25, 2019, Nice, France. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3343031.3350574

4. D. Borth, R. Ji, T. Chen, T. Breuel, and S-F Chang, “Large-scale Visual Sentiment Ontology and Detectors Using Adjective Noun Pairs,” ACM MM, 2013.

5. Hand, Emily M., Carlos Domingo Castillo and Rama Chellappa, “Doing the Best We Can With What We Have: Multi-Label Balancing With Selective Learning for Attribute Prediction,” AAAI Conference on Artificial Intelligence (2018).

6. Francesco Gelli, Tiberio Uricchio, Xiangnan He, Alberto Del Bimbo, and Tat-Seng Chua, “Learning Visual Elements of Images for Discovery of Brand Posts,” ACM Trans. Multimedia Comput. Commun. Appl. 16, 2, Article 56 (May 2020), 21 pages. https://doi.org/10.1145/3385413.

7. Yazdanbakhsh, O., & Dick, S., “A Deep Neuro-Fuzzy Network for Image Classification,”ArXiv, abs/2001.01686.

8. Xu, Jingyi, Zilu Zhang, Tal Friedman, Yitao Liang, and Guy Broeck, "A semantic loss function for deep learning with symbolic knowledge," In International conference on machine learning, pp. 5502-5511. PMLR, 2018.

9. Fan, Lixin, "Revisit fuzzy neural network: Demystifying batch normalization and ReLU with generalized hamming network," Advances in Neural Information Processing Systems 30 (2017).

10. Yang, R., Yu, J., Yin, J. et al. An FA-SegNet Image Segmentation Model Based on Fuzzy Attention and Its Application in Cardiac MRI Segmentation. Int J Comput Intell Syst 15, 24 (2022). https://doi.org/10.1007/s44196-022-00080-x.

11. X. Chen, D. Li, P. Wang and X. Yang, "A Deep Convolutional Neural Network With Fuzzy Rough Sets for FER," IEEE Access, vol. 8, pp. 2772-2779, 2020, doi: 10.1109/ACCESS.2019.2960769.

12. E. Soria-Olivas, J. D. Martin-Guerrero, G. Camps-Valls, A. J. Serrano-Lopez, J. Calpe-Maravilla and L. Gomez-Chova, "A low-complexity fuzzy activation function for artificial neural networks," in IEEE Transactions on Neural Networks, vol. 14, no. 6, pp. 1576-1579, Nov. 2003, doi: 10.1109/TNN.2003.820444.