Improving Evidential Sources with Contextual Corrections

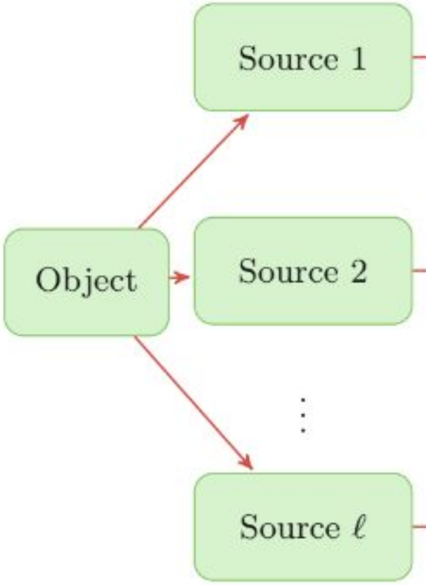
HSE, 2022

Improvement with Contextual Corrections

On Improving a Group of Evidential Sources with Different Contextual Corrections, Mutmainah et al. 2022

- **Information fusion:** combine different heterogeneous sources of information, to obtain a better understanding of the situation under evaluation

- **Classifiers as sources of information**, optimize directly the performance of the combination using different corrections among CD, CR and CN.



Dempster-Shafer theory

- Mass function (provided by classifiers) ***m(A)***, universe ***Ω***:

$$\sum_{A \subseteq \Omega} m(A) = 1$$

- Belief (evidence for) and plausibility (evidence that does not contradict):

$$Bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B)$$

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) = Bel(\Omega) - Bel(\overline{A})$$

Contextual Corrections

- **Unnormalized Dempster's rule -** combine independent sources in the same ***Ω***

$$(m_1 \bigcirc\!\!\!\!\!\cap\, m_2)(A) = m_{1 \bigcirc\!\!\!\!\cap\, 2}(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad \forall A \subseteq \Omega$$

- **Contextual correction** - discounting a source reliable only to a degree

$$\begin{aligned} {}^{\beta}m &= \beta\, m + \alpha\, m_{\Omega} \\ &= \begin{cases} A \mapsto \beta\, m(A) & \forall A \subset \Omega \\ \Omega \mapsto \beta\, m(\Omega) + \alpha \end{cases} \end{aligned}$$

Contextual Corrections

- Contour functions of each contextual correction - CD, CR, CN

Corrections	Contour functions
CD	${}^{\beta}pl(\omega) = 1 - (1 - pl(\omega))\beta_\omega$
CR	${}^{\beta}pl(\omega) = pl(\omega)\beta_\omega$
CN	${}^{\beta}pl(\omega) = 0.5 + (pl(\omega) - 0.5)(2\beta_\omega - 1)$

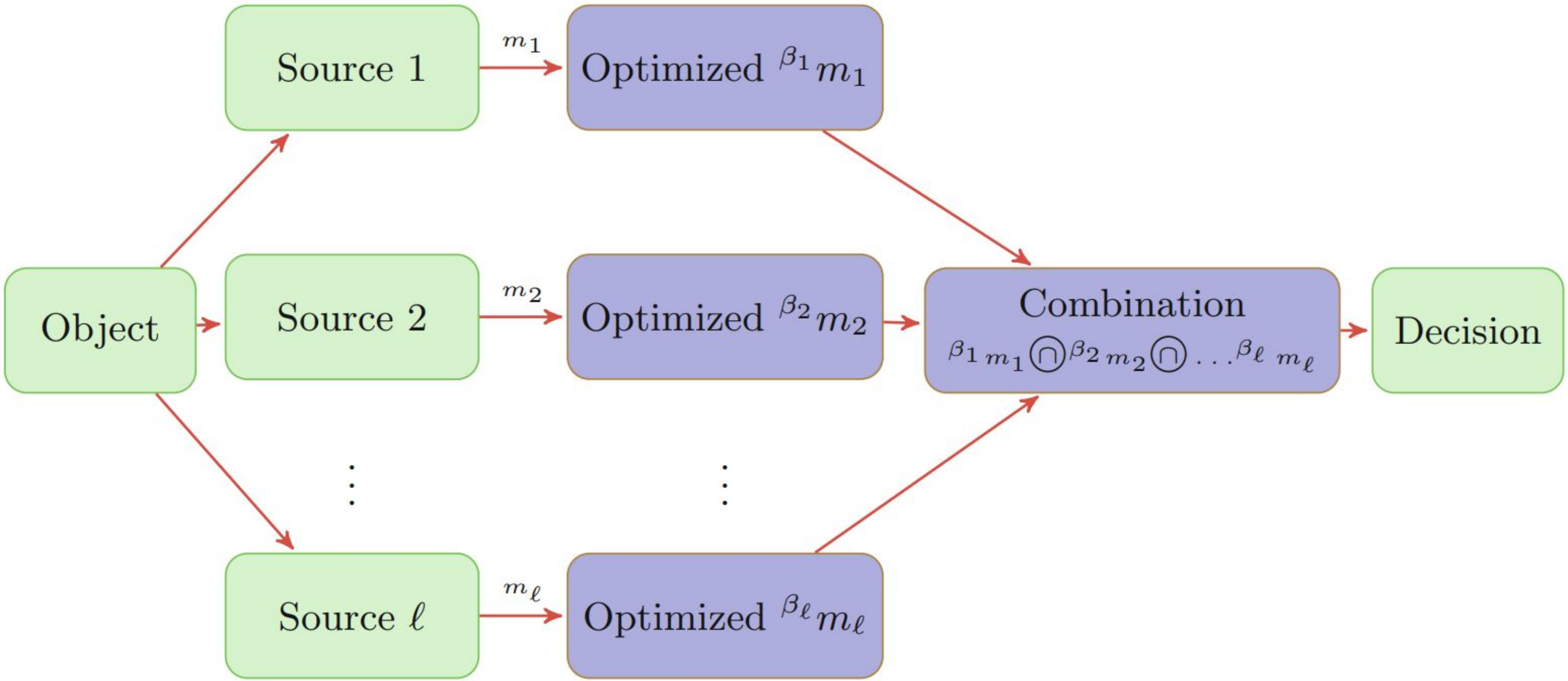
Error function

- Discrepancy between the corrected outputs and the true classes of the objects

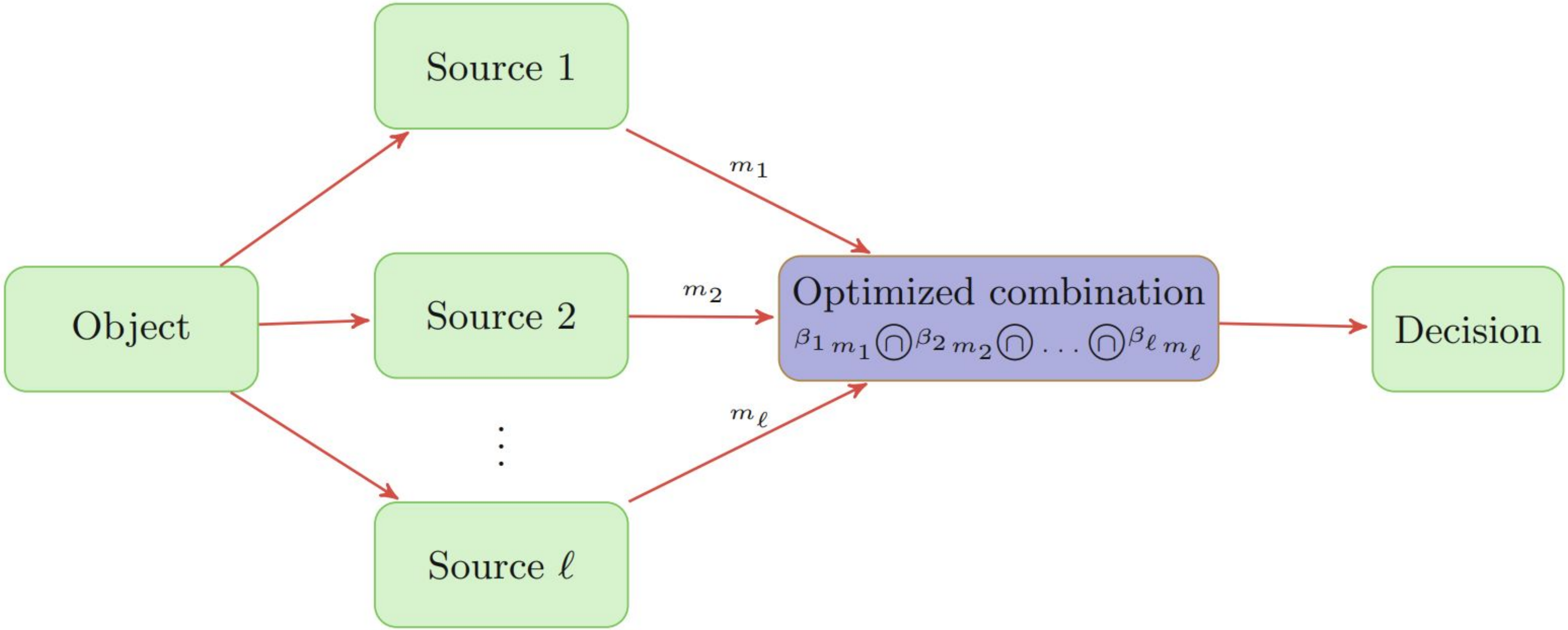
$$E_{pl}(\beta) = \sum_{i=1}^{n} \sum_{k=1}^{K} ({}^{\beta}pl\{o_i\}(\{\omega_k\}) - \delta_{i,k})^2$$

- Where ***pl{o**$_i$**}*** is the output of a model on object ***o**$_i$*

Scheme 1 - independent



Scheme 2 - concurrent



Error function

- Discrepancy between the corrected outputs and the true classes of the objects

$$E_{pl}(\beta) = \sum_{i=1}^{n} \sum_{k=1}^{K} ({}^{\beta}pl\{o_i\}(\{\omega_k\}) - \delta_{i,k})^2$$

- Updated function for simultaneous training

$$E_{pl}(\beta_1, \ldots, \beta_\ell) = \sum_{i=1}^{n} \sum_{k=1}^{K} ({}^{\beta_1}pl_1\{o_i\}(\{\omega_k\}) \times \\ \times \ldots \times {}^{\beta_\ell}\, pl_\ell\{o_i\}(\{\omega_k\}) - \delta_{i,k})^2$$

Paper results

Tested on several small datasets, 10 fold cross-val, EkNN and ENN - brief mention

Data	No cont. correction	Scheme 1 correction	Scheme 2 correction	Scheme 2 only CD	Scheme 2 only CR	Scheme 2 only CN
Synthetic (900 obj)	10.953 (3.094)	<u>11.690</u> (2.748)	9.496 (2.492)	**9.491** (2.525)	10.961 (3.076)	10.882 (2.913)
Haberman (306 obj)	6.054 (2.381)	<u>6.742</u> (1.664)	**5.515** (1.725)	5.886 (2.096)	5.717 (2.070)	5.577 (1.849)
Iris (150 obj)	0.503 (0.619)	<u>0.569</u> (0.608)	**0.467 (0.715)**	0.471 (0.712)	0.503 (0.619)	0.503 (0.619)
Lympho (140 obj)	2.305 (1.077)	<u>2.748</u> (0.929)	2.253 (1.058)	**2.239** (1.045)	2.322 (1.086)	2.322 (1.059)

...

***Table 1.1:** Epl on various datasets from the paper*

Paper results

Tested on several small datasets, 10 fold cross-val, EkNN and ENN - brief mention

Data	No cont. correction	Scheme 1 correction	Scheme 2 correction	Scheme 2 only CD	Scheme 2 only CR	Scheme 2 only CN
Synthetic (900 obj)	<u>0.850</u> (0.052)	0.857 (0.046)	0.857 (0.046)	**0.858** (0.046)	<u>0.850</u> (0.052)	<u>0.850</u> (0.052)
Haberman (306 obj)	0.755 (0.103)	<u>0.747</u> (0.112)	0.758 (0.096)	**0.762** (0.097)	0.756 (0.103)	0.759 (0.104)
Iris (150 obj)	**0.971** (0.062)	**0.971** (0.062)	<u>0.968</u> (0.064)	0.969 (0.063)	**0.971** (0.062)	**0.971** (0.062)
Lympho (140 obj)	**0.815** (0.139)	<u>0.805</u> (0.149)	<u>0.805</u> (0.143)	0.806 (0.143)	0.814 (0.139)	**0.815** (0.135)

...

Table 1.2: *Utility on various datasets from the paper*

Experiment setup

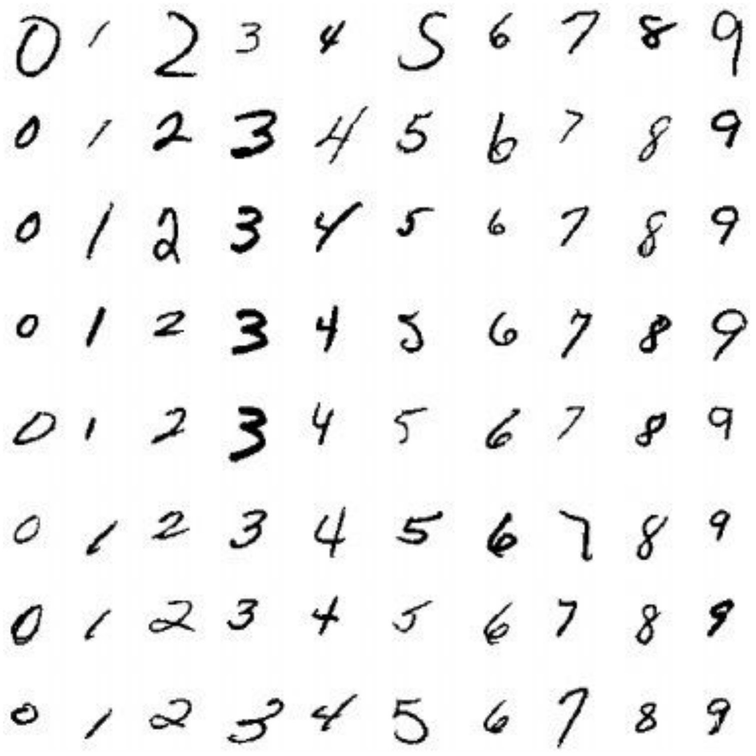
Datasets:

- **MNIST**, 30.000 training,
  30.000 cross-validation and fine-tuning
- **FashionMNIST**, 30.000 training,
  30.000 cross-validation and fine-tuning
- **CIFAR10**, 25.000 training,
  25.000 cross-validation and fine-tuning

Universe: 10 classes (digits 0…9)

Models: ECNN and a ENN

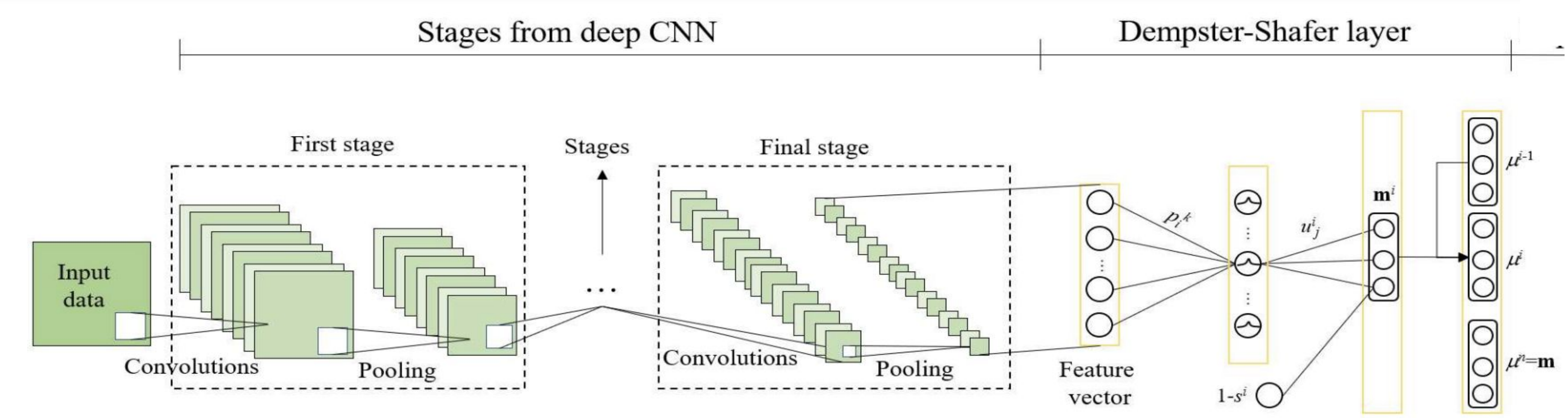
Metric: Epl for training and validation



Experiment setup

ECNN and **ENN**, added Dempster-Shafer layer, trained in two phases

- 20 epochs for probabilistic head - reliable
- 100 epochs for evidential head - highly unpredictable



Error of ENNs with Corrections

ECNN and **ENN**.; **5 fold** cross-validation

Models	No cont. correction	Scheme 1 correction	Scheme 2 correction	Scheme 2 only CD	Scheme 2 only CR	Scheme 2 only CN
MNIST (60.000 obj)	3778.89 (98.95)	<u>5500.31</u> (148.73)	3133.84 (153.67)	**3102.65** (345.86)	3696.77 (341.49)	3164.96 (172.30)
FashionMNIST (60.000 obj)	3875.15 (421.75)	<u>4553.35</u> (290.12)	**3658.21** (417.60)	3762.23 (457.83)	3768.29 (384.40)	3661.94 (412.61)
CIFAR10 (50.000 obj)	4344.89 (207.70)	<u>4533.96</u> (162.24)	**4122.45** (311.65)	4156.81 (310.72)	4318.55 (210.37)	4131.32 (302.10)

Table 2.1: *Epl on MNIST, experiment*

Utility of ENNs with Corrections

ECNN and **ENN**.; **5 fold** cross-validation

Models	No cont. correction	Scheme 1 correction	Scheme 2 correction	Scheme 2 only CD	Scheme 2 only CR	Scheme 2 only CN
MNIST (60.000 obj)	0.776 (0.042)	<u>0.724</u> (0.054)	0.760 (0.025)	**0.776** (0.046)	**0.776** (0.043)	0.713 (0.074)
FashionMNIST (60.000 obj)	**0.785** (0.021)	0.776 (0.024)	<u>0.737</u> (0.057)	<u>0.737</u> (0.057)	**0.785** (0.021)	0.738 (0.057)
CIFAR10 (50.000 obj)	0.450 (0.003)	0.454 (0.001)	0.461 (0.011)	**0.464** (0.009)	<u>0.449</u> (0.004)	0.459 (0.012)

Table 2.2: *Utility on MNIST, experiment*

Comparing Evidential and Probabilistic NNs

Larger amount of data: **MNIST** 60.000 obj.; **5 fold** cross-validation

Models	No cont. correction	Scheme 1 correction	Scheme 2 correction	Scheme 2 only CD	Scheme 2 only CR	Scheme 2 only CN
CNN+FCN	<u>3709.00</u> (11.67)	3698.68 (11.61)	**3696.87** (11.62)	**3696.87** (11.62)	3711.10 (11.68)	3699.26 (11.63)
ECNN+ENN	3778.89 (98.95)	<u>5500.31</u> (148.73)	3133.84 (153.67)	**3102.65** (345.86)	3696.77 (341.49)	3164.96 (172.30)

Table 3: *Epl on MNIST, experiment*

Conclusion

- Evidential NNs require complex training
- Contextual correction scores are better when corrections are done in tandem
- Strategy limited when the number of sources to be combined increases
- Evidential NNs combined using this technique can achieve better utility that probabilistic NNs

Sources

- On Improving a Group of Evidential Sources with Different Contextual Corrections
  Siti Mutmainah , Samir Hachour , Frédéric Pichon , David Mercier, 2022
- An evidential classifier based on Dempster-Shafer theory and deep learning
  Zheng Tong, Philippe Xu, Thierry Denœux, 2021
- Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence, ThierryDenœux, 2008