Copula-based conformal prediction for object detection

A more efficient approach

Bruce Cyusa Mukama, Soundouss Messoudi, Sylvain Rousseau, Sébastien Destercke

Heudiasyc, CNRS, Compiègne, France

Université de Technologie de Compiègne, France

Milan, September 11, 2024

Table of Contents

(Safety-critical) Object detection

- Sub-task N°1: predict object categories, e.g., car, pedestrian.
- Sub-task N°2: predict object locations, e.g., $[(10, 15), (35, 75)]$.

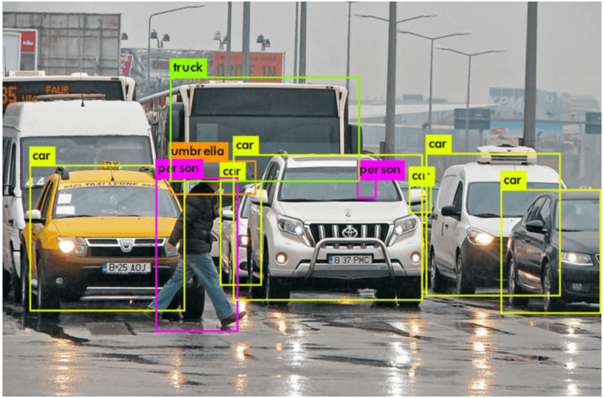

Figure 1: Detecting pedestrians & vehicles.

Other (safety-critical) applications

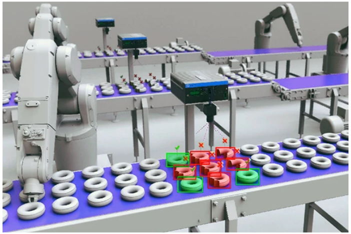

Figure 2: Quality control.

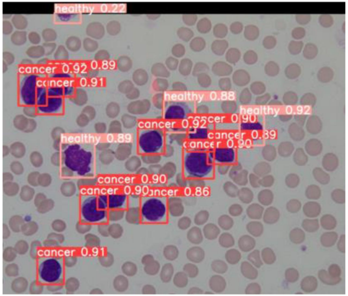

Figure 3: Medical diagnosis.

Figure 4: Plant monitoring.

- We need UQ because failure in these systems can result in catastrophes!

Object detection (under the hood)

Object detectors are shipped without any rigorously calibrated UQ mechanism:

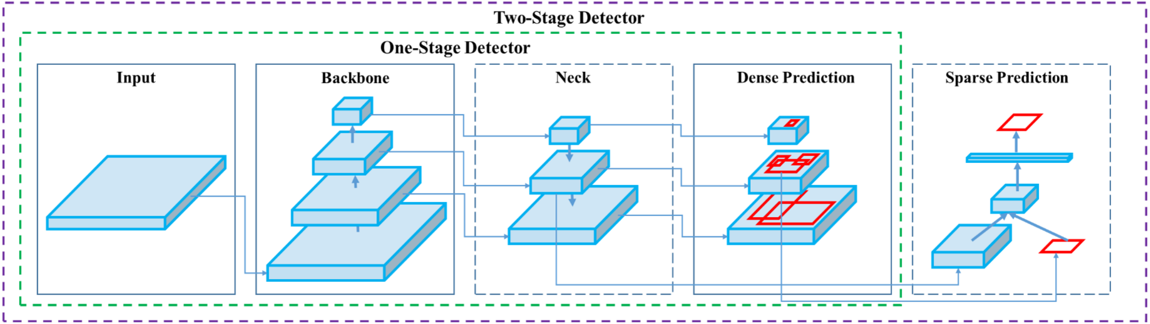

Figure 5: The typical architecture of object detectors [2].

Table of Contents

Conformal bounding boxes

What are bounding box confidence regions?

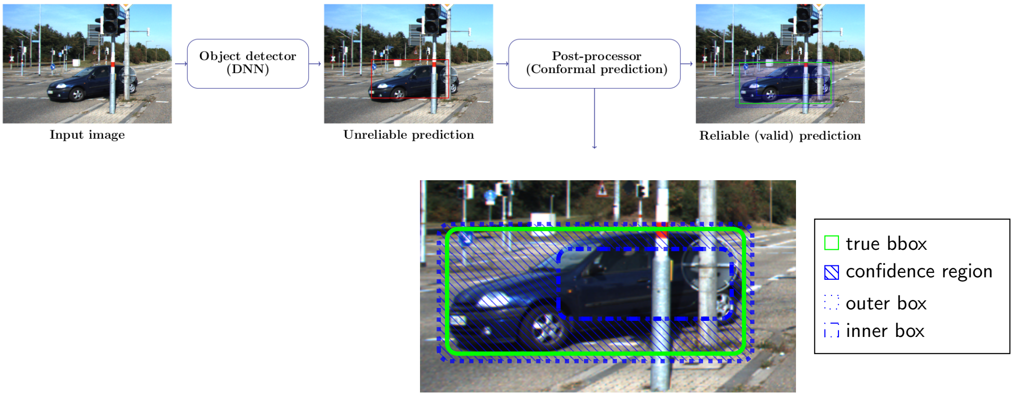

Figure 6: A bounding box confidence region (bottom) and its inference pipeline (top).

Reliable bounding boxes

A formal definition of the problem:

- the available data (boxes) $\rightarrow \{(\widehat{b}_i, b_i)\}_{i=1}^{n}$ with $b_i = [\underline{x}_i, \underline{y}_i, \overline{x}_i, \overline{y}_i]$,
- the non-conformity scores $\rightarrow \alpha_i = |b_i - \widehat{b}_i| \in \mathbb{R}^4$,
- the desired confidence level $\rightarrow 1 - \epsilon^g \in (0, 1)$,
- the conformal prediction region: $\mathcal{B}_i^\epsilon = [\widehat{\underline{b}}_i, \widehat{\overline{b}}_i] \in \mathbb{R}^{2\times 4}$,
- the goal:

$$P(b_{n+1} \in \mathcal{B}_{n+1}^\epsilon) \geq 1 - \epsilon^g \quad (1)$$

$$P(|\underline{x}_{n+1} - \widehat{\underline{x}}_{n+1}| \leq \alpha_s^1, \ldots, |\overline{y}_{n+1} - \widehat{\overline{y}}_{n+1}| \leq \alpha_s^4) \geq 1 - \epsilon^g \quad (2)$$

$$F(\alpha_s^1, \ldots, \alpha_s^4) \geq 1 - \epsilon^g \quad (3)$$

Copula-based conformal bounding boxes

We can use copulas to deduce dimension-wise confidence levels $\{1 - \epsilon^d\}_{d=1}^4$ [5].

Sklar's theorem [6]

Every joint c.d.f. is composed of (d) marginal c.d.f(s) and their dependency model $\mathcal{C} : [0, 1]^d \to [0, 1]$, i.e., the copula [4].

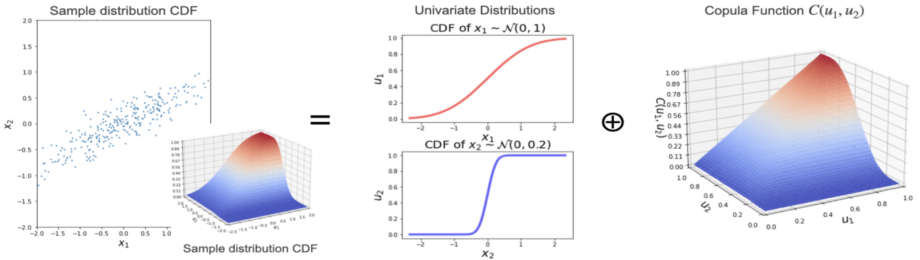

Figure 7: An illustration of Sklar's theorem, from [7].

Copula-based conformal bounding boxes

To achieve perfect calibration:

$$
\begin{aligned}
P(b_i \in \mathcal{B}_i^{\epsilon}) = F(\alpha_s^1, \ldots, \alpha_s^4) = 1 - \epsilon^g \\
\mathcal{C}(F^1(\alpha_s^1), \ldots, F^4(\alpha_s^4)) = \mathcal{C}(1 - \epsilon^1, \ldots, 1 - \epsilon^4) = 1 - \epsilon^g
\end{aligned}
\tag{4}
$$

We can solve (4) and compute dimension-wise quantiles:

$$
[\alpha_s^1, \ldots, \alpha_s^4] = [Q^1((1 - \epsilon^1) \times (n+1)/n), \ldots, Q^4((1 - \epsilon^4) \times (n+1)/n)] \tag{5}
$$

We define the prediction region's bounds as follows:

$$
\underline{\widehat{b}}_i \leftarrow [\underline{\hat{x}}_i + \alpha_s^1,\ \underline{\hat{y}}_i + \alpha_s^2,\ \hat{\bar{x}}_i - \alpha_s^3,\ \hat{\bar{y}}_i - \alpha_s^4] \tag{6}
$$

$$
\hat{\bar{b}}_i \leftarrow [\underline{\hat{x}}_i - \alpha_s^1,\ \underline{\hat{y}}_i - \alpha_s^2,\ \hat{\bar{x}}_i + \alpha_s^3,\ \hat{\bar{y}}_i + \alpha_s^4] \tag{7}
$$

Table of Contents

Tested approaches

Preexisting approaches [1]:

- Multiple hypothesis testing: $\epsilon^1 = \cdots = \epsilon^4 = \epsilon^g/4$,
- Dimensionality reduction:

  $\alpha_i = max(|b_i - \widehat{b}_i|) \in \mathbb{R}$ and $\alpha_s^1 = \cdots = \alpha_s^4 = Q((1 - \epsilon^g) \times (n+1)/n)$.

Our (copula) approaches:

- Independent copula: $C_\pi(u^1, \ldots, u^m) = \prod_{t=1}^m u^t$,
- Gumbel copula:

  $C_G(u^1, \ldots, u^m) = \exp\left(\sum_{t=1}^m (-\ln u^t)^\theta\right)^{\frac{1}{\theta}}$,
- Empirical copula:

  $C_E(u^1, \ldots, u^m) = \frac{1}{n} \sum_{i=1}^n \prod_{t=1}^m \mathbb{1}_{u_i^t \leq u^t}$.

Test results on synthetic data

Settings:

- $\alpha_i = |b_i - \widehat{b}_i| \sim \mathcal{U}(\Omega)$ with $\Omega = [0, 2.8] \times [0, 2.5] \times [0, 8] \times [0, 2.5]$

Results:

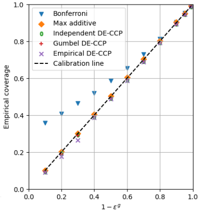

Figure 8: Calibration.

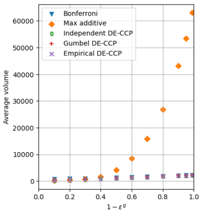

Figure 9: Regions' volumes.

- our approach is more robust to the disparity of ranges between the dimensions!

Test results on benchmarks

We also use popular (real-life) benchmarks:

Figure 10: An example from the KITTI dataset [3].

Figure 11: An example from the BDD100K dataset [8].

- KITTI is smaller and was collected with a single platform (Germany),
- BDD100K is larger and was collected with multiple platforms (USA).

Test results on benchmarks

Our approach yields smaller prediction regions:

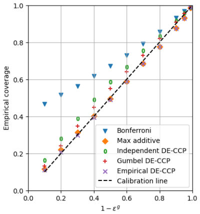

Figure 12: Calibration on KITTI.

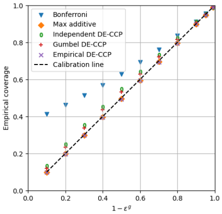

Figure 13: Calibration on BDD100K.

$1-\epsilon^g$	Bonferroni	Max additive	Independent DE-CCP	Gumbel DE-CCP	Empirical DE-CCP
0.99	6.96e+09±5.47e+09	1.06e+10±1.34e+10	3.78e+09±3.56e+09	3.45e+09±3.76e+09	**7.33e+08±1.02e+09**
0.95	2.10e+07±2.10e+07	4.33e+05±2.88e+05	6.97e+06±9.28e+06	3.44e+06±4.07e+06	**2.87e+05±1.85e+05**
0.90	4.42e+05±4.99e+05	4.86e+04±1.64e+04	1.84e+05±1.24e+05	1.21e+05±6.99e+04	**3.48e+04±1.11e+04**
0.80	2.31e+04±1.01e+04	6.95e+03±1.31e+03	1.56e+04±5.45e+03	1.14e+04±3.30e+03	**5.42e+03±7.63e+02**

Table 1: Regions' volumes on KITTI.

$1-\epsilon^g$	Bonferroni	Max additive	Independent DE-CCP	Gumbel DE-CCP	Empirical DE-CCP
0.99	1.59e+08±9.23e+06	1.73e+08±9.97e+06	1.51e+08±7.40e+06	**1.41e+08±7.79e+06**	1.41e+08±8.60e+06
0.95	9.11e+06±4.91e+05	6.44e+06±3.23e+05	8.29e+06±4.79e+05	7.49e+06±3.99e+05	**6.22e+06±3.24e+05**
0.90	1.62e+06±2.76e+04	9.84e+05±1.55e+03	1.36e+06±6.67e+03	1.21e+06±8.24e+01	**8.88e+05±5.10e+03**
0.80	2.46e+05±2.75e+02	1.40e+05±1.23e+03	1.93e+05±6.70e+01	1.73e+05±4.26e+01	**1.27e+05±6.38e+02**

Table 2: Regions' volumes on BDD100K.

Table of Contents

Key takeaways

Copula-based conformal object detection's advantages:

- robustness to the disparity of dimension ranges,
- higher efficiency on popular benchmarks.

The limitations:

- the guarantees only apply to detected objects,
- the classification task is not (yet) addressed.

Future directions:

- exploring vine & hierarchical copulas,
- tracking moving objects.

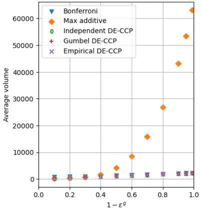

Figure 14: Regions' volumes.

References

[1] Léo Andéol, Thomas Fel, Florence De Grancey, and Luca Mossina.
Confident object detection via conformal prediction and conformal risk control: an application to railway signaling.
In *Conformal and Probabilistic Prediction with Applications*, pages 36–55. PMLR, 2023.

[2] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao.
Yolov4: Optimal speed and accuracy of object detection.
arXiv preprint arXiv:2004.10934, 2020.

[3] Andreas Geiger, Philip Lenz, and Raquel Urtasun.
Are we ready for autonomous driving? the kitti vision benchmark suite.
In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012.

[4] Christian Genest and Anne-Catherine Favre.
Everything you always wanted to know about copula modeling but were afraid to ask.
Journal of hydrologic engineering, 12(4):347–368, 2007.

[5] Soundouss Messoudi, Sébastien Destercke, and Sylvain Rousseau.
Copula-based conformal prediction for multi-target regression.
Pattern Recognition, 120:108101, 2021.

[6] M Sklar.
Fonctions de répartition à n dimensions et leurs marges.
In *Annales de l'ISUP*, volume 8, pages 229–231, 1959.

[7] Sophia Huiwen Sun and Rose Yu.
Copula conformal prediction for multi-step time series prediction.
In *The Twelfth International Conference on Learning Representations*, 2023.

[8] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell.
Bdd100k: A diverse driving dataset for heterogeneous multitask learning.
In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2636–2645, 2020.

[9] Ruiyao Zhang, Ping Zhou, and Tianyou Chai.
Improved copula-based conformal prediction for uncertainty quantification of multi-output regression.
Journal of Process Control, 129:103036, 2023.

Backup materials

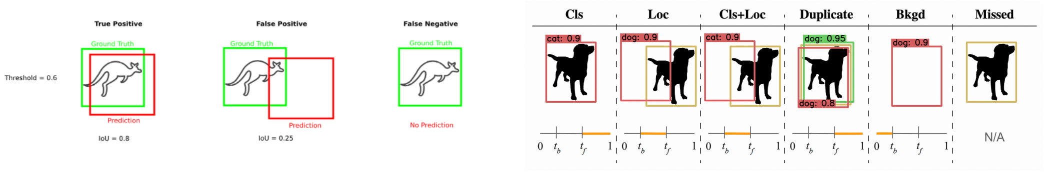

Figure 15: Examples of object detection errors.

Backup materials

How to select a solution-tuple among many candidate confidence levels:

- we can explicitly minimize confidence region sizes [9],

$$\underset{\epsilon^1,\ldots,\epsilon^4}{\arg\min} \prod_{d=1}^{4} \left(2 \times \alpha_s^d\right) \quad \text{s.t.} \quad \begin{cases} \mathcal{C}(1-\epsilon^1,\ldots,1-\epsilon^4) \geq 1-\epsilon \\ \epsilon^d \in \ (0,\epsilon^d] \end{cases} \tag{8}$$

Backup materials

Algorithm 1 The generic calibration procedure for box-wise SCP

Require: a global significance level ϵ^g, an object detector f_θ, a dataset D

1: Split the dataset D in two subsets: D_{train} & $D_{\text{cal}} = \{(X_i, Y_i)\}_{i=1}^n$,

2: Fit or fine tune f_θ on D_{train},

3: Follow Algorithm 2 to compute bounding box dissimilarity scores $\{\alpha_{i,j}\}_{i=1}^n$,

4: Compute conformal quantiles $\{\alpha_s^1, \alpha_s^2, \alpha_s^3, \alpha_s^4\}$ from $\{\alpha_{i,j}\}_{i=1}^n$ and ϵ^g,

5: For any new predicted box $\hat{B}_{n+1,j}$, infer an inner box $\underline{\hat{B}}_{n+1,j}$ and an outer box $\hat{\bar{B}}_{n+1,j}$:

$$\underline{\hat{B}}_{i,j} = \{\underline{\hat{x}}_{i,j} + \alpha_s^1,\ \underline{\hat{y}}_{i,j} + \alpha_s^2,\ \hat{\bar{x}}_{i,j} - \alpha_s^3,\ \hat{\bar{y}}_{i,j} - \alpha_s^4\} \tag{9}$$

$$\hat{\bar{B}}_{i,j} = \{\underline{\hat{x}}_{i,j} - \alpha_s^1,\ \underline{\hat{y}}_{i,j} - \alpha_s^2,\ \hat{\bar{x}}_{i,j} + \alpha_s^3,\ \hat{\bar{y}}_{i,j} + \alpha_s^4\} \tag{10}$$

6: Yield bounding box prediction regions $\mathcal{I}(\hat{B}_{n+1,j}) \leftarrow [\underline{\hat{B}}_{n+1,j}, \hat{\bar{B}}_{n+1,j}]$

Backup materials

Algorithm 2 Computing bounding box dissimilarity scores

Require: a detection threshold ρ_{th}, an overlap threshold IoU_{th}, a trained object detector f_θ, a calibration dataset D_{cal}.

```
for X_i ∈ D_cal do
    Predict the bounding boxes: Ŷ_i = f_θ(X_i),
    for B_{i,j} ∈ Y_i, B̂_{i,j} ∈ Ŷ_i do
        if IoU(B_{i,j}, B̂_{i,j}) ≥ IoU_th and ρ_{i,j} ≥ ρ_th then
            Pair B_{i,j} with B̂_{i,j}
        end if
        α_{i,j} ← {|x̂_{i,j} − x_{i,j}|, |ŷ_{i,j} − y_{i,j}|, |x̄_{i,j} − x̂̄_{i,j}|, |ȳ_{i,j} − ŷ̄_{i,j}|}
    end for
end for
```

where line 7 reads: $\alpha_{i,j} \leftarrow \{|\hat{\underline{x}}_{i,j} - \underline{x}_{i,j}|, |\hat{\underline{y}}_{i,j} - \underline{y}_{i,j}|, |\overline{x}_{i,j} - \hat{\overline{x}}_{i,j}|, |\overline{y}_{i,j} - \hat{\overline{y}}_{i,j}|\}$