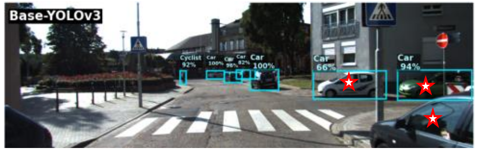
BEA: Revisiting anchor-based object detection DNN using Budding Ensemble Architecture

Qutub Syed Sha[1, 2], Neslihan Kose[1], Rafael Rosales[1], Michael Paulitsch[1], Korbinian Hagn[1], Florian Geissler[1], Yang Peng[1], Gereon Hinz[2], Alois Knoll[2]

1. Intel Labs, Munich Germany, 2. Technical University of Munich, Munich Germany

Motivation

Deterministic object detection models and ensembles struggle with confidence score calibration and objects on corner.

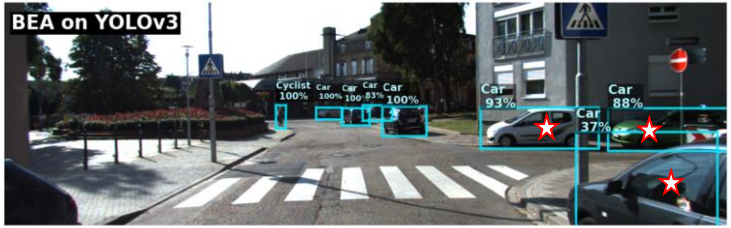


★ Example objects to compare base and BEA model performance

Contributions

- **Budding Ensemble Architecture (BEA)** outperforms state of the art models in terms of accuracy.
- **AP50-based retention curves** are introduced to measure the quality of calibration for object detection models.
- **Novel Tandem loss function** is introduced to the BEA model which increases the overall accuracy by ~6% and OOD detection at least by ~300%.

BEA results



The BEA shows better out-of-distribution image detection than the vanilla and ensemble models.

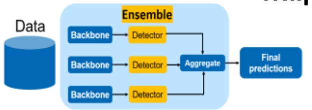
BEA applied to **Validation** dataset

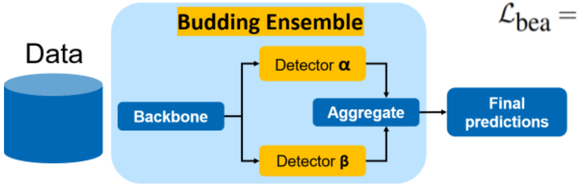
BEA applied to **shifted** dataset

Models (input size 416×416)	mAP_{raw} (%) ↑	AP50 ↑		UE (%) ↓	AP50-based Retention curve AUC (%) ↑	Out-of-distribution detection (OOD) AUC-ROC (%)↑		
		$AP50_{raw}$	$AP50_{\mathcal{U}_{pred}}$			CityPersons $\mathcal{U}_{near-ood}$	BDD100K $\mathcal{U}_{near-ood}$	COCO $\mathcal{U}_{far-ood}$
Base-YOLOv3	51.72	87.4	78.2	11.96	53.1	35*	40.16*	20.21
YOLOv3 3 Ensemble	54.58	89	82.94	9.23	58.7	28.79*	32.44*	20.5*
YOLOv3 5 Ensemble	**55.1**	89.27	82.97	9.03	59.3	28.6*	12.19*	10.21*
BEA-YOLOv3	**54.83 ± 0.28**	**89.3 ± 0.28**	**85.79 ± 0.13**	**4.55 ± 0.02**	**73.9 ± 1.1**	**98.75 ± 2.3**	**86.71 ± 1.7**	**97.33 ± 0.9**

BEA method

"Budding Ensemble Architecture" uses a common backbone to feed an ensemble of detectors trained with our novel tandem loss function.





Tandem loss function

$$\mathcal{L}_{tandem} = \mathcal{L}_{tq} + \mathcal{L}_{ta}$$

$$\mathcal{L}_{bea} = \mathcal{L}_{conv} + \mathcal{L}_{tandem}$$

$$\mathfrak{L}_{tq}(\hat{\phi}) = \Sigma_{i=1}^{S^2} \Sigma_{j=1}^{B} \mathbb{1}_{ij}^{noobj} \frac{2}{\sqrt{(\hat{\phi}_i^{\alpha} - \hat{\phi}_i^{\beta})^2}},$$

$$\mathfrak{L}_{ta}(\hat{\phi}) = \Sigma_{i=1}^{S^2} \Sigma_{j=1}^{B} \mathbb{1}_{ij}^{obj} \frac{\sqrt{(\hat{\phi}_i^{\alpha} - \hat{\phi}_i^{\beta})^2}}{2},$$

$\mathfrak{L}_{tq}(\hat{\phi})$ Tandem quelling loss

$\mathfrak{L}_{ta}(\hat{\phi})$ Tandem aiding loss

$\mathcal{L}_{conv}$ Existing conventional loss function of the model