6 Supplementary Material: BEA

6.1 Evaluation of Tandem Loss on Confidence Prediction in BEA using YOLOv3:

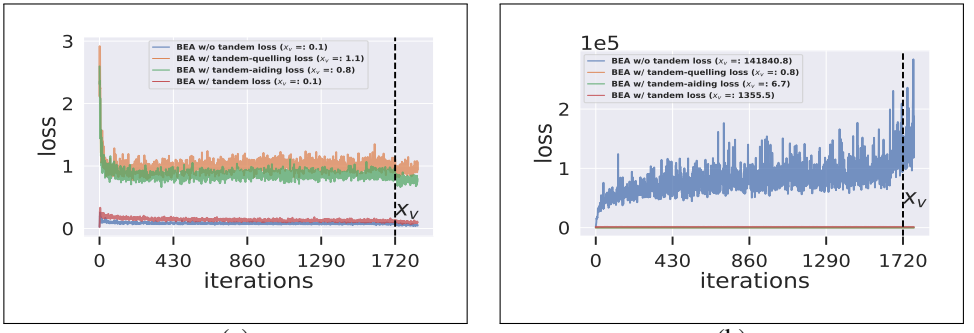
Figure 8 illustrates the impact of Tandem Loss-$\mathcal{L}_{\textbf{tandem}}$ on confidence prediction loss, which is one of the regression losses of YOLOv3. The experiment was conducted on Budding-Ensemble Architecture (BEA) to evaluate the usability of Tandem Loss. The loss of predicted confidence score reflects the model's proficiency in detecting an object by a particular anchor. Figure 8 displays the monitoring of $\mathcal{L}_{\textbf{ta}}$, while Figure 8 displays the monitoring of $\mathcal{L}_{\textbf{tq}}$ losses. The application and monitoring of these losses depend on the configuration indicated in the legends of the corresponding figures. Figure 8 demonstrates that the BEA with Tandem Loss configuration shows superior performance in reducing the variance of respective positive predictions between α and β detectors by decreasing the $\mathcal{L}_{\textbf{ta}}$ compared to the separate factors of $\mathcal{L}_{\textbf{tandem}}$ loss, including the absence of $\mathcal{L}_{\textbf{tandem}}$ loss. This is also shown by the data points along a vertical line denoted by X_V in the legends. Similarly, the Figure in 8 illustrates that the BEA without the Tandem Loss results in the decreased variance of negative predictions between the α and β detectors, which is not desirable as we prefer to have high variance between the corresponding negative predictions. In Figure 8, it is worth noting that the BEA configuration with individual factors of $\mathcal{L}_{\textbf{tandem}}$ loss produces a preferred result in increasing the variance between negative predictions. However, this also hurts $\mathcal{L}_{\textbf{ta}}$ loss as shown in Figure 8. This indicates that using both $\mathcal{L}_{\textbf{ta}}$ and $\mathcal{L}_{\textbf{tq}}$ together leads to better performance in improving the confidence scores of True Positives and reducing the scores of False Positives. This results in better-calibrated prediction outcomes.

6.2 Extended Ablation Study of Table 1 with OOD results on YOLOv3

$\mathcal{L}_{\text{tandem}}$		$mAP_{raw}\uparrow$	$AP50_{raw}\uparrow$	$AP50_{\mathcal{U}_{pred}}\uparrow$	UE (%) $\downarrow$	Retention Curve AUC (%) $\uparrow$	Out-of-distribution detection (OOD) AUC-ROC (%)$\uparrow$		
$\mathcal{L}_{\text{ta}}$	$\mathcal{L}_{\text{tq}}$						CityPersons $\mathcal{U}_{near-ood}$	BDD100K $\mathcal{U}_{near-ood}$	COCO $\mathcal{U}_{far-ood}$
✗	✗	52.36	87.96	79.11	9.65	56.9	81.05	77.21	94.81
✓	✗	54.07	88.56	79.44	11.05	54.2	77.01	75.97	93.63
✗	✓	54.82	88.31	82.15	9.03	57.1	72.8	73.79	91.61
✓	✓	**54.83**	**89.2**	**85.79**	**4.55**	**73.9**	**98.75**	**86.71**	**97.33**

Table 2: Extended Ablation study on BEA-YOLOv3 with on YOLOv3: Effects on the accuracy, uncertainty error, calibration and out-of-distribution shown by $\mathcal{L}_{\text{tandem}}$.

Table 2 provides an extended ablation study to understand the dependency of $\mathcal{L}_{\textbf{ta}}$ and $\mathcal{L}_{\textbf{tq}}$ loss functions individually. This comprehensive study presents out-of-distribution (OOD) results for various configurations of Tandem Loss ($\mathcal{L}_{\textbf{tandem}}$). The table highlights the importance of adding $\mathcal{L}_{\textbf{ta}}$ and $\mathcal{L}_{\textbf{tq}}$ Tandem Loss together leading to improved prediction accuracy and calibration of confidence score, higher uncertainty estimates, and better OOD detection performance.



(a) (b)

Figure 8: Monitoring objectness-based losses - Impact of Tandem Loss ($\mathcal{L}_{\mathbf{tandem}}$) across various BEA-YOLOv3 configurations) : (a) Monitoring of $\mathcal{L}_{\mathbf{ta}}$ on predicted confidence scores for positive predictions from the Tandem Detectors; (b) Monitoring of $\mathcal{L}_{\mathbf{tq}}$ on predicted confidence scores for negative predictions from the Tandem Detectors

Model	Parameters (MM)
Base-YOLOv3	61.5
M-Ensemble YOLOv3	M×61.5
BEA-YOLOv3	82.53
Base-SSD	25.5
M-Ensemble SSD	M×25.5
BEA-SSD	30.63
BEA-SSD Arch2	27.5

Table 3: Computational-Overhead Analysis of BEA

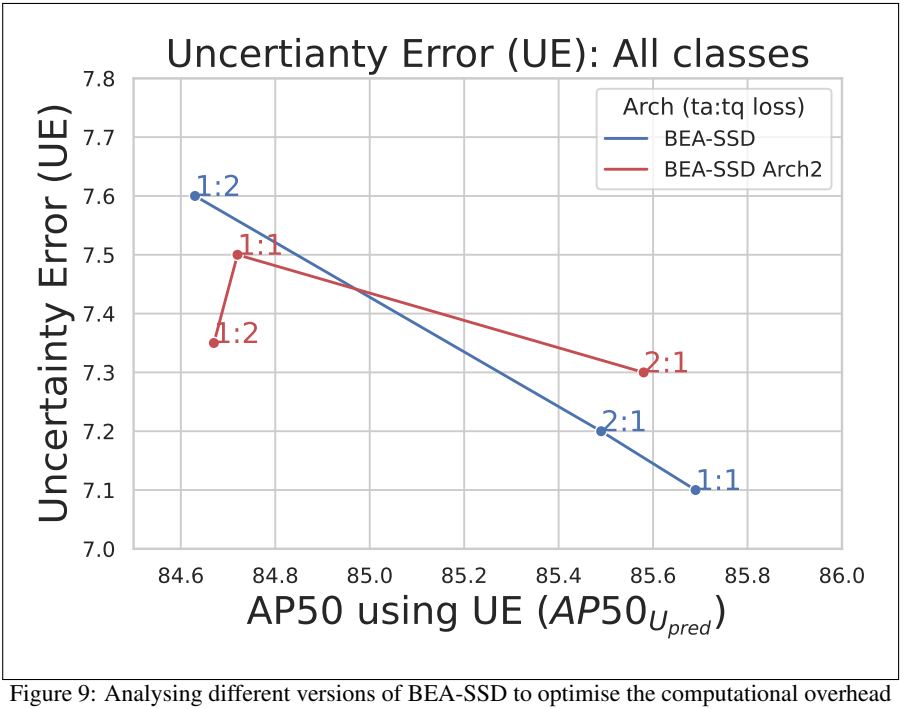
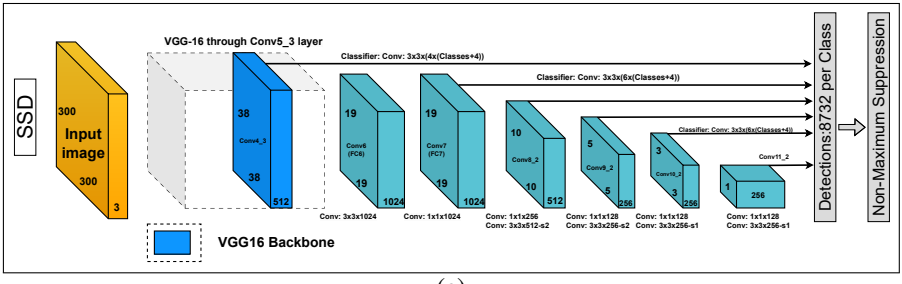
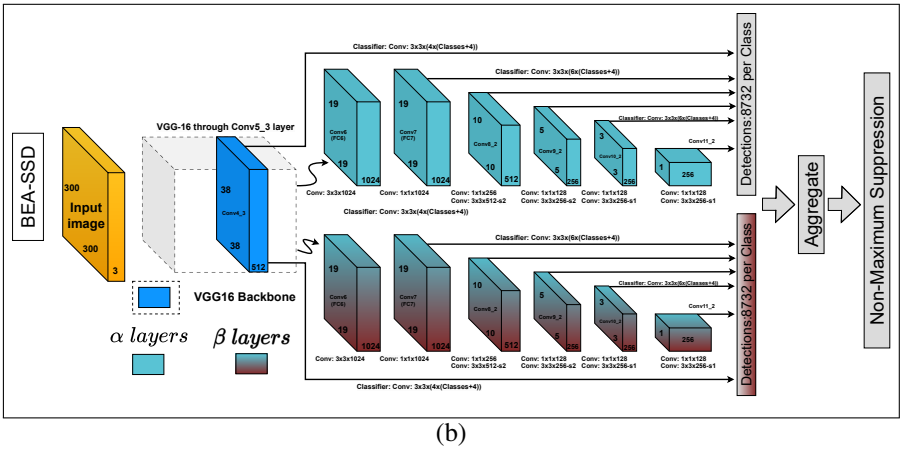


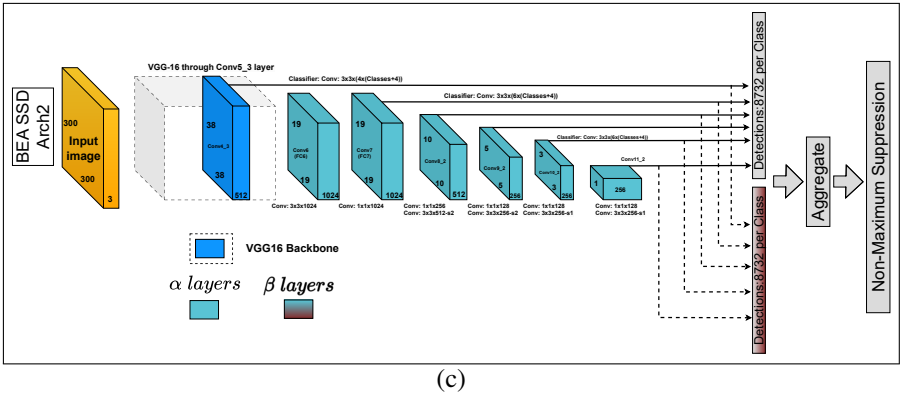
Figure 9: Analysing different versions of BEA-SSD to optimise the computational overhead



(a)



(b)



(c)

Figure 10: Different versions of SSD : (a) Base-SSD (Vanilla version); (b) BEA-SSD; (c) BEA-SSD Arch 2

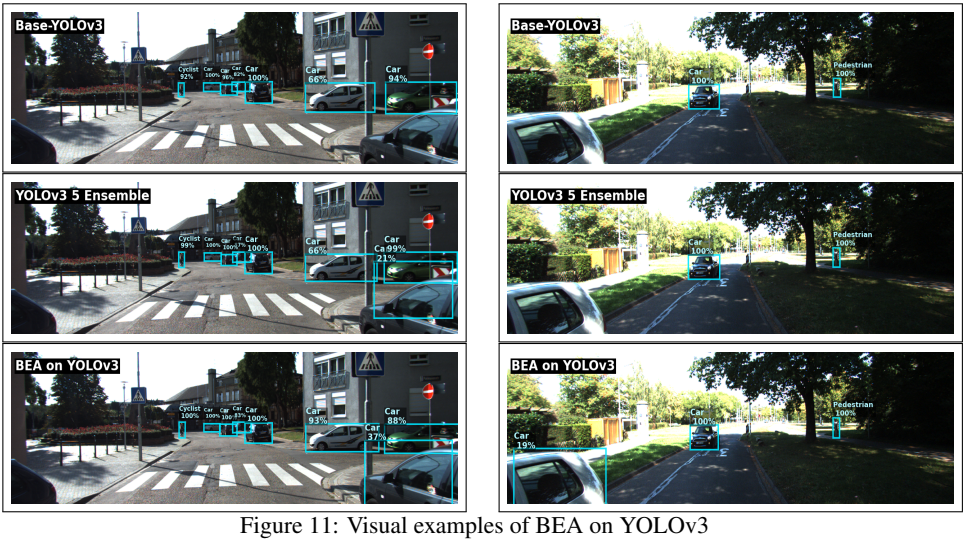


Figure 11: Visual examples of BEA on YOLOv3

6.3 SSD architecture with BEA:

This section explores the potential of splitting the architecture during the creation of BEA (Backend for Analytics) to further minimize overhead. We specifically examine the possibility of splitting the architecture at a later stage, beyond the backbone, to not only reduce computational overhead but also maintain similar levels of uncertainty estimation and calibration. The BEA-SSD results shown in Table 1 by duplicating the entire detector after VGG backbone as shown in Figure 10-b. The BEA-SSD has 20.1% more parameters than the Base-SSD. The Figure 10-c is optimised version of BEA and Table 3 shows that BEA-SSD Arch2 has just 7.8% more parameters than the Base-SSD. By employing the same loss functions and ratios, the performance of BEA-SSD Arch2 experiences a noticeable decline. However, it is possible to restore the performance to the original BEA-SSD levels by effectively controlling the ratios of $\mathcal{L}_{\mathbf{ta}}$ and $\mathcal{L}_{\mathbf{tq}}$, as illustrated in Figure 9.

6.4 Visual Inspection of Budding-Ensemble Architecture's performance: Figure 11, 12, 13

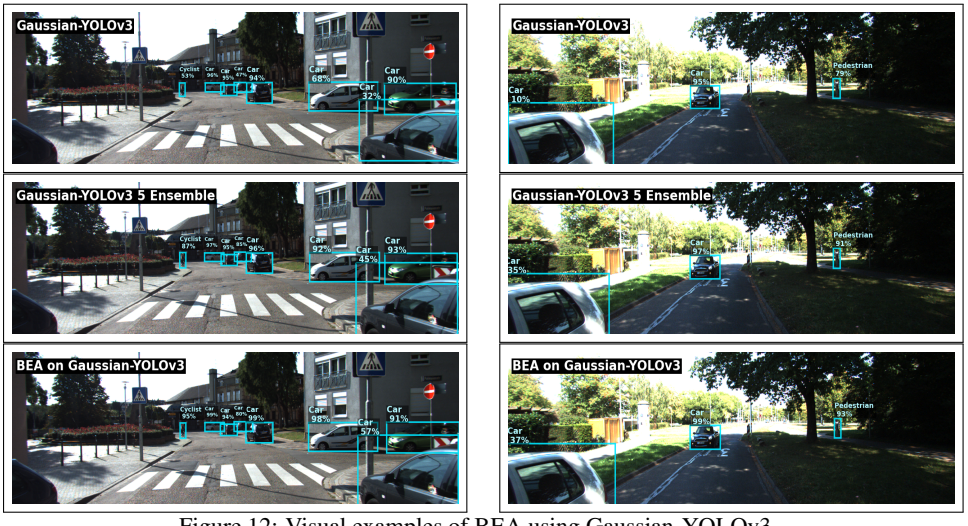


Figure 12: Visual examples of BEA using Gaussian-YOLOv3

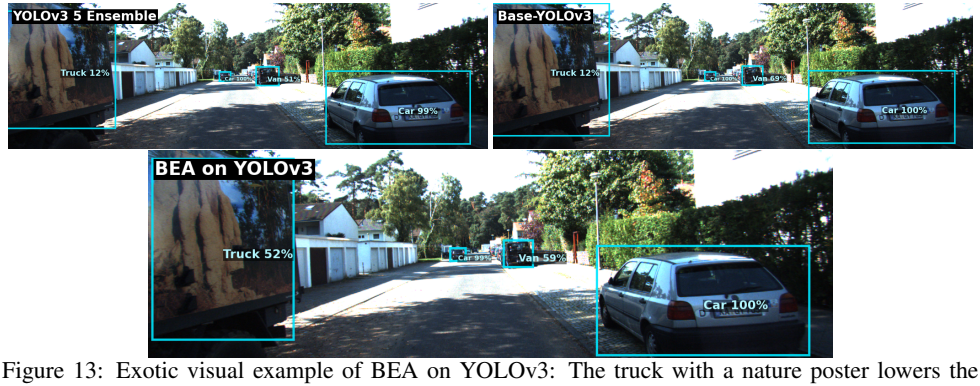


Figure 13: Exotic visual example of BEA on YOLOv3: The truck with a nature poster lowers the confidence score of the prediction in non BEA models.