

Crowd Counting

Team D – Project 10

Thomas Le Monnier de Gouville

Simon Lembeye

Robin Michard

Victor Rambaud

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

Objectif

Tâche : estimer le nombre de personnes présentes sur une image de foule.

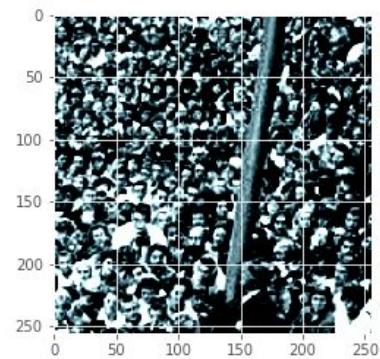
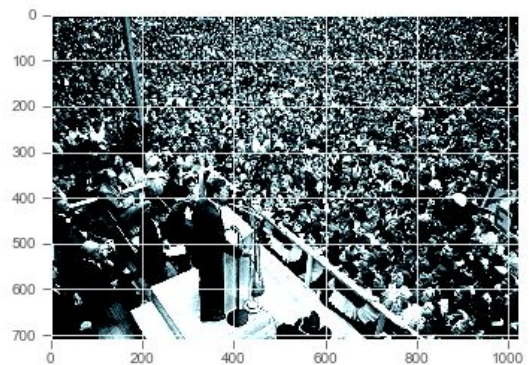
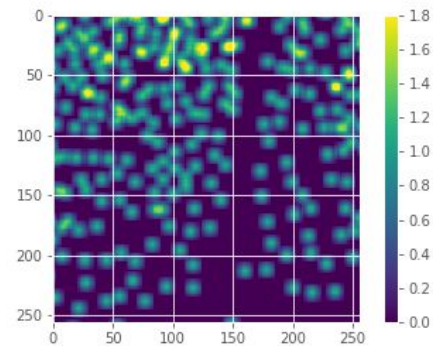
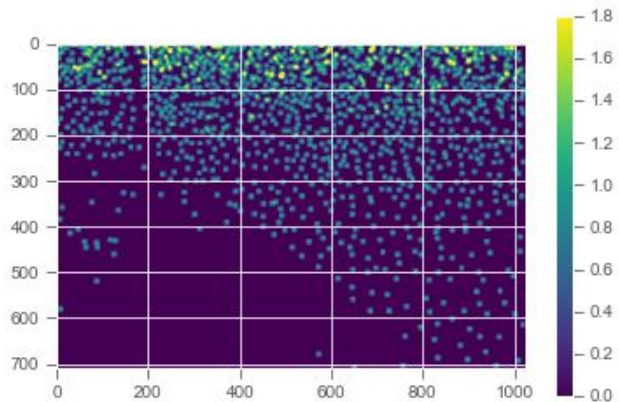
Objectif du projet : Confronter les méthodes / articles suivants sur cette tâche.

- **Ground Truth Density** - CSRNet: Dilated Convolutional Neural Networks for Understanding the Highly Congested Scenes, CVPR 2018.
- **Bayesian Loss** - Bayesian Loss for Crowd Count Estimation with Point Supervision.
- **Aleatoric Loss** - What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?

Ground Truth Density

- Estimation d'une densité prédéterminée
- Utilisation d'un Deep CNN - **CSRNet**
- Structure du vgg16 avec des couches supplémentaires.

Configurations of CSRNet			
A	B	C	D
input(unfixed-resolution color image)			
front-end (fine-tuned from VGG-16)			
conv3-64-1			
conv3-64-1			
max-pooling			
conv3-128-1			
conv3-128-1			
max-pooling			
conv3-256-1			
conv3-256-1			
conv3-256-1			
max-pooling			
conv3-512-1			
conv3-512-1			
conv3-512-1			
back-end (four different configurations)			
conv3-512-1	conv3-512-2	conv3-512-2	conv3-512-4
conv3-512-1	conv3-512-2	conv3-512-2	conv3-512-4
conv3-512-1	conv3-512-2	conv3-512-2	conv3-512-4
conv3-256-1	conv3-256-2	conv3-256-4	conv3-256-4
conv3-128-1	conv3-128-2	conv3-128-4	conv3-128-4
conv3-64-1	conv3-64-2	conv3-64-4	conv3-64-4
conv1-1-1			



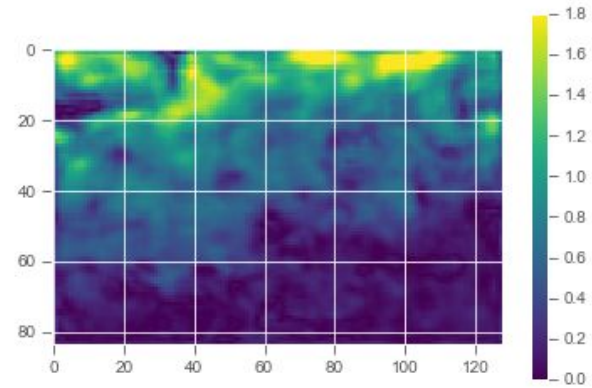
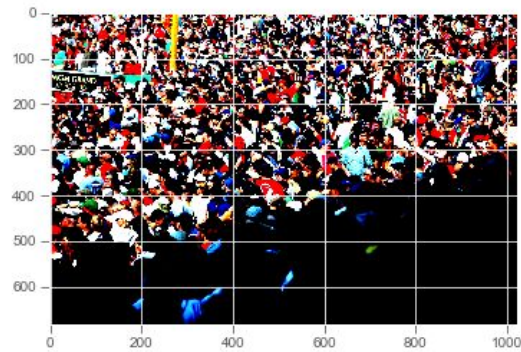
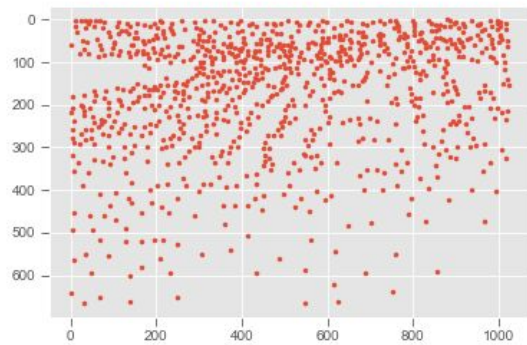
Visualisation des données - Ground Truth

Bayesian Loss

- Utilise uniquement les **coordonnées des têtes** pour calculer la loss
- Calcule la probabilité qu'une tête soit associée à chaque pixel
- Utilisation possible d'un “**background**”
- L'article d'origine utilise un réseau **vgg19** avec une extension

$$\begin{aligned} E[c_n] &= E\left[\sum_{m=1}^M c_n^m\right] = \sum_{m=1}^M E[c_n^m] \\ &= \sum_{m=1}^M p(y_n | \mathbf{x}_m) \mathbf{D}^{est}(\mathbf{x}_m). \end{aligned}$$

$$\mathcal{L}^{Bayes} = \sum_{n=1}^N \mathcal{F}(1 - E[c_n])$$



Visualisation des données – Bayes Loss

Aleatoric Loss

- **Méthode générique** - Loss adaptable aux deux réseaux précédents
- L'incertitude aléatoire capture le bruit inhérent aux observations
- Le réseau "apprend" le terme s_i qui permet de minimiser l'influence des outliers, en augmentant leur incertitude

$$\mathcal{L}_{BNN}(\theta) = \frac{1}{D} \sum_i \frac{1}{2} \exp(-s_i) \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|^2 + \frac{1}{2} s_i.$$

Dataset – ShanghaiTech A

- **Dataset référence** sur cette tâche.
- Echantillonnage **aléatoire**

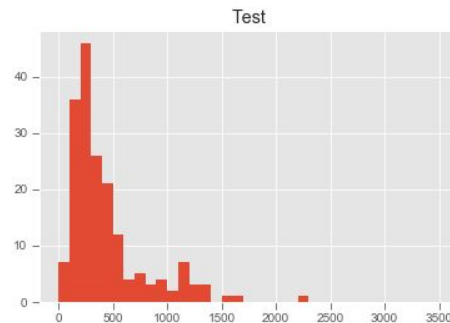
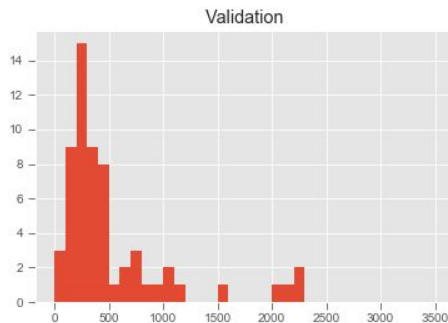
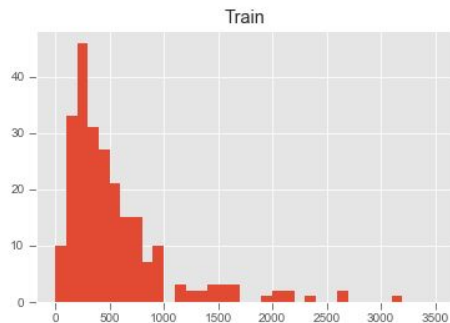
Métriques :

- Mean Square Error
- Mean Absolute Error

Train set : 240 images - moyenne = 545

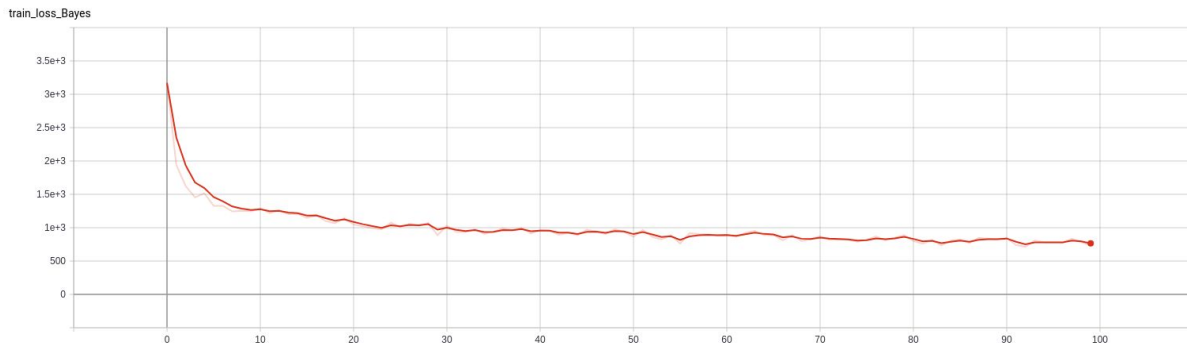
Validation set : 60 images - moyenne = 523

Test set : 126 images - moyenne = 423



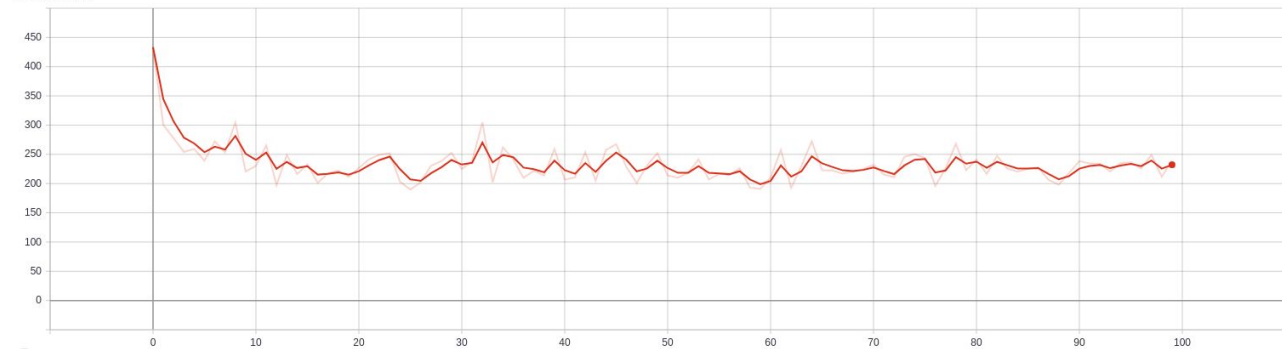
Training Method

- Data Augmentation : Crop + Flip Aléatoire
- Choix des hyperparamètres avec des trains courts (50 epochs) sur chaque methodes.
- Métrique d'évaluation : $2 * \text{MSE} + \text{MAE}$
- Les poids du réseau sont conservés pour l'époch donnant la meilleure métrique d'évaluation sur le validation set.

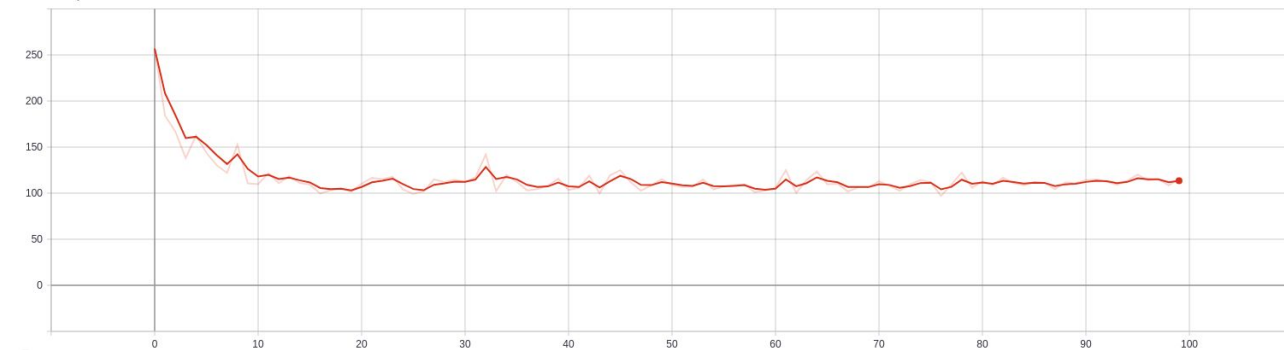


Example : Bayes+ – CSRnet Training

val_MSE_Bayes



val_MAE_Bayes



Résultats

Méthode	Ground Truth	Bayes	Bayes	Bayes+	Bayes Aleatoric
Réseau	CSRNet	CSRNet	vgg19	CSRNet	CSRNet
Validation MSE	454.9	189.8	189.9	187.2	157.6
Validation MAE	338.8	104.1	99.22	102.6	96.4
Test MSE	308.8	116,24	120.6	116.3	115.5
Test MAE	265.98	79,3	84,0	77.0	77.8