INSTITUT NATIONAL DE L'INFORMATION GÉOGRAPHIQUE ET FORESTIÈRE

l'information grandeur nature

# Segmentation sémantique de grands nuages de points avec Superpoint Graph

## JR 2018

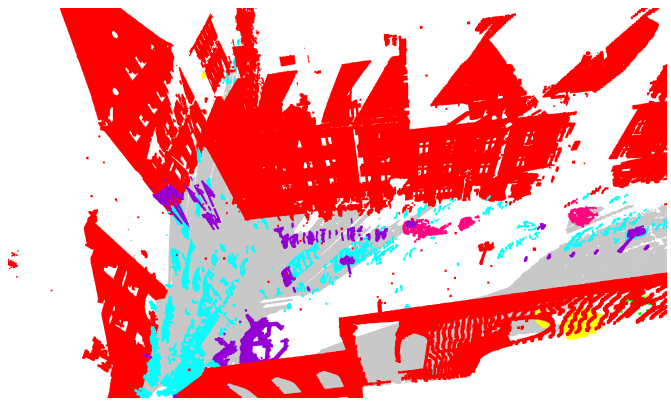
Loic Landrieu

MATIS - IGN

Mars 2018

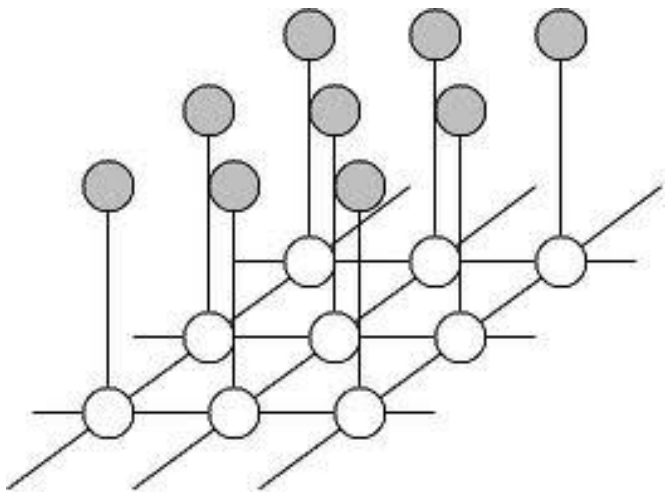
# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense



# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense
- Approche traditionelle: prédiction point par point + regularisation

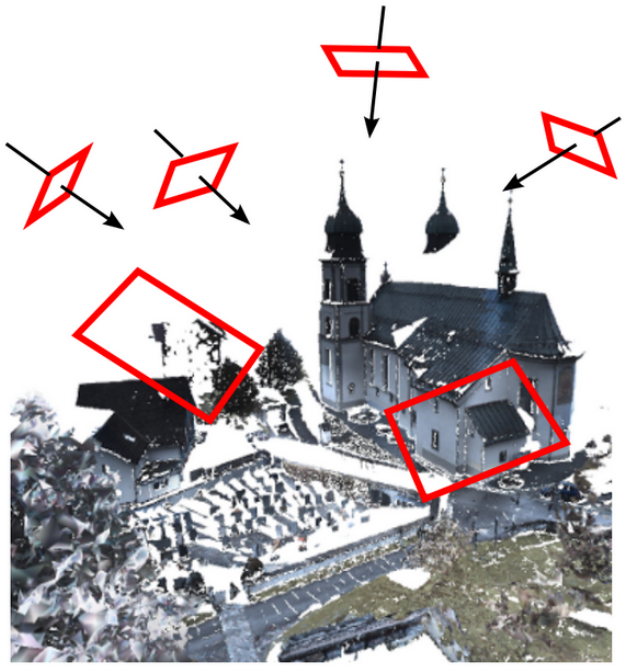


# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense
- Approche traditionelle: prédiction point par point + regularisation
- La révolution Deep Learning:

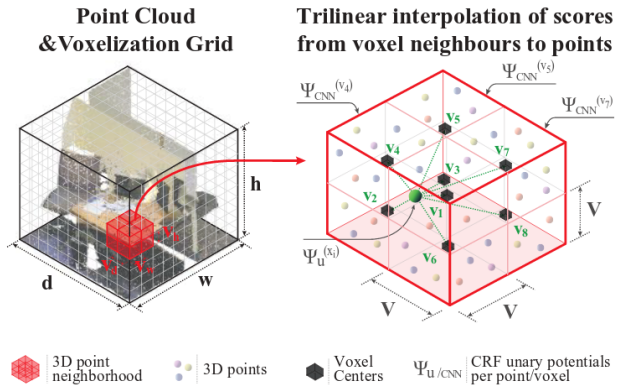
# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense
- Approche traditionelle: prédiction point par point + regularisation
- La révolution Deep Learning:
- SnapNet: efficace mais 3D $\rightarrow$ 2D



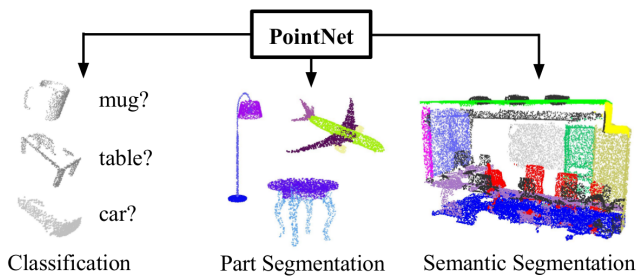
# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense
- Approche traditionelle: prédiction point par point + regularisation
- La révolution Deep Learning:
  - SnapNet: efficace mais 3D $\rightarrow$ 2D
  - SegCloud: efficace mais perd détails



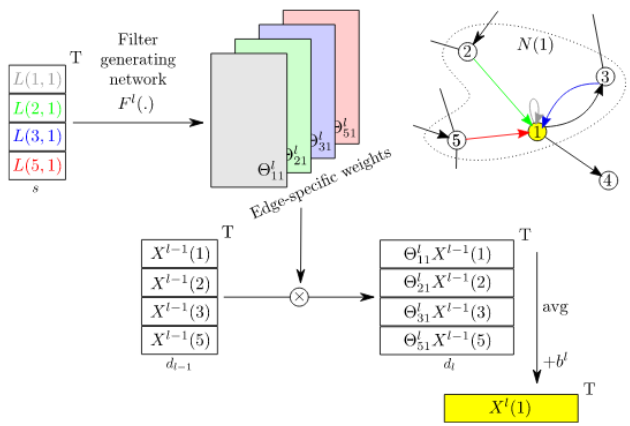
# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense
- Approche traditionelle: prédiction point par point + regularisation
- La révolution Deep Learning:
  - SnapNet: efficace mais 3D $\rightarrow$ 2D
  - SegCloud: efficace mais perd détails
  - PointNet: précis mais coûteux



# Segmentation sémantique à grande échelle

- **Objectif:** prédiction dense
- Approche traditionelle: prédiction point par point + regularisation
- La révolution Deep Learning:
  - SnapNet: efficace mais 3D $\rightarrow$ 2D
  - SegCloud: efficace mais perd détails
  - PointNet: précis mais coûteux
  - ECC: structuré mais coûteux



# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points

# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points
- **Idée:** découper le nuage de point en formes simples

# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points
- **Idée:** découper le nuage de point en formes simples
- **Avantages**:

# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points
- **Idée:** découper le nuage de point en formes simples
- **Avantages**:
  - Formes simples plus faciles à classifier que des points

# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points
- **Idée:** découper le nuage de point en formes simples
- **Avantages**:
  - Formes simples plus faciles à classifier que des points
  - Relations entre formes plus riches qu'entre points

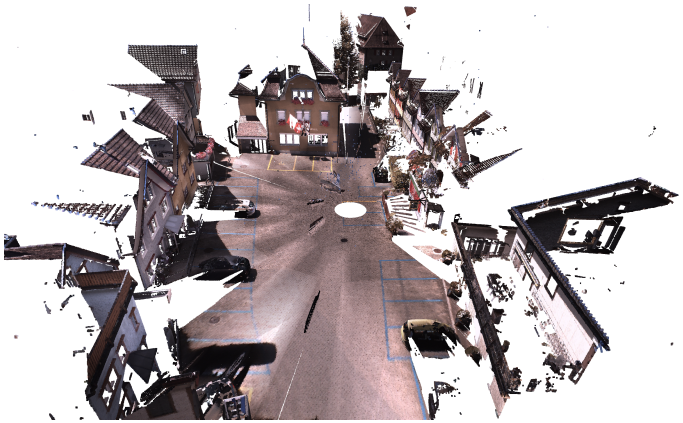
# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points
- **Idée:** découper le nuage de point en formes simples
- **Avantages**:
  - Formes simples plus faciles à classifier que des points
  - Relations entre formes plus riches qu'entre points
  - Beaucoup moins de formes que de points

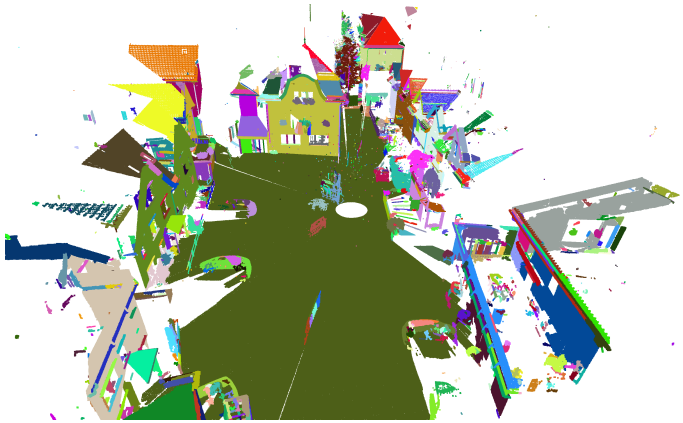
# Principe de SPG

- **Observation:** les méthodes les plus précises ne sont pas appicables pour les grands nuages de points
- **Idée:** découper le nuage de point en formes simples
- **Avantages**:
  - Formes simples plus faciles à classifier que des points
  - Relations entre formes plus riches qu'entre points
  - Beaucoup moins de formes que de points
  - Permet de modéliser les interactions à longues distances

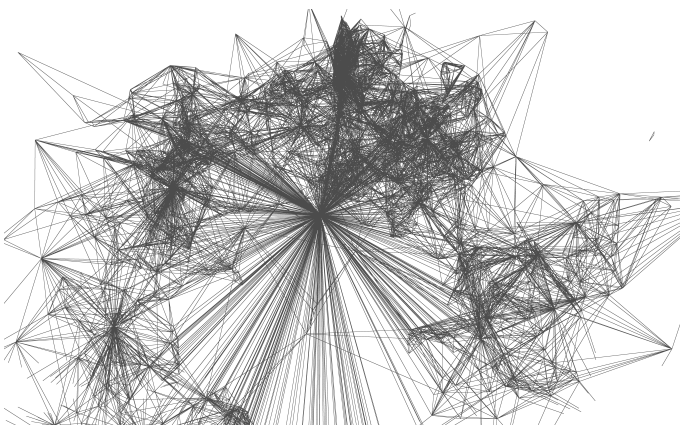
# Etapes de superpoint graph



**(a)** RGB point cloud



**(b)** Simple shapes



**(c)** Superpoint graph

**(d)** Segmentation

# Les 3 sub-problèmes

- Segmentation semantique réduite à 3 sub-problèmes:

# Les 3 sub-problèmes

- Segmentation semantique réduite à 3 sub-problèmes:
- **Partition géometrique :** en formes simples.

  Complexité: très hautes (nuages de $10^8$ points)

  Algorithme: $\ell_0$-cut pursuit

# Les 3 sub-problèmes

- Segmentation semantique réduite à 3 sub-problèmes:

- **Partition géometrique :** en formes simples.

  Complexité: très hautes (nuages de $10^8$ points)

  Algorithme: $\ell_0$-cut pursuit

- **Représentation de formes:** apprentissage de descripteurs

  Complexité: très basse (formes subsamplée à 128 points)

  Algorithme: PointNet

# Les 3 sub-problèmes

- Segmentation semantique réduite à 3 sub-problèmes:

- **Partition géometrique :** en formes simples.

  Complexité: très hautes (nuages de $10^8$ points)

  Algorithme: $\ell_0$-cut pursuit

- **Représentation de formes:** apprentissage de descripteurs

  Complexité: très basse (formes subsamplée à 128 points)

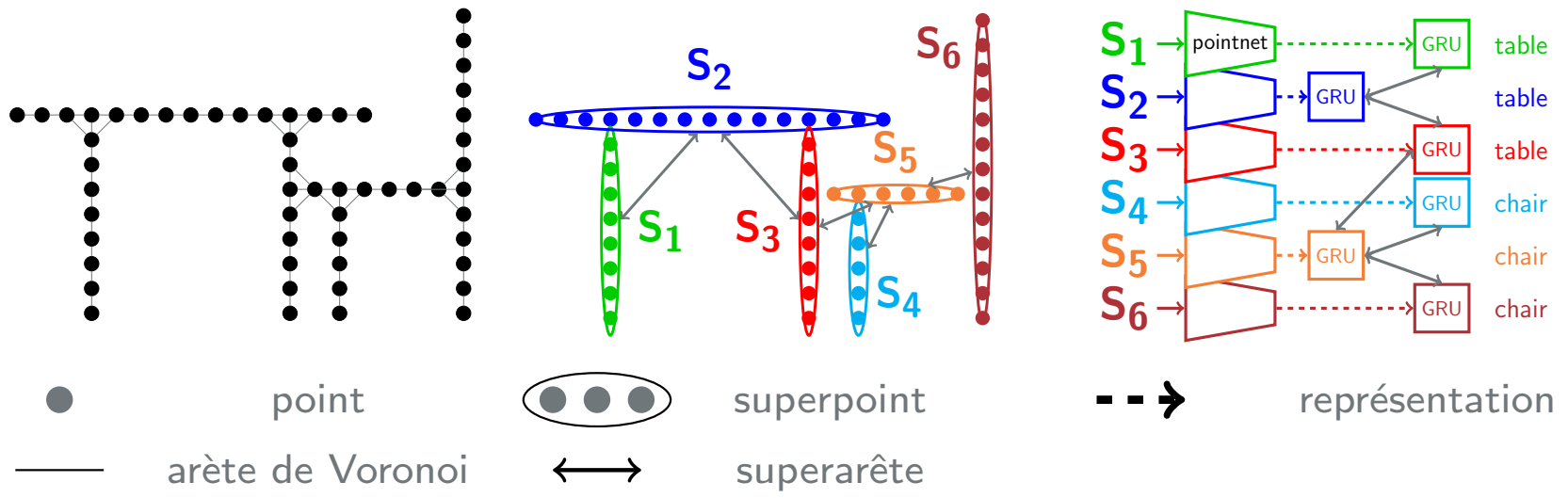
  Algorithme: PointNet

- **Segmentation contextuelle:** utilisation du contexte

  Complexité: faible (superpoint graph $\sim$ 1000 noeuds)

  Algorithme: ECC avec Gated Reccurent Unit (GRU)

# Les 3 sub-problèmes



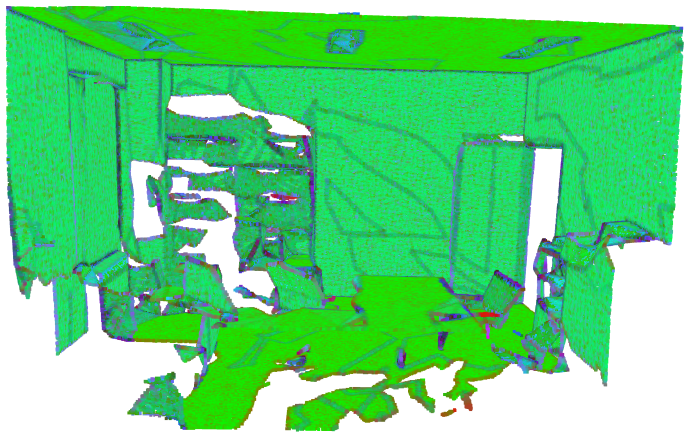

**(a)** Nuage de point **(b)** Superpoint graph **(c)** Réseau

# Partition géometrique

- Nuage de points $C$, position $p_i$, RGB: $o_i$

# Partition géometrique

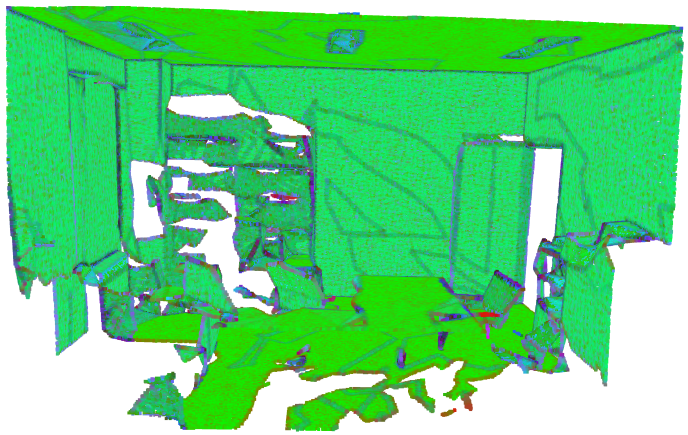
- Nuage de points $C$, position $p_i$, RGB: $o_i$
- **Step1:** Descripteurs géométriques $f_i$



linearity, planarity, verticality

# Partition géometrique

- Nuage de points $C$, position $p_i$, RGB: $o_i$
- **Step1:** Descripteurs géométriques $f_i$

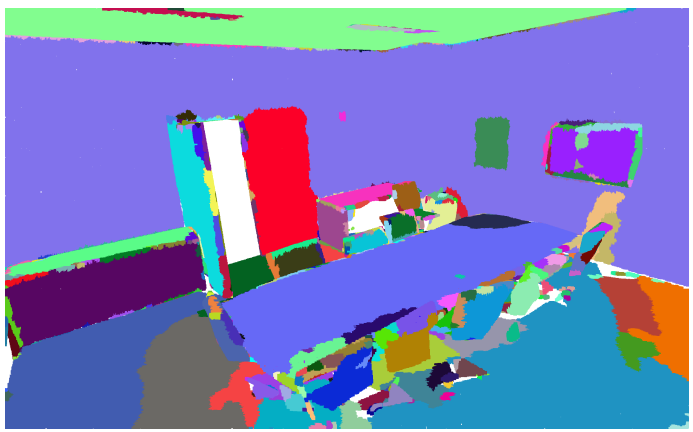


linearity, planarity, verticality

# Partition géometrique

- Nuage de points $C$, position $p_i$, RGB: $o_i$
- **Step1:** Descripteurs géométriques $f_i$
- **Step3:** Approximation constantes par morceaux ($\ell_0$-cut pursuit):

$$\arg\min_{g \in \mathbb{R}^{dn}} \sum_{i \in C} \|g_i - f_i\|^2 + \mu \sum_{(i,j) \in E_{\text{nn}}} [g_i - g_j \neq 0]$$

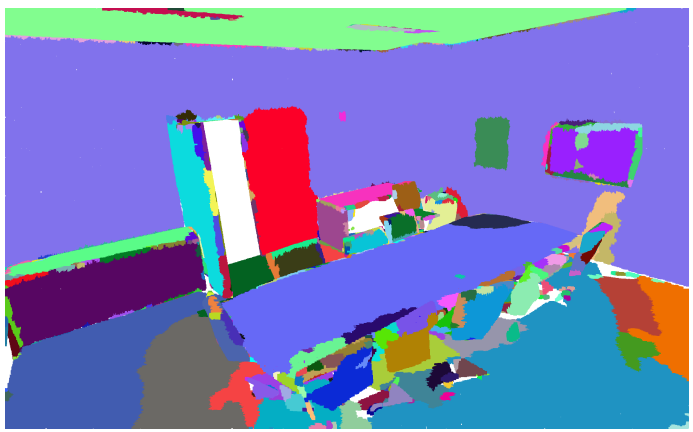


# Partition géometrique

- Nuage de points $C$, position $p_i$, RGB: $o_i$
- **Step1:** Descripteurs géométriques $f_i$
- **Step3:** Approximation constantes par morceaux ($\ell_0$-cut pursuit):

$$\arg\min_{g \in \mathbb{R}^{dn}} \sum_{i \in C} \|g_i - f_i\|^2 + \mu \sum_{(i,j) \in E_{\text{nn}}} [g_i - g_j \neq 0]$$

- Composantes constantes: formes simples.



# Construction du Superpoint Graph

- Superpoint Graph $\mathcal{G} = (\mathcal{S}, \mathcal{E}, F)$

# Construction du Superpoint Graph

- Superpoint Graph $\mathcal{G} = (\mathcal{S}, \mathcal{E}, F)$
- Superpoints $\mathcal{S}$: formes simples

# Construction du Superpoint Graph

- Superpoint Graph $\mathcal{G} = (\mathcal{S}, \mathcal{E}, F)$
- Superpoints $\mathcal{S}$: formes simples
- Superedges $\mathcal{E}$: relation d'adjacence entre superpoints

# Construction du Superpoint Graph

- Superpoint Graph $\mathcal{G} = (\mathcal{S}, \mathcal{E}, F)$
- Superpoints $\mathcal{S}$: formes simples
- Superedges $\mathcal{E}$: relation d'adjacence entre superpoints
- $G_{\text{vor}} = (C, E_{\text{vor}})$ graphe de Voronoi

# Construction du Superpoint Graph

- Superpoint Graph $\mathcal{G} = (\mathcal{S}, \mathcal{E}, F)$
- Superpoints $\mathcal{S}$: formes simples
- Superedges $\mathcal{E}$: relation d'adjacence entre superpoints
- $G_{\text{vor}} = (C, E_{\text{vor}})$ graphe de Voronoi
- $F$: descripteurs des superedges:

| Descripteurs | Taille | Description |
|---|---|---|
| mean offset | 3 | $\text{mean}_{m \in \delta(S,T)}\ \delta_m$ |
| offset deviation | 3 | $\text{std}_{m \in \delta(S,T)}\ \delta_m$ |
| centroid offset | 3 | $\text{mean}_{i \in S}\ p_i - \text{mean}_{j \in T}\ p_j$ |
| length ratio | 1 | $\log \text{length}\,(S)\,/\text{length}\,(T)$ |
| surface ratio | 1 | $\log \text{surface}\,(S)\,/\text{surface}\,(T)$ |
| volume ratio | 1 | $\log \text{volume}\,(S)\,/\text{volume}\,(T)$ |
| point count ratio | 1 | $\log \lvert S \rvert / \lvert T \rvert$ |

# Représentation de formes simples

- Une représentation de taille 32 par forme

# Représentation de formes simples

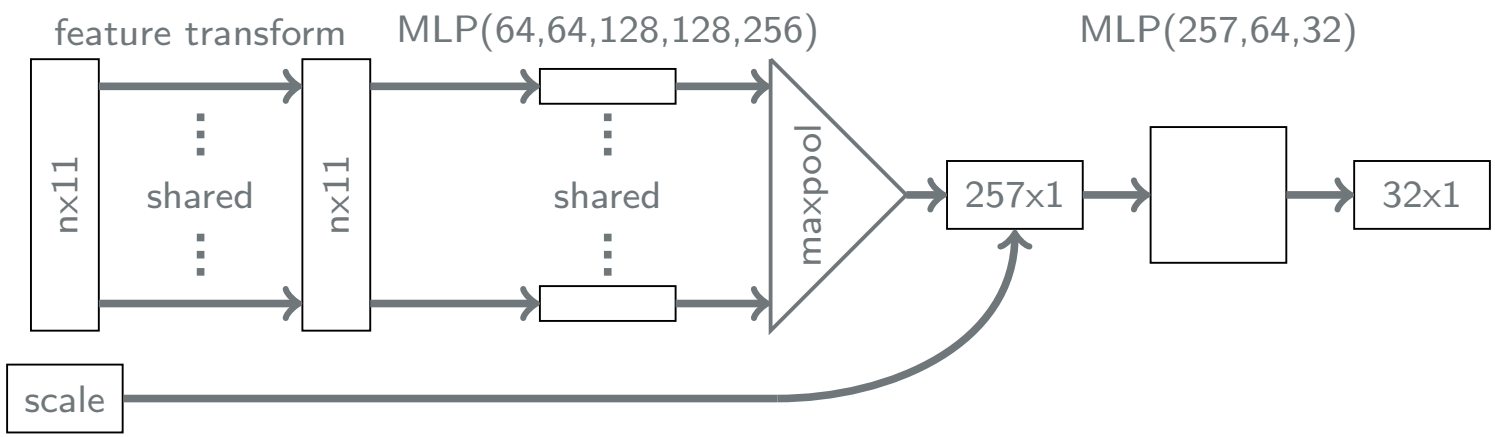
- Une représentation de taille 32 par forme
- Chaque forme est subsamplée à 128 points

# Représentation de formes simples

- Une représentation de taille 32 par forme
- Chaque forme est subsamplée à 128 points
- Chaque point: $(p_i, o_i, f_i)$

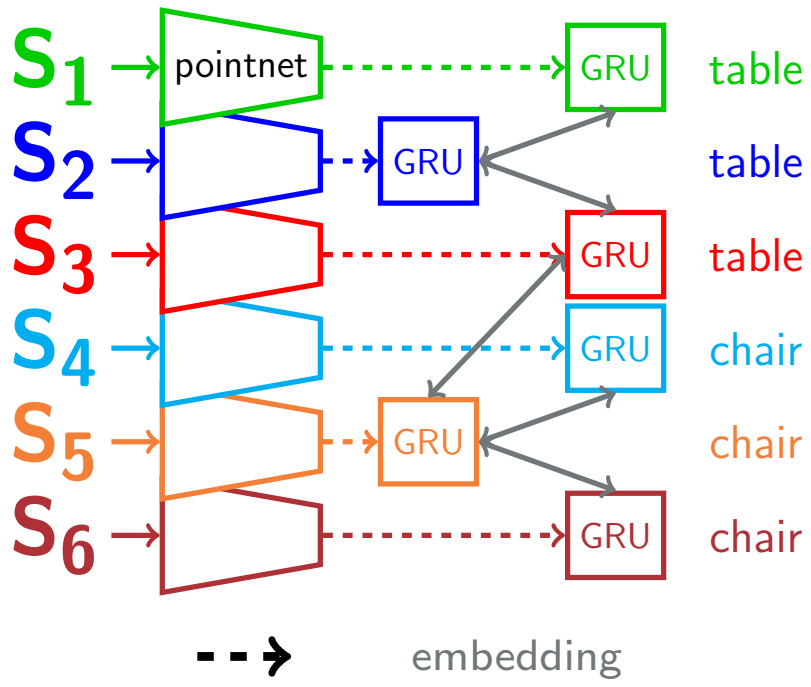
# Représentation de formes simples

- Une représentation de taille 32 par forme
- Chaque forme est subsamplée à 128 points
- Chaque point: $(p_i, o_i, f_i)$
- PointNet réduit pour représenter les formes simples



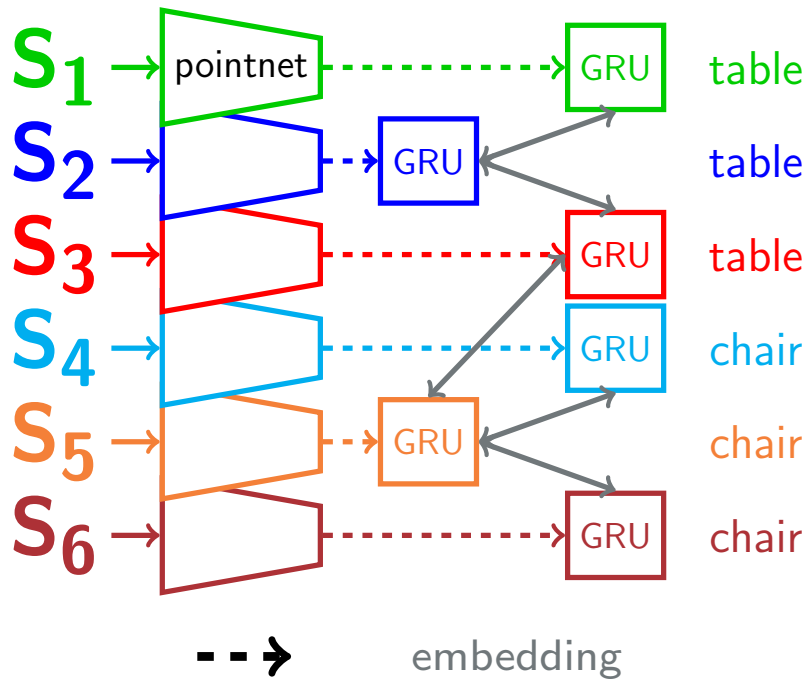
# Segmentation contextuelle

- ECC: passage de messages entre voisins dans $\mathcal{G}$



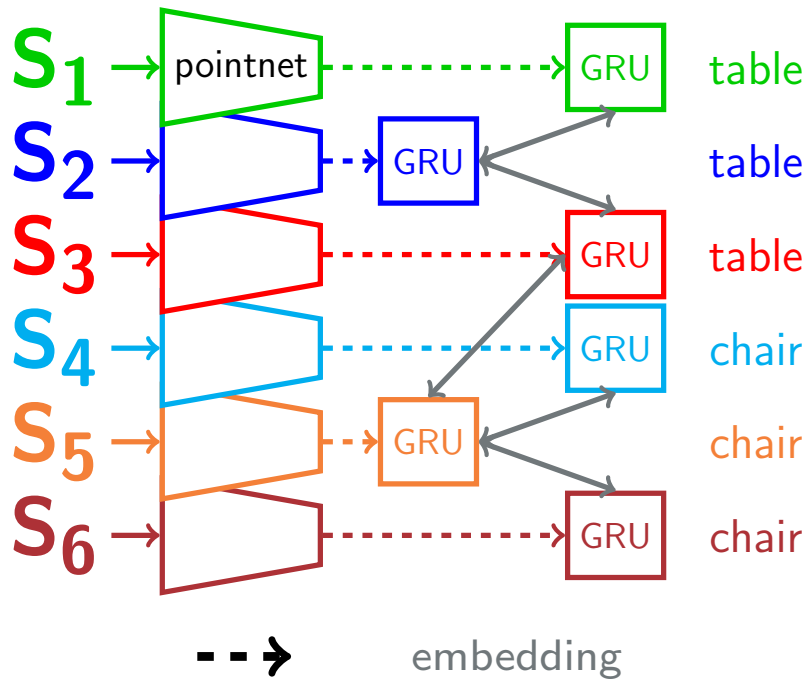
# Segmentation contextuelle

- ECC: passage de messages entre voisins dans $\mathcal{G}$
- Etat interne + descripteurs d'arètes $\rightarrow$ message



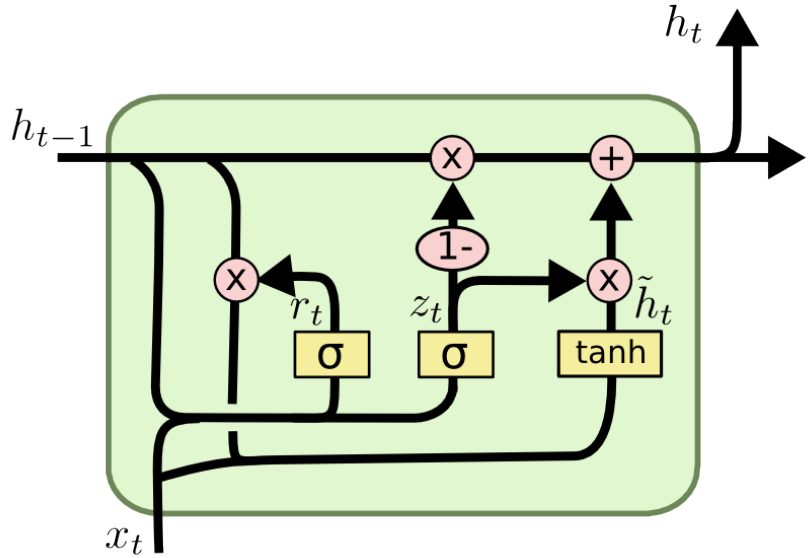
# Segmentation contextuelle

- ECC: passage de messages entre voisins dans $\mathcal{G}$
- Etat interne + descripteurs d'arètes $\rightarrow$ message
- Messages de taille 32



# Segmentation contextuelle

- ECC: passage de messages entre voisins dans $\mathcal{G}$
- Etat interne + descripteurs d'arètes $\rightarrow$ message
- Messages de taille 32
- Gated Recurent Unit avec Input Gating




$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

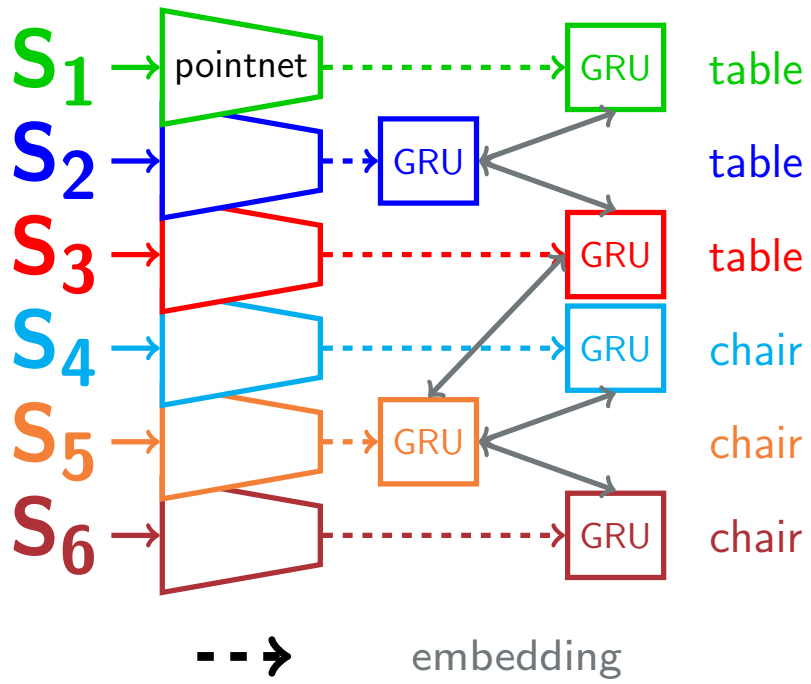
$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

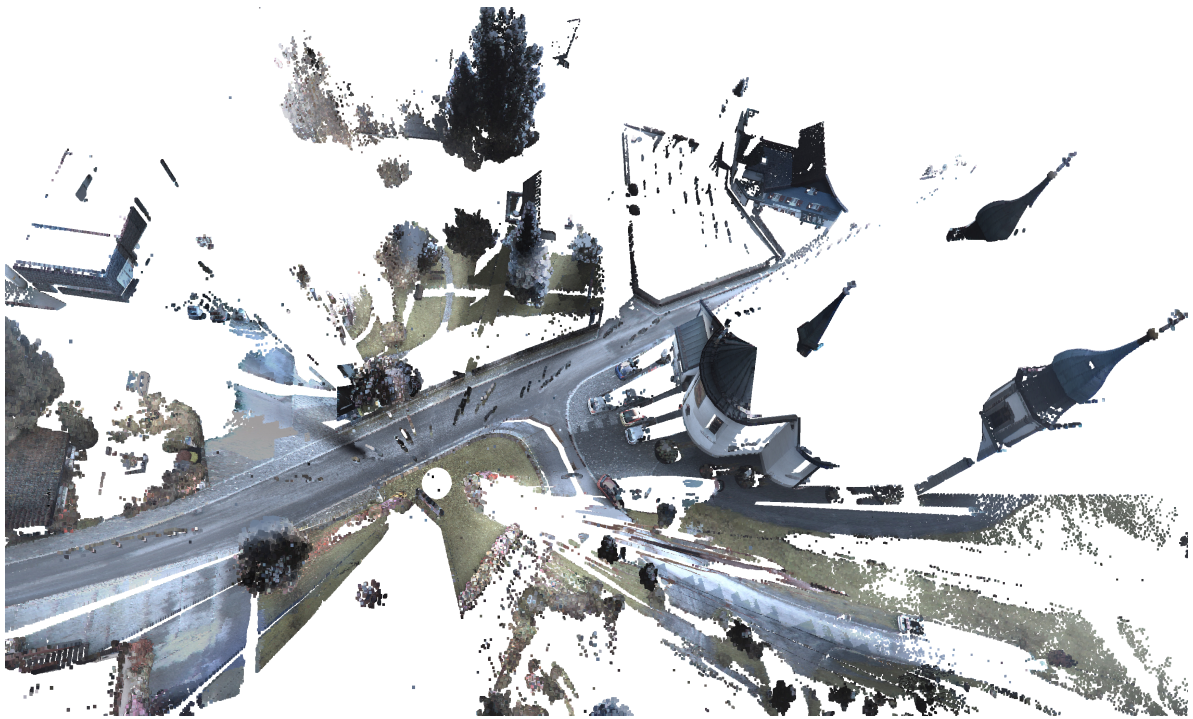
# Segmentation contextuelle

- ECC: passage de messages entre voisins dans $\mathcal{G}$
- Etat interne + descripteurs d'arètes $\rightarrow$ message
- Messages de taille 32
- Gated Recurent Unit avec Input Gating
- Concaténation de 10 iterations = analyse multi-échelle



# Résultat qualitatif results: Semantic3D

Semantic3D: 3 milliards de points sur 30 nuages



- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

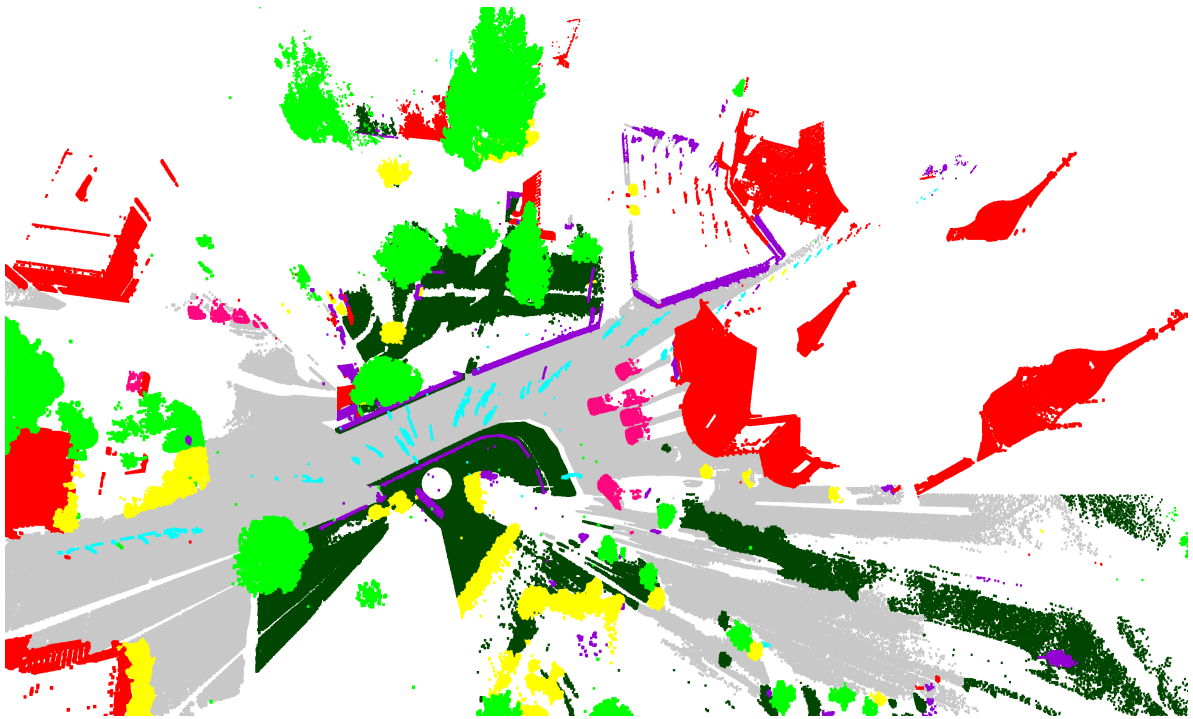
# Résultat qualitatif results: Semantic3D

Semantic3D: 3 milliards de points sur 30 nuages

- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

# Résultat qualitatif results: Semantic3D

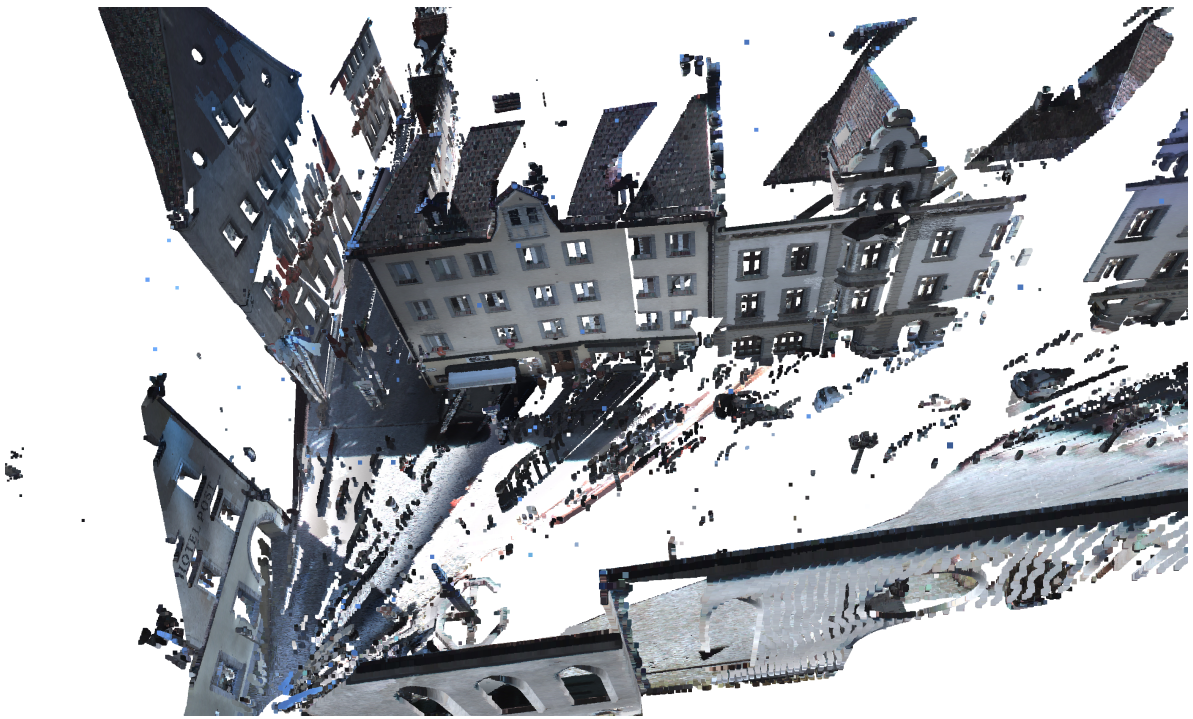
Semantic3D: 3 milliards de points sur 30 nuages



- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

# Résultat qualitatif results: Semantic3D

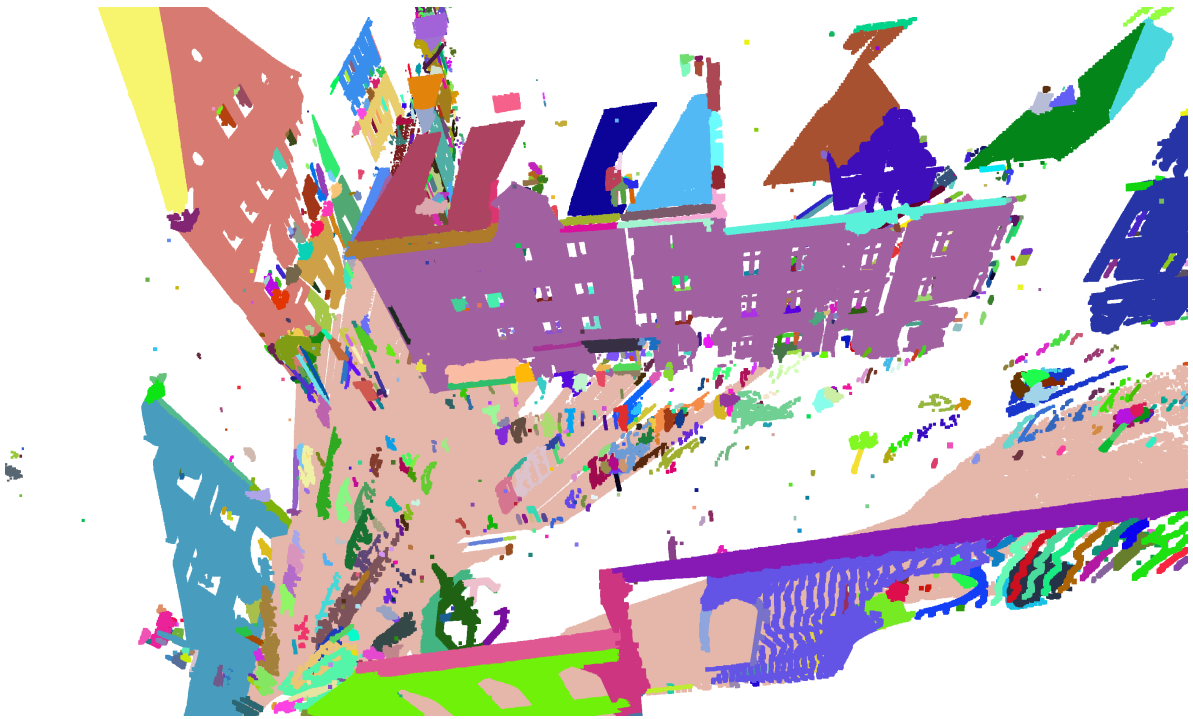
Semantic3D: 3 milliards de points sur 30 nuages



- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

# Résultat qualitatif results: Semantic3D

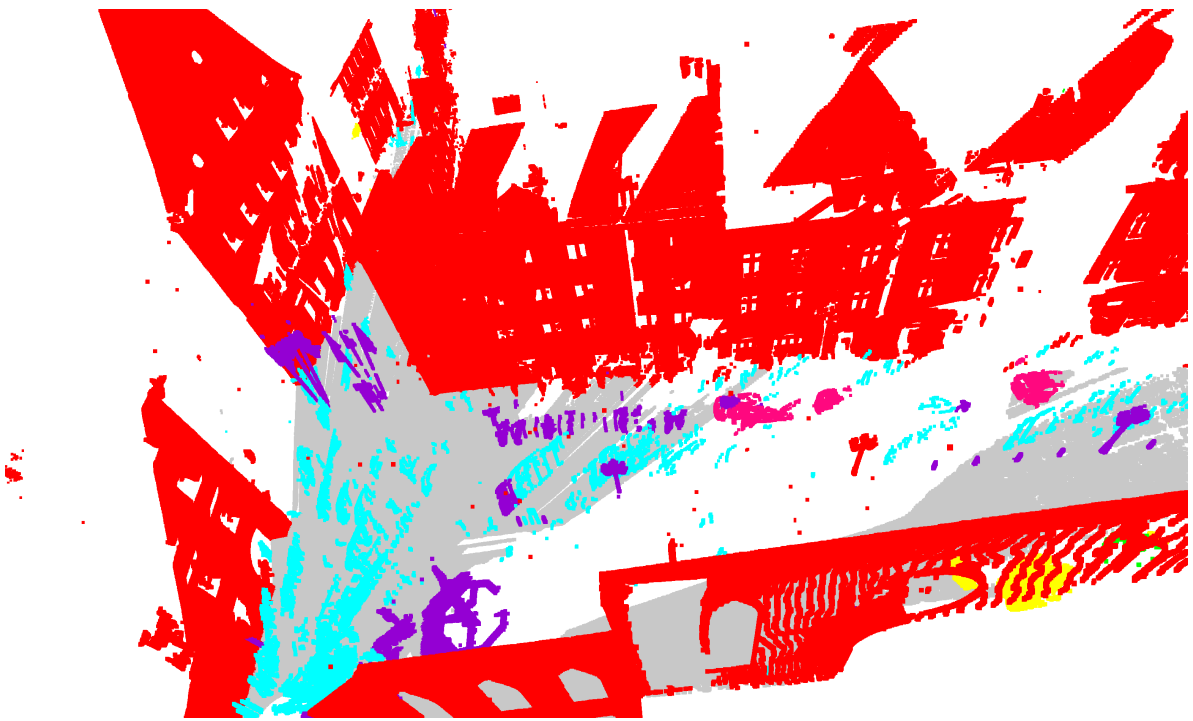
Semantic3D: 3 milliards de points sur 30 nuages



- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

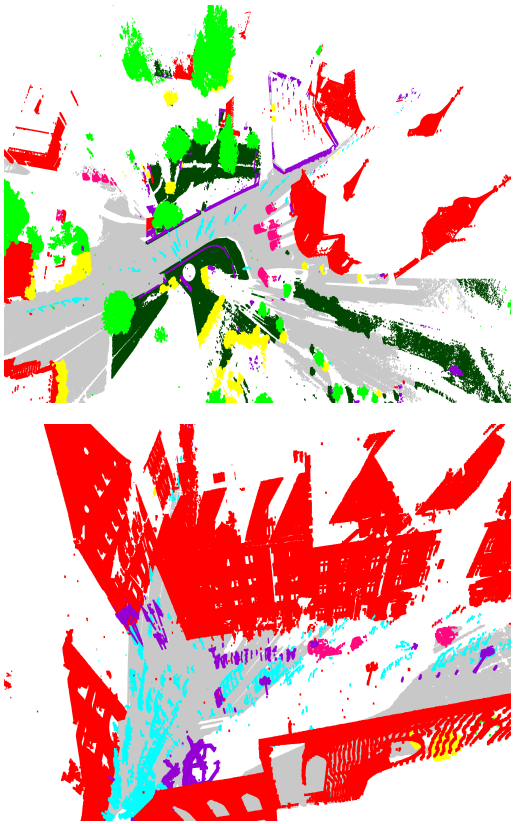
# Résultat qualitatif results: Semantic3D

Semantic3D: 3 milliards de points sur 30 nuages

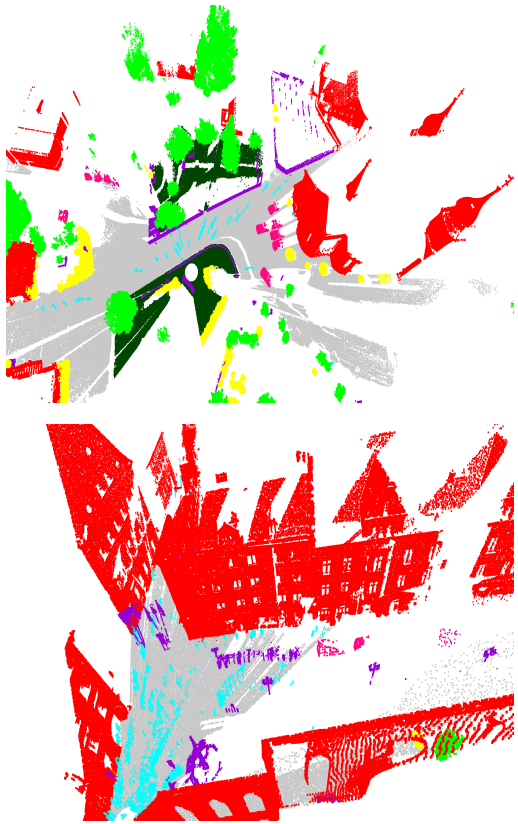


- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

# Résultat qualitatif results: Semantic3D



Prediction



Ground truth

- route
- herbe
- arbre
- buisson
- bâtiment
- aménagement
- artefact
- voiture

# Résultats quantitatifs: Semantic3D

| **Methode** | OA | mIoU | route | herbe | arbre | buisson | bâtiment | aménagement | artefact | voiture |
|---|---|---|---|---|---|---|---|---|---|---|
| reduced test set: 78 699 329 points | | | | | | | | | | |
| TMLC-MSR | 86.2 | 54.2 | 89.8 | 74.5 | 53.7 | 26.8 | 88.8 | 18.9 | 36.4 | 44.7 |
| DeePr3SS | 88.9 | 58.5 | 85.6 | 83.2 | 74.2 | 32.4 | 89.7 | 18.5 | 25.1 | 59.2 |
| SnapNet | 88.6 | 59.1 | 82.0 | 77.3 | 79.7 | 22.9 | 91.1 | 18.4 | 37.3 | 64.4 |
| SegCloud | 88.1 | 61.3 | 83.9 | 66.0 | 86.0 | 40.5 | 91.1 | 30.9 | 27.5 | 64.3 |
| SPG (Ours) | **94.0** | **73.2** | **97.4** | **92.6** | **87.9** | **44.0** | **93.2** | **31.0** | **63.5** | **76.2** |
| full test set: 2 091 952 018 points | | | | | | | | | | |
| TMLC-MS | 85.0 | 49.4 | 91.1 | 69.5 | 32.8 | 21.6 | 87.6 | 25.9 | 11.3 | 55.3 |
| SnapNet | 91.0 | 67.4 | 89.6 | **79.5** | 74.8 | 56.1 | 90.9 | 36.5 | 34.3 | 77.2 |
| SPG (Ours) | **92.9** | **76.2** | **91.5** | 75.6 | **78.3** | **71.7** | **94.4** | **56.8** | **52.9** | **88.4** |

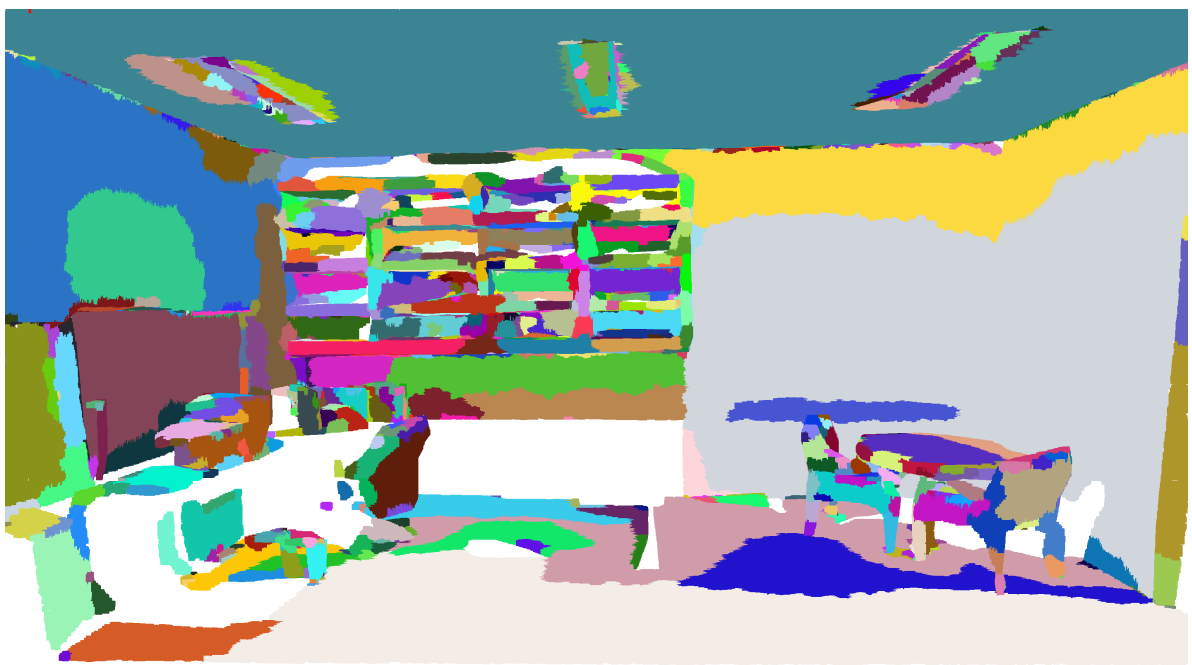
# Résultat qualitatif results: S3DIS

Intérieur, 3 batiments, 6 étages, 200+ pièces, 600 000 000+ points

- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

# Résultat qualitatif results: S3DIS

Intérieur, 3 batiments, 6 étages, 200+ pièces, 600 000 000+ points



- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

# Résultat qualitatif results: S3DIS

Intérieur, 3 batiments, 6 étages, 200+ pièces, 600 000 000+ points

- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

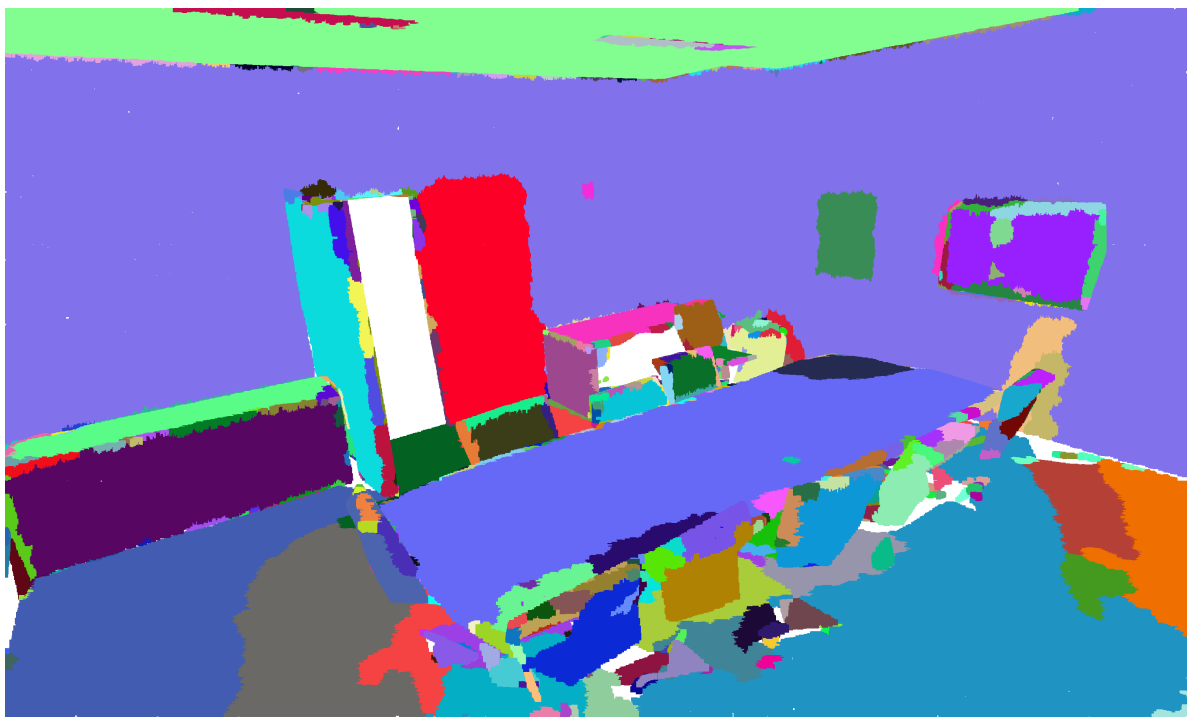
# Résultat qualitatif results: S3DIS

Intérieur, 3 batiments, 6 étages, 200+ pièces, 600 000 000+ points

- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

# Résultat qualitatif results: S3DIS

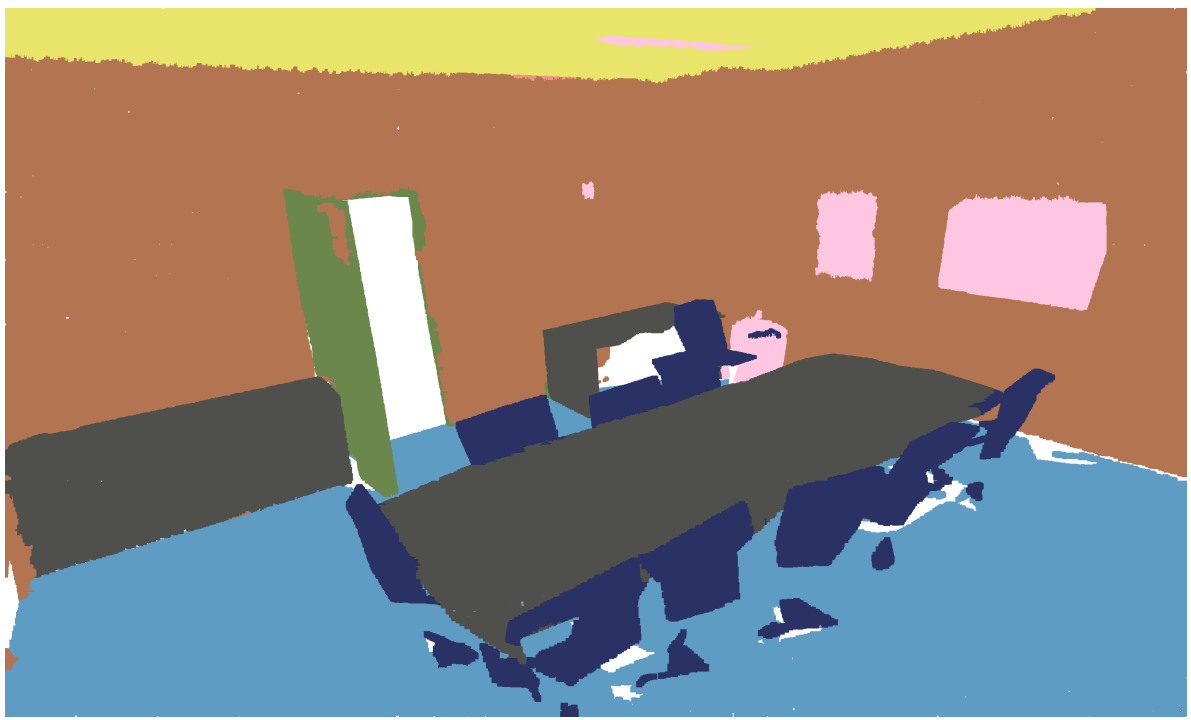
Intérieur, 3 batiments, 6 étages, 200+ pièces, 600 000 000+ points



- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

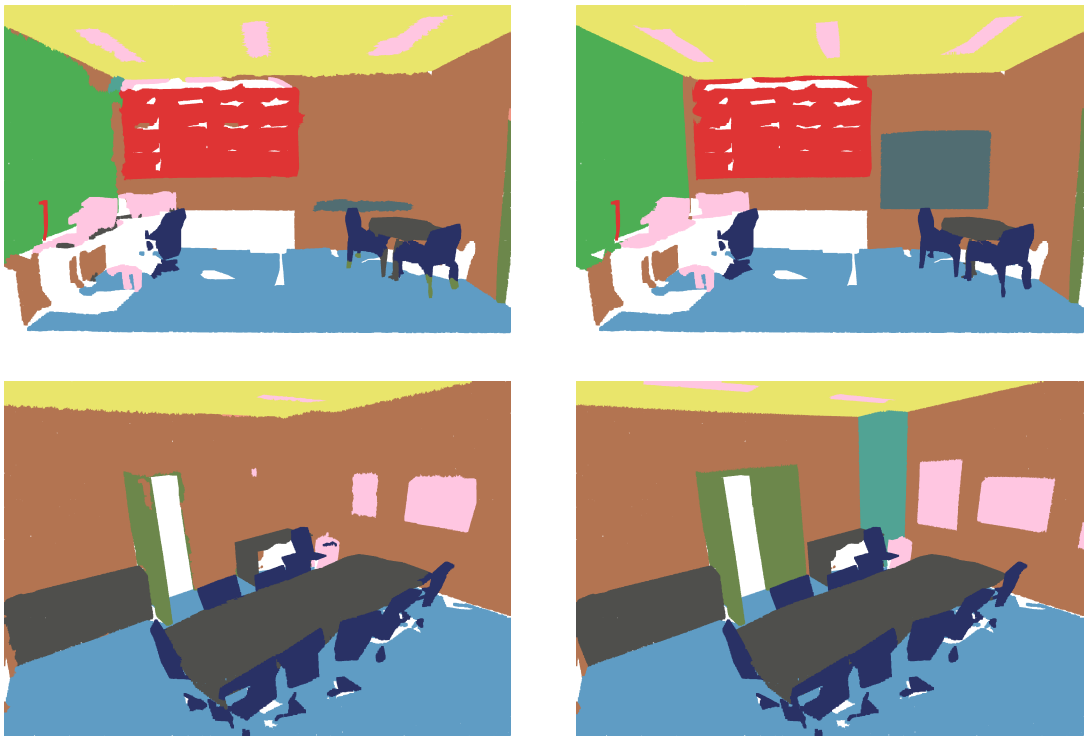
# Résultat qualitatif results: S3DIS

Intérieur, 3 batiments, 6 étages, 200+ pièces, 600 000 000+ points



- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

# Résultats qualitatif: S3DIS



- plafond
- sol
- mur
- colonne
- poutre
- fenêtre
- porte
- table
- chaise
- meuble
- tableau
- autre

# Résultats quantitatif: S3DIS

| **Méthode** | OA | mAcc | mIoU | porte | tableau |
|---|---|---|---|---|---|
| A5 PointNet | – | 48.5 | 41.1 | 10.7 | **26.3** |
| A5 SEGCloud | – | 57.3 | 48.9 | 23.1 | 13.0 |
| **A5 SPG** | 86.4 | **66.5** | **58.0** | **61.5** | 2.1 |
| PointNet | 78.5 | 66.2 | 47.6 | 51.6 | 29.4 |
| Engelmann | 81.1 | 66.4 | 49.7 | 51.2 | **30.0** |
| **SPG** | **85.5** | **73.0** | **62.1** | **68.4** | 8.7 |

# Résultats quantitatif: S3DIS

| **Méthode** | OA | mAcc | mIoU | porte | tableau |
|---|---|---|---|---|---|
| A5 PointNet | – | 48.5 | 41.1 | 10.7 | **26.3** |
| A5 SEGCloud | – | 57.3 | 48.9 | 23.1 | 13.0 |
| **A5 SPG** | 86.4 | **66.5** | **58.0** | **61.5** | 2.1 |
| PointNet | 78.5 | 66.2 | 47.6 | 51.6 | 29.4 |
| Engelmann | 81.1 | 66.4 | 49.7 | 51.2 | **30.0** |
| **SPG** | **85.5** | **73.0** | **62.1** | **68.4** | 8.7 |

| Étape | Nuage entier | 2 cm | 3 cm | 4 cm |
|---|---|---|---|---|
| Voxelisation | 0 | 40 | 24 | 16 |
| Calcul des descripteurs | 439 | 194 | 88 | 43 |
| Partition géometrique | 3428 | 1013 | 447 | 238 |
| Calcul du SPG | 3800 | 958 | 436 | 252 |
| Inférence | 240 | 110 | 60 | 50 |
| Total | 7907 | 2315 | 1055 | 599 |
| mIoU 6-fold | 54.1 | 60.2 | 62.1 | 57.1 |

# Perspectives

- **De nouveaux data sets:** architecure très versatile

# Perspectives

- **De nouveaux data sets:** architecure très versatile
- **Améliorer la partition:** vers une partition sémantique apprise

# Perspectives

- **De nouveaux data sets:** architecure très versatile
- **Améliorer la partition:** vers une partition sémantique apprise
- **Applications à d'autres problèmes:** très grandes images, videos, etc...

# Perspectives

- **De nouveaux data sets:** architecure très versatile
- **Améliorer la partition:** vers une partition sémantique apprise
- **Applications à d'autres problèmes:** très grandes images, videos, etc...
- Repo sur GitHub: `github.com/loicland/superpoint-graph`

# Ablation study

| Modèle | mAcc | mIoU |
|---|---|---|
| Best | 73.0 | 62.1 |
| Perfect | 92.0 | 88.2 |
| Unary | 50.8 | 40.0 |
| iCRF | 51.5 | 40.7 |
| CRF – ECC | 66.4 | 56.4 |
| GRU13 | 69.1 | 58.5 |
| NoInputGate | 68.6 | 57.5 |
| NoConcat | 69.3 | 57.7 |
| NoEdgeFeat | 50.1 | 39.9 |
| ECC – mem | 70.2 | 59.4 |

**Table:** Étude d'ablation des différents choix et apports de cet article (validation croisée sur S3DIS)

# Ablation study

| | | | | |
|---|---|---|---|---|
| *Superpoints max* | 256 | **512** | 1024 | |
| mIoU | 57.9 | 62.1 | 60.4 | |
| *Limit des superarêtes* | 1 m | 5 m | **10 m** | |
| mIoU | 61.0 | 61.3 | 62.1 | |
| *Regularisation* $\mu$ | 0.01 | 0.02 | **0.3** | 0.4 |
| # superpoints | 785 010 | 385 091 | 251 266 | 186 108 |
| perfect mIoU | 90.6 | 88.2 | 86.6 | 85.2 |
| mIoU | 59.1 | 59.2 | 62.1 | 58.8 |
| *Modèle* | LSTM | LSTM+IG | GRU | **GRU+IG** |
| mIoU | 61.0 | 59.6 | 57.5 | 62.1 |
| *Taille des superpoint* | 1-40 | 40-128 | 128-1000 | $\geq$ 1000 |
| proportion | 25.2% | 34.7% | 31% | 9.2% |

**Table:** Évaluation sur S3DIS par validation croisée de différentes configurations. Les choix retenus dans notre méthode sont en gras.