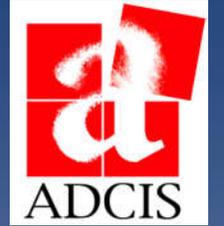


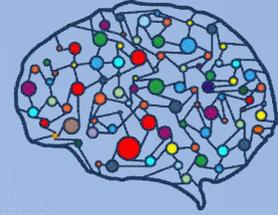
Segmentation automatique et reconstruction de conduites forcées à partir d'images de survol par drone



Aurélien Ducournau
ADCIS S.A.

3 rue Martin Luther King – 14280 Saint-Contest – France





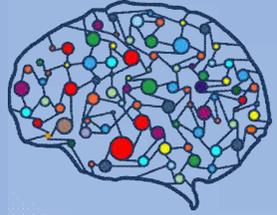
- Surveillance et inspection automatique d'infrastructures via l'analyse d'image acquises par survol de drones
 - Conduites forcées en aval de barrage hydraulique (EDF)
 - Clôtures aéroportuaires (Aéroport Caen-Carpiquet)
 - Lignes électriques moyenne tension (ENEDIS)



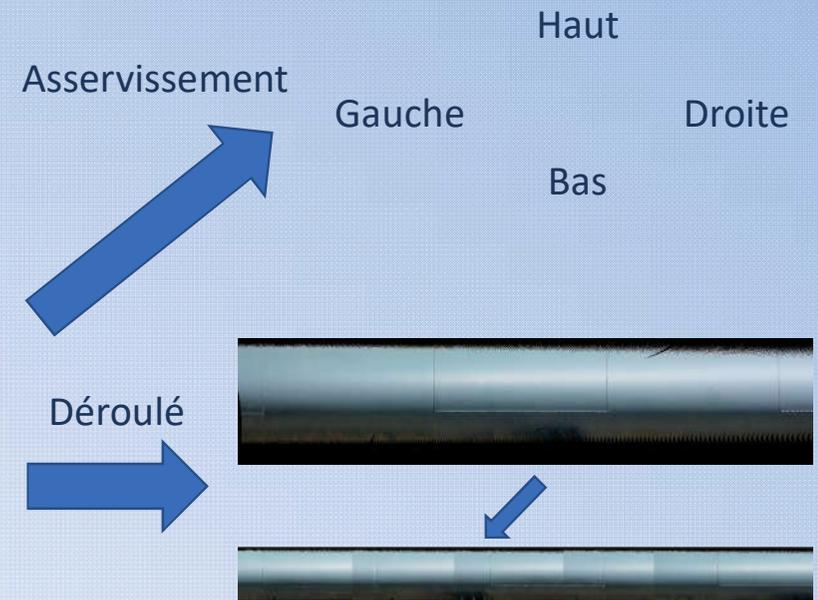
www.ceos-systems.com



Conduites forcées : problématique



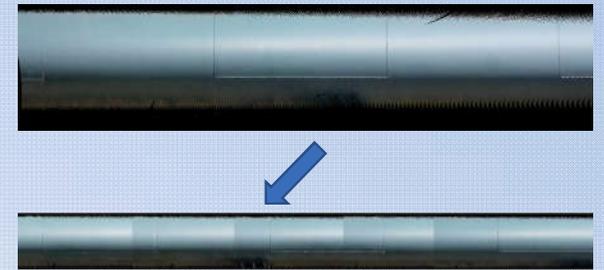
- Segmentation des conduites forcées pour :
 - Asservissement de la nacelle du drone
 - 5 images/seconde maximum pour temps réel
 - Reconstruction de la conduite (déroulé)
 - Pour traitement futur (détections d'anomalies)



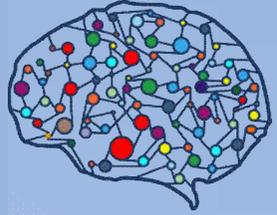
Segmentation



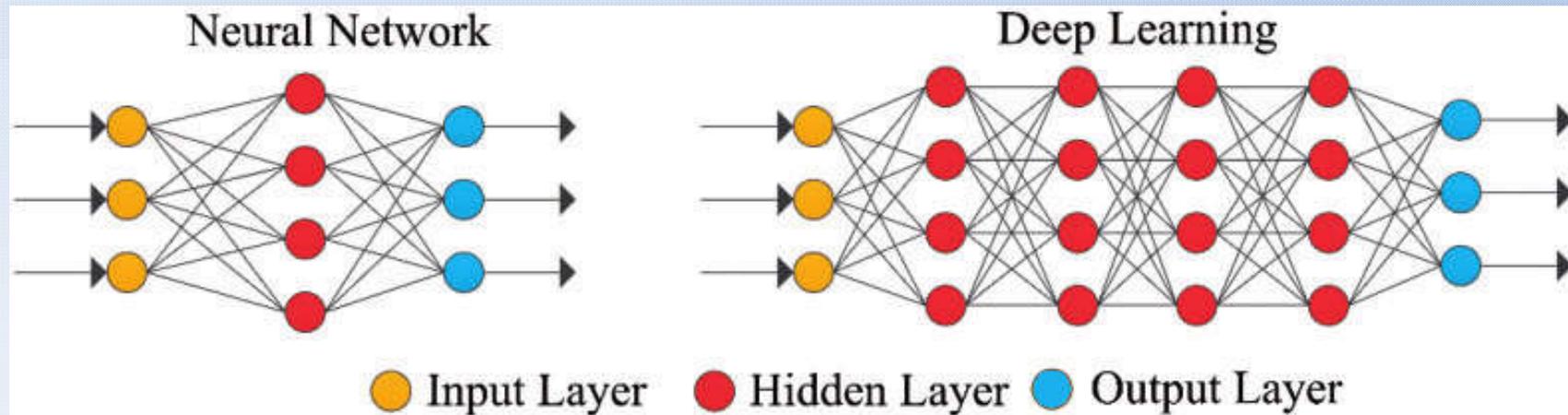
Déroulé



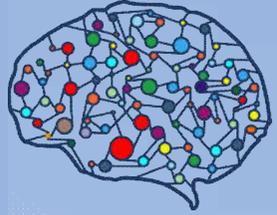
Deep Learning ?



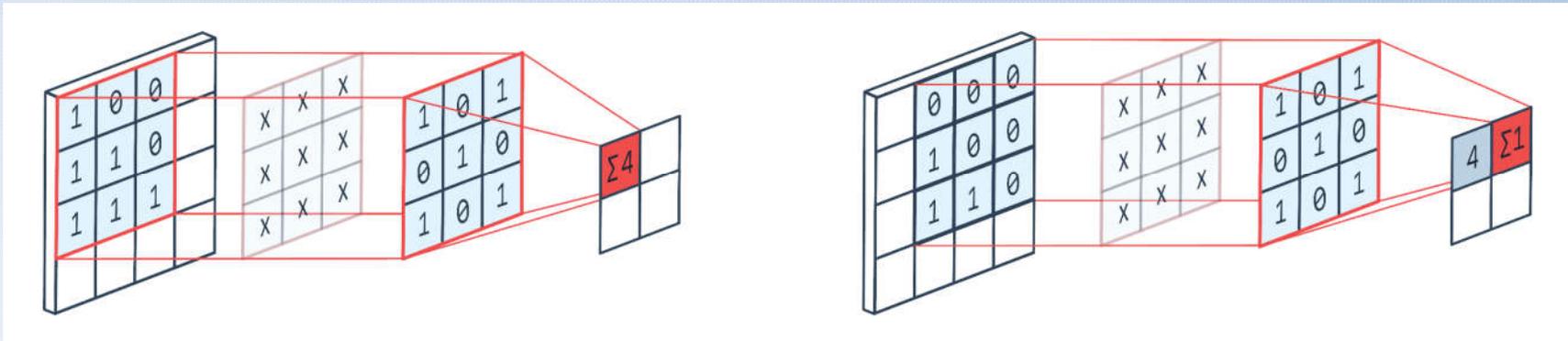
- Neurone artificiel : entrées pondérées puis sommées (fonction linéaire) → sortie
- Réseau de neurones : organisation en couches pour fonctions non linéaires
 - Davantage de couches = fonction plus complexe → Deep Neural Networks (DNN)
- Apprentissage par minimisation de fonction d'erreur vs vérité terrain
 - Descente de gradient + rétro propagation → nécessite grande base de données annotée



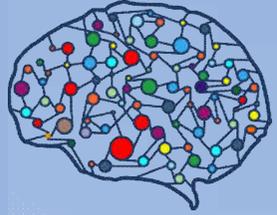
Deep Learning appliqué aux images



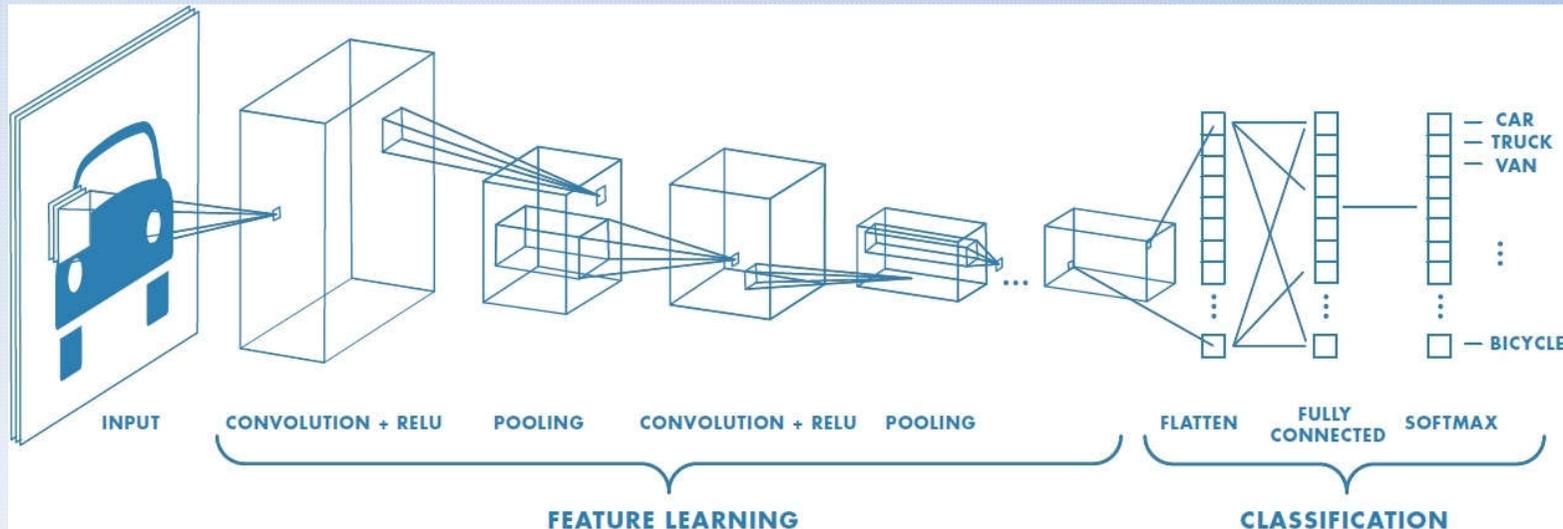
- Image = #entrées > 1M → DNN inutilisable en pratique
 - Solution : réduire la taille des données d'entrée
- Couche à convolution + pooling
 - Beaucoup moins de paramètres (#valeurs du filtre <<< #pixels)
 - Réduction progressive de l'image → génération de caractéristiques de plus en plus fines
 - Prise en compte de la structure locale des images



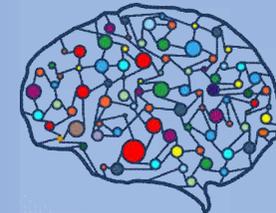
Convolutional Neural Networks (CNN)



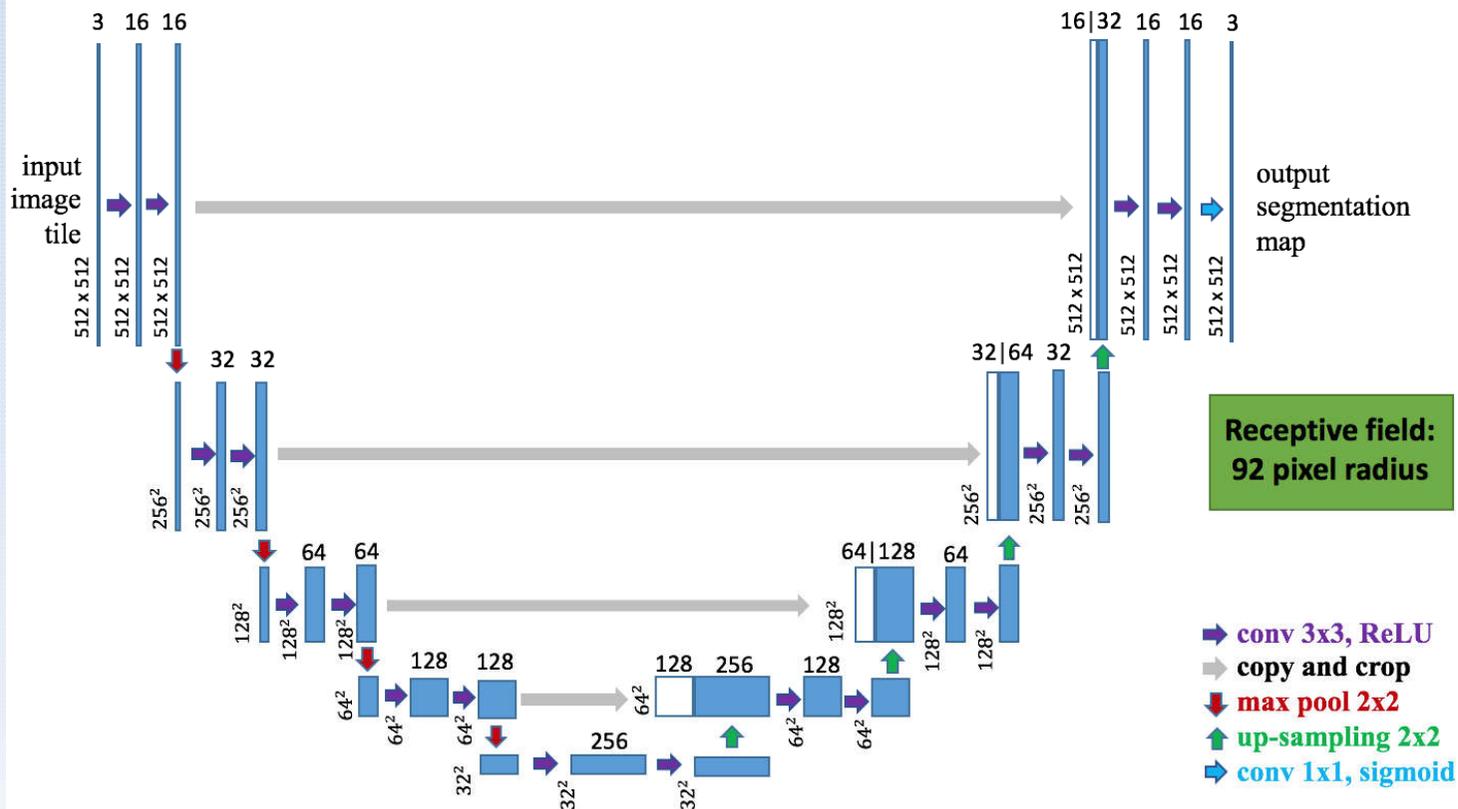
- 1^{ère} partie convolutive → extraction caractéristiques (feature map)
- 2^{ème} partie fully-connected (FC) → classification (NN standard)
- Réseaux récents presque entièrement convolutifs



CNN pour la segmentation d'images ?



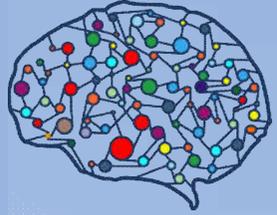
Ronneberger et al., 2015. « U-Net: Convolutional Networks for Biomedical Image Segmentation »



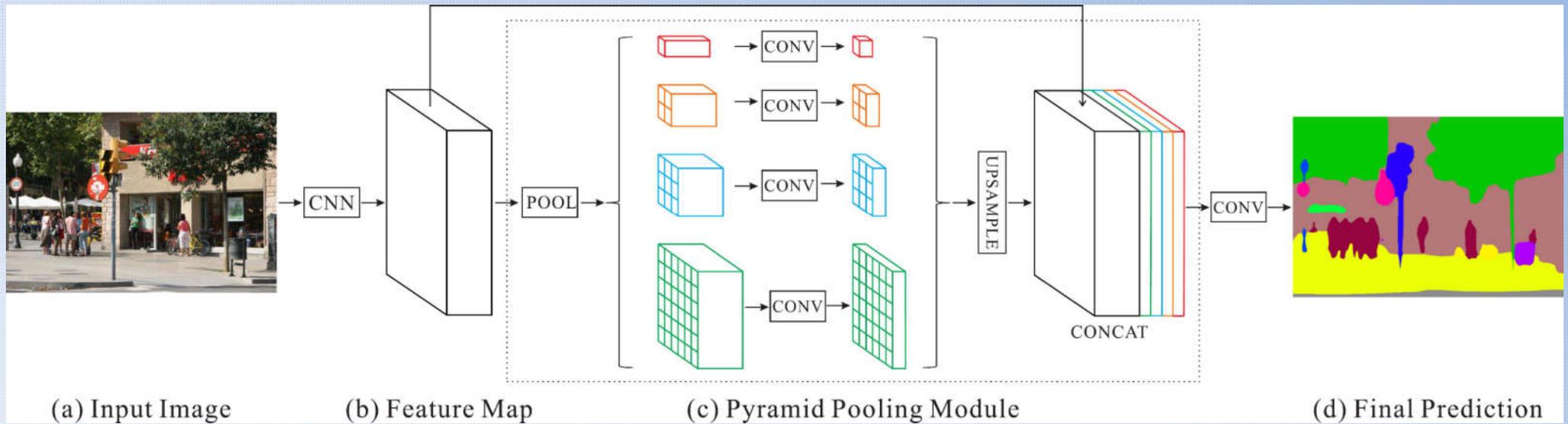
- Type auto-encodeur
 - Encodage : feature map
 - Décodage : localisation
- Sortie : probabilité / pixel
- Copie des feature maps pour meilleure localisation
- Encodage = CNN standard
 - *Transfer learning*



PSPNet



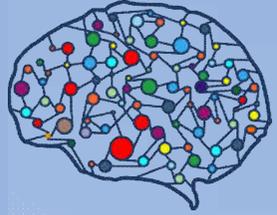
Zhao et al., 2017. « Pyramid Scene Parsing Network »



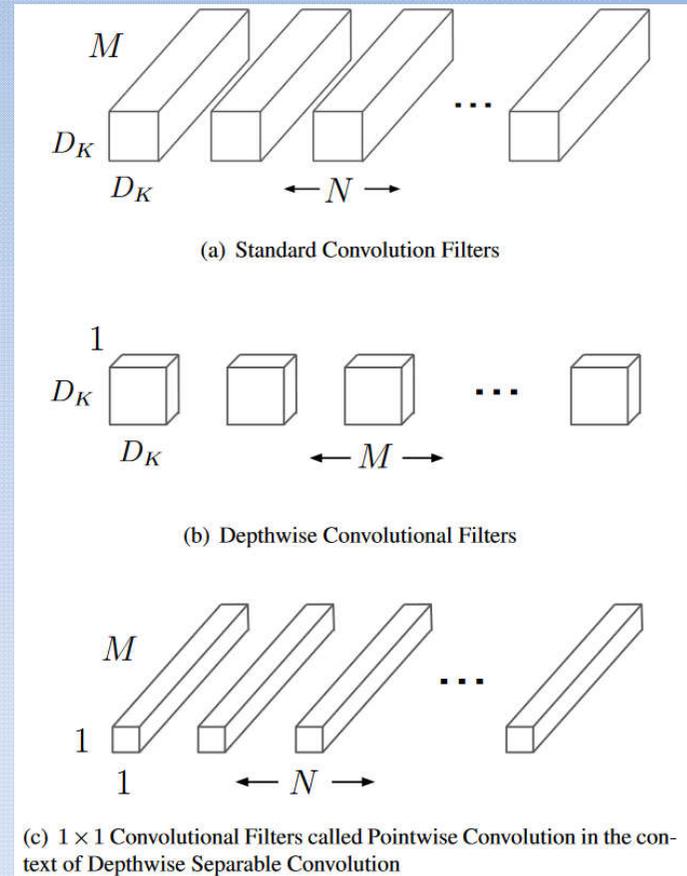
- Feature map = $1/8$ image
- Décodage remplacé par un pooling à différentes échelles → un seul upsampling
- CNN pré-entraîné (transfer learning)

MobileNet V2

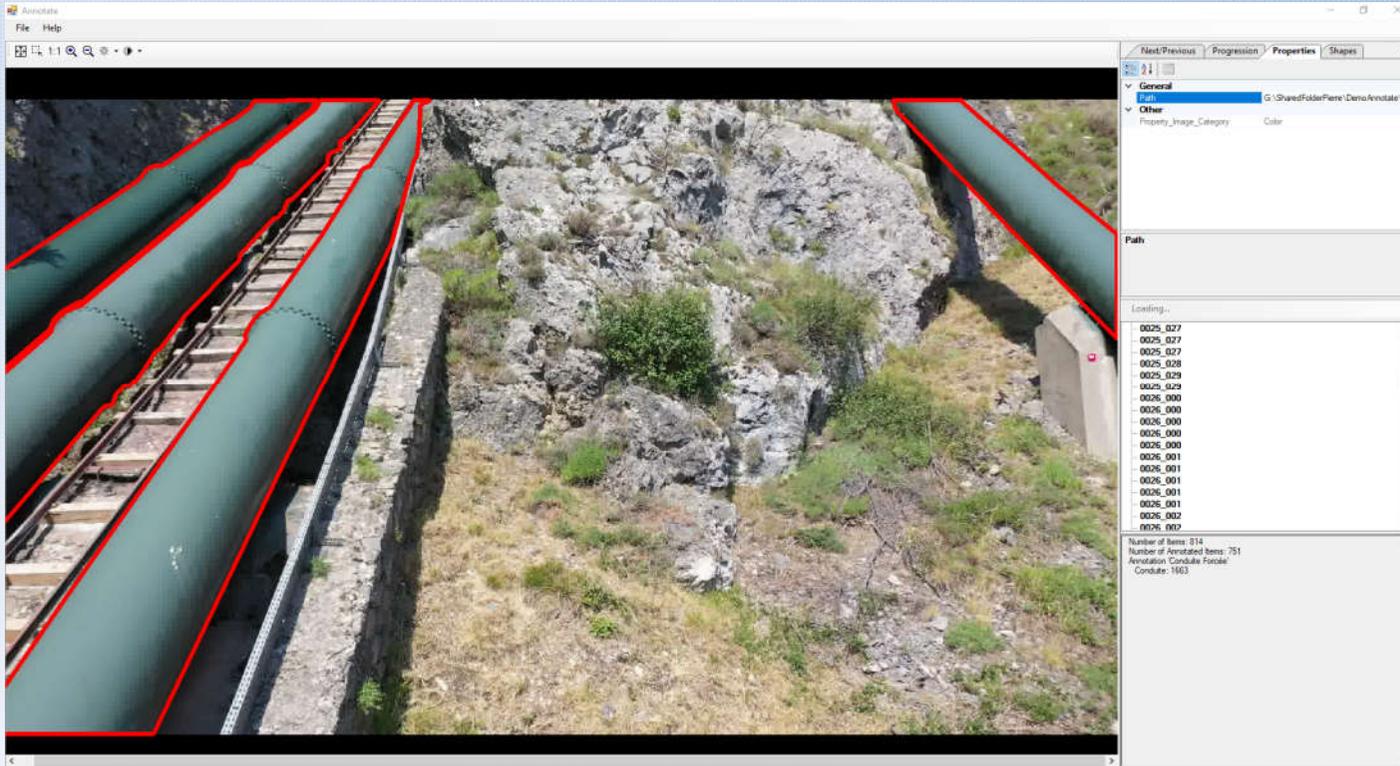
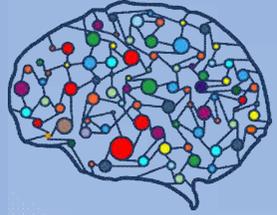
Sandler et al., 2019. « MobileNetV2: Inverted Residuals and Linear Bottlenecks »



- Approximer couches de convolution classique par depthwise (dw) + pointwise (pw) convolutions
 - Convolution (N filtres) : $D_F \times D_F \times M \rightarrow D_F \times D_F \times M \times N$ avec un kernel de taille $D_K \times D_K \times N$
 - Dw : un filtre / canal ($D_K \times D_K \times 1$)
 - Pw : combine les sorties de chaque dw ($1 \times 1 \times N$)
 - Gain en complexité de facteur 8-9 pour filtres 3x3
- MobileNet : convolution + 13 couches dw + pw
- MobileNet V2 :
 - pw remplacé par expansion + réduction des canaux avant et après dw \rightarrow agit comme décompression/compression



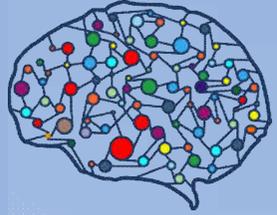
Annotation & Apprentissage



- 1922 images annotées (Saint Dalmas de Tende)
- Logiciel *Annotate* (ADCIS)
- PSPNet + MobileNet V2
 - Transfer learning depuis modèle MobileNet V2 appris sur base ImageNet
- Augmentation de données
 - Crop/flip aléatoire
 - Distorsion couleur/luminosité
 - Ajout de bruit, ...
- Images 320x480 pixels

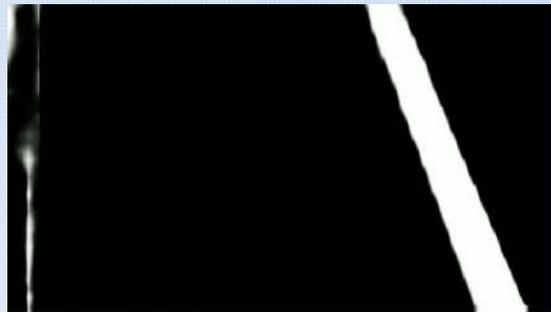
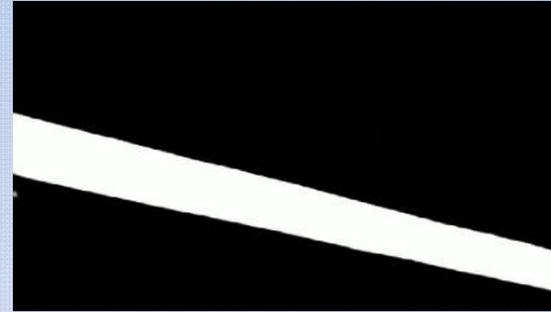
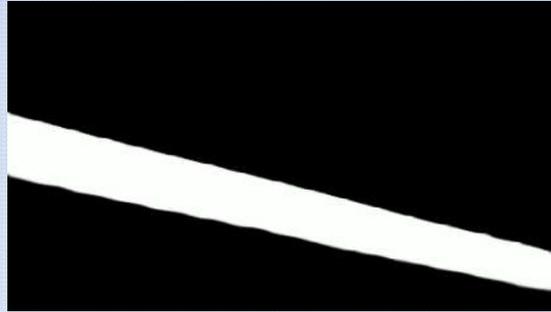


Résultats segmentation



PSPNet + MobileNet V2

Unet + ResNet V1 101*

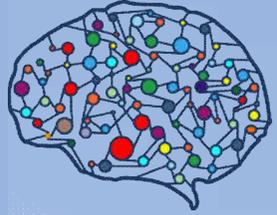


Sur Nvidia Jetson TX2 :

- PSPNet : 15 FPS
- Unet : 3,4 FPS

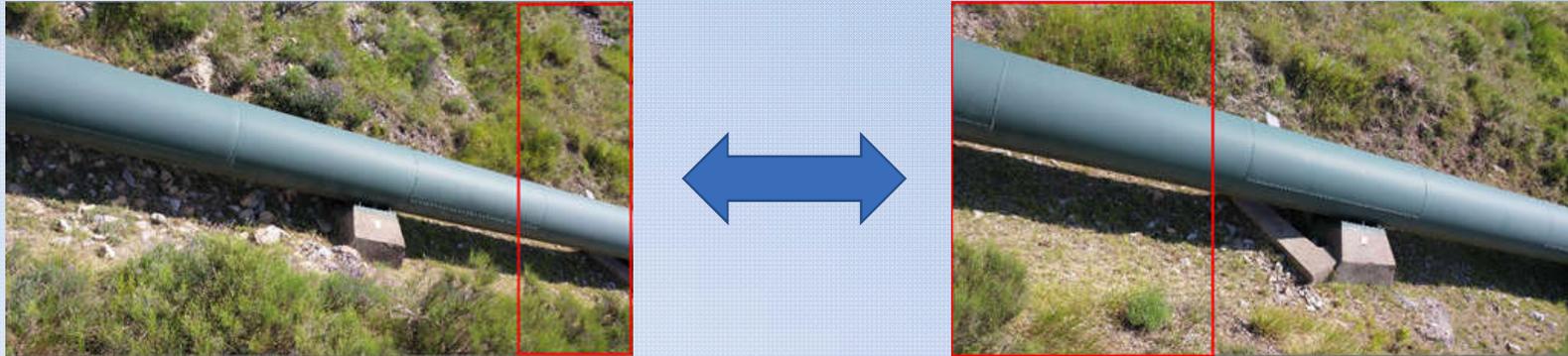
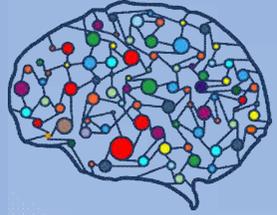
*He et al., 2015. « Deep Residual Learning for Image Recognition »

Asservissement de la nacelle du drone



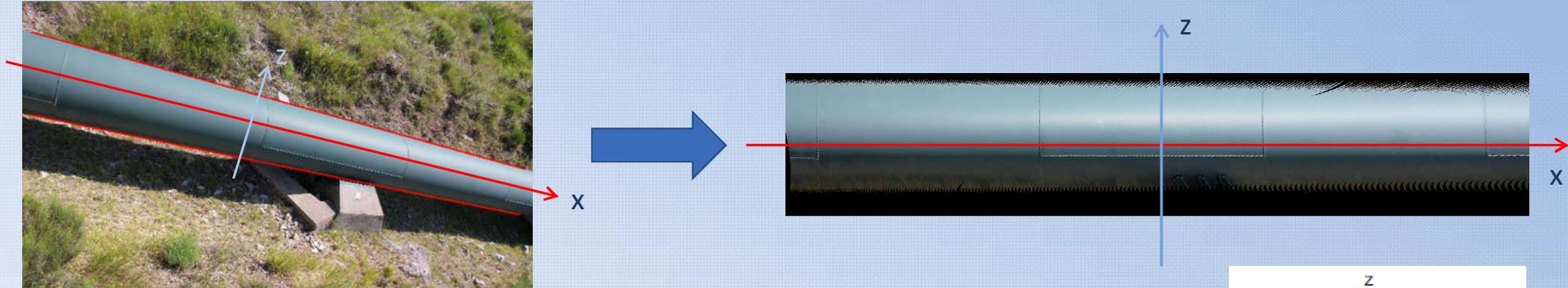
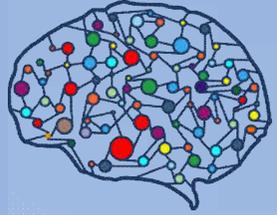
- Seuil haut (0,95-0,99) pour extraction d'un masque binaire
- Ouverture morphologique pour éliminer les petits objets
- Extraction de la plus grande composante connexe et de son barycentre
- Envoi d'une instruction au drone (haut/bas) en fonction de sa position

Déroulé et reconstruction de conduite

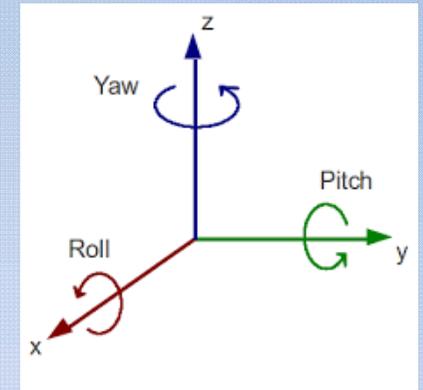


- Objectif : recaler la même section sur 2 images successives pour reconstruire l'intégralité de la conduite
 - Non directement superposables dû à la perspective → transformation nécessaire
- Pas de contrainte de temps → modèle Unet + ResNet V1 101
 - Image pleine résolution (3840x2160 pixels) → plus de précision

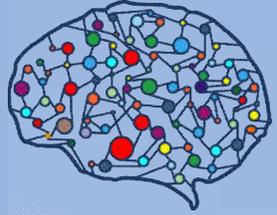
Déroulé et reconstruction de conduite (2)



- Extraction des contours haut et bas (approximation par segments)
- Conduites rectilignes + diamètre constant + pas de distorsion caméra
- La position et l'orientation des 2 droites permet de déduire :
 - Translation le long de l'axe vertical (z) et le long de l'axe de la caméra (y)
 - Rotation autour des axes y (pitch) et z (yaw)
 - x et roll sont invariants (axe du cylindre)

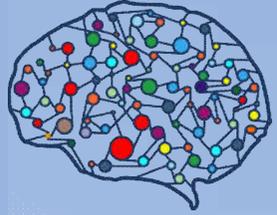


Déroulé et reconstruction de conduite (3)



- Après projection, les parties communes de 2 images successives sont égales à une translation près (car diamètre constant)
- Recalage par cross-corrélation sur bande rouge
- Avant recalage, application d'un chapeau haut de forme
 - Mettre en évidence les structures visibles sur la conduite et s'abstraire du fond
 - Changement de l'orientation relative de la lumière incidente et de l'angle d'observation
- Méthode limitée à la reconstruction d'un cylindre

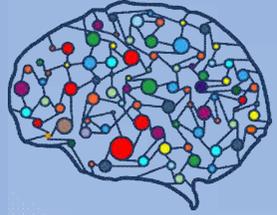
Conclusion



- Utilisation de réseaux de neurones convolutifs (CNN) pour résoudre le problème de segmentation de conduites forcées
 - Modèle limité par un manque de données
- Application à l'embarqué pour asservissement de la nacelle du drone
 - 15 FPS largement suffisant pour un contrôle en temps réel (simulation)
 - Pas encore de test en situation réelle
- Application à la reconstruction des conduites
 - Précision suffisante pour un recalage
 - Reconstruction limitée à un cylindre (pas de déformation – vannes, etc.)
 - Rendu perturbé par la variation d'intensité de lumière



Perspectives



- Augmentation de la base de données
 - Davantage de sites et campagnes acquisition + annotations
 - Augmentation de la robustesse des modèles
- Tests d'asservissement en situation réelle
- Amélioration de la reconstruction
 - Techniques + génériques (stéréo ?) pour détails hors conduite
 - Amélioration du rendu (variation lumière)
- Application de la reconstruction à la détection d'anomalies
 - Pas de données exploitables pour l'instant

