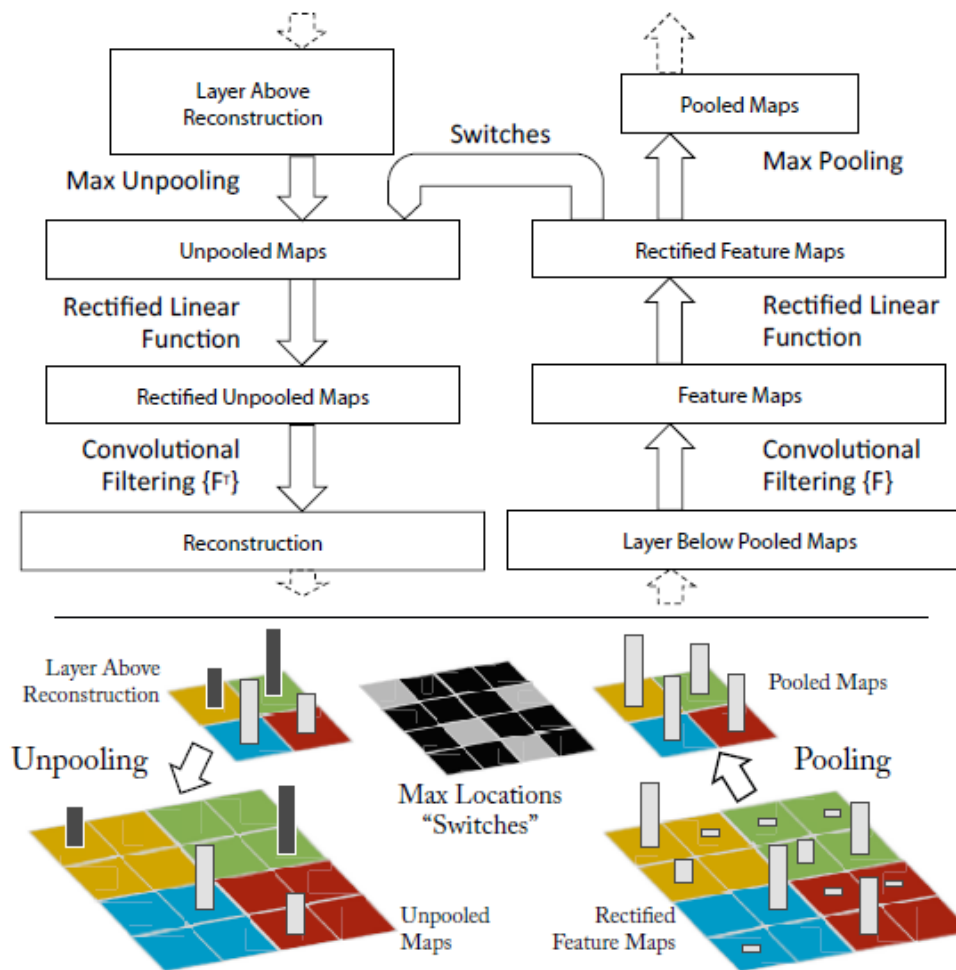


Decision explanation		Local	Global	방법	
※ Backpropagation-based methods	Gradient	√ Deconv net		Model-agnostic	
		√ Guided backpropagation			
	Input x Gradient	√ LRP		Model-specific	모든 pixel에 대해 pixel-wise relevance 구함
		√ DeepLIFT		Model-agnostic	모든 뉴런에 대해 reference 기준으로 attribution 구함
		√ Integrated Gradients		Model-agnostic	두 가지 공리를 만족하는 attribution method (Sensitivity, Implementation Invariance), path integral
	Activation map	√ CAM		Model-agnostic	Global average pooling
		√ Grad-CAM			weight로 gradient 사용
		√ Grad-CAM++			feature map의 weight + pixel의 weight
		√ SmoothGrad			removing noise by adding noise
※ Perturbation-based methods		√ LIME	SP LIME	Model-agnostic	local interpretable model-agnostic explanation
		√ SHAP	모든 local instances의 조합에 대한 평균		additive feature attribution, Shapley values, 평균과 차이
※ Hybrid methods		√ DeepLIFT SHAP, LIME SHAP			methods + SHAP
		√ Smooth Grad-CAM++			SmoothGrad + Grad-CAM++

- Visualizing and Understanding Convolutional Networks (DeconvNet)

- ✓ 이미지로부터 특징이 어떻게 추출되고 학습되어 가는지 시각화를 통한 접근 방법
- ✓ Deconvolution이라는 개념을 사용하며
- ✓ Conv-ReLU-pooling(max) 과정의 반대로 수행
- ✓ Switch variables를 정의하여 max-pooling전에 이전 feature map에서 가장 큰 값들의 위치를 저장



- Guided backpropagation

