Image processing applications in the field of autonomous driving

Created by: Róbert Varga
robertvarga87@gmail.com

Brief introductory

- **Driver Assistance System**
  - Interaction of a system with the human driver
  - various forms of task division between the human and the automatic system

  Operation modes:

  Informing functions:

  - only affect the vehicle guidance “indirectly,” namely via the driver.
  - regularly perform the task of environment perception in vehicle guidance as is simultaneously performed by the driver.
  - information is made available to the driver via the so-called human-machine interface.

Brief introductory

- **Driver Assistance System**
  - Interaction of a system with the human driver
  - various forms of task division between the human and the automatic system

  Operation modes:

  Continuously automated functions:
  - characterized as directly intervening
  - influence the vehicle control over long periods or parts of the journey.
  - information via a human-machine interface but also immediately perceptible for the driver, partly via a feedback from the controls (e.g., on the steering wheel)

Brief introductory

- **Driver Assistance System**
  - Interaction of a system with the human driver
  - various forms of task division between the human and the automatic system

Operation modes:

Intervening emergency functions:

- In sudden emergencies, to be described as near-collision situations, the human can only react with a delay
- more specifically referred to as emergency functions, are superior to drivers, and their action is temporarily not subject to human control.

Brief introductory

- **Driver Assistance System**

Operation modes:

Operation type A: Informing and warning functions	Operation type B: Continuously automating functions	Operation type C: Intervening emergency functions (near-accident situations)
Take only indirect influence on vehicle control via the driver	Take immediate control over the vehicle. Division of tasks between the human driver and the function (usually convenience functions – control always remains overrideable)	Take immediate control over the vehicle in near-accident situations that de facto cannot be controlled/ handled by the driver (usually safety functions)
Examples: • Traffic sign recognition (display of current speed limit) • Lane departure warning (e.g. Vibration on the steering)	Examples: • Adaptive Cruise Control (ACC) • Lane keep assist (via steering interventions)	Examples: • Automatic emergency braking (system triggered) • Emergency steer assist (under development) • Emergency stop assist (driver suddenly not capable of acting)

Brief introductory

- **Driver Assistance System**

Levels of Automated Driver Assistance System

Level	Name	Narrative definition	Execution of steering and acceleration/ deceleration	Monitoring of driving environment	Fallback performance of *dynamic driving task*	System capability (*driving modes*)	BASt level	NHTSA level
Human driver monitors the driving environment								
0	No Automation	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a	Driver only	0
1	Driver Assistance	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes	Assisted	1
2	Partial Automation	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes	Partially automated	2
Automated driving system ("system") monitors the driving environment								
3	Conditional Automation	the *driving mode*-specific performance by an *automated driving system* of all aspects of the *dynamic driving task* with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes	Highly automated	3
4	High Automation	the *driving mode*-specific performance by an *automated driving system* of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes	Fully automated	3/4
5	Full Automation	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**	-	

Short URL: http://cyberlaw.stanford.edu/loda

Brief introductory

- **Driver Assistance System**

  Vehicle dynamincs sensors:

  - Wheel speed sensor
  - Steering angle sensor
  - Yaw rate and acceleration sensor
  - Brake pressure sensor
  - Etc.

Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Ultrasound sensors

  – The basic function of an ultrasonic parking assistance system

  Radar

  – electromagnetic wave will undergo a frequency shift if the observer and transmitter move relative to each other.

  – Thus a RADAR beam to an arbitrary distance **r** and back again to the receiver travels a real figure z of $z = 2r/\lambda$ wave lengths in total

  – Therefore a phase lag of $\phi = -2\pi z$ arises

  – If r now changes with $\dot{r}$, then the phase also experiences a change of $\phi\_ = 2\pi z\_ = 4\pi \dot{r} / \lambda$

Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Radar

- The Doppler effect is expressed as the frequency change $f_{Doppler}$ which is proportional to the relative velocity and to the reciprocal value of the wave length $\lambda = f_0/c$ (speed of light c), where the frequency shift is positive when approaching ($\dot{r} < 0$) and negative when departing.

$$f_{\text{Doppler}} = -2\dot{r}/\lambda = -2\dot{r}f_0/c$$

Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Lidar

  Light Detection And Ranging (LIDAR)

  – optical measurement principle to localize and measure the distance of objects in space. Basically it is similar to a RADAR-system, but instead of using microwaves LIDAR uses ultraviolet, infrared or beams within the visible light spectrum.

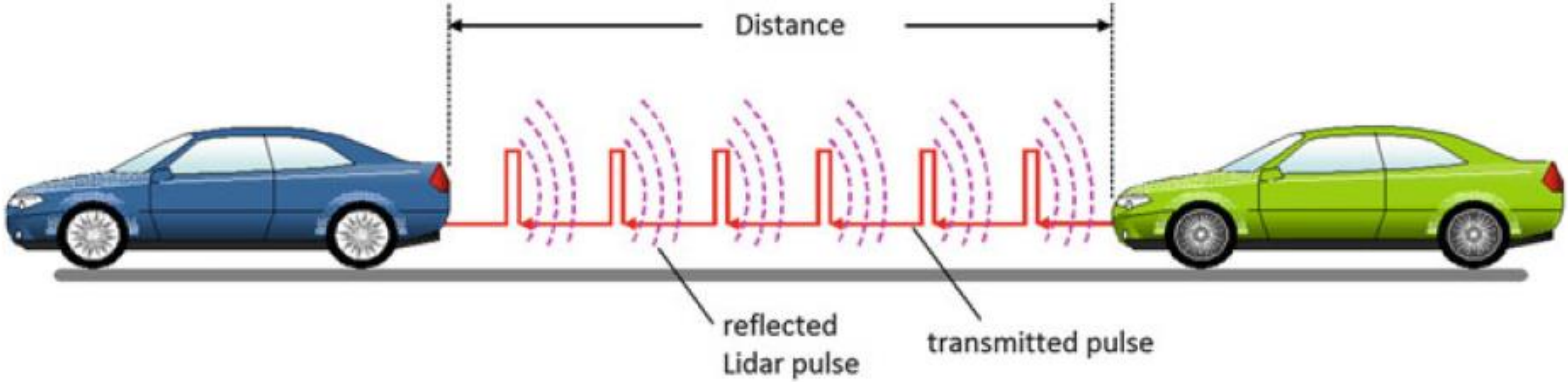
  – can be used for a limited visual detection of objects by analyzing the light intensity, visibility measurement by analyzing the shape of the reflected LIDAR pulse, day/night detection as background illumination

Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Lidar



Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Camera:

  Applications:

  - Interior Monitoring
  - Environmental Sensing
- full recognition of all relevant road users, road scenery, and road signs in order to be able to react accordingly

Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Camera:

  - **Operates on the visible spectral range:**
  - high-beam assist
  - Traffic sign recognition
  - lane detection
  - Object detection
  - **or infrared spectral range:**
  - disadvantage of cameras operating in the visible spectral range is the insufficient sensitivity at very low light conditions

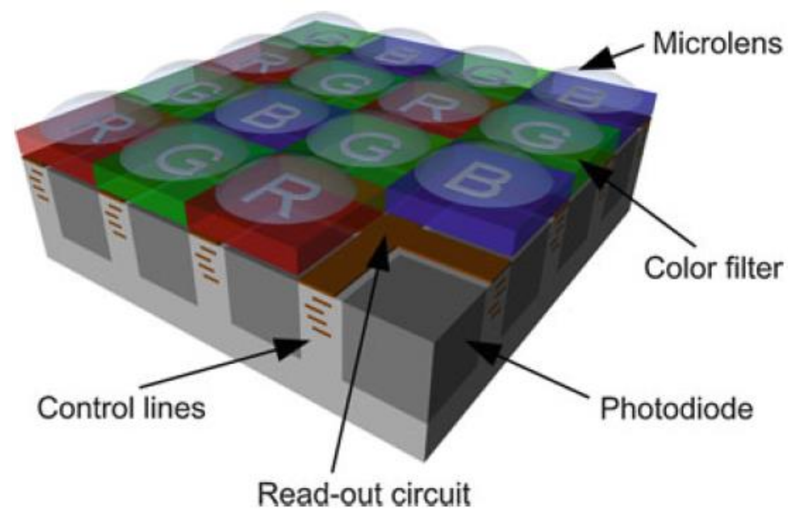
Brief introductory

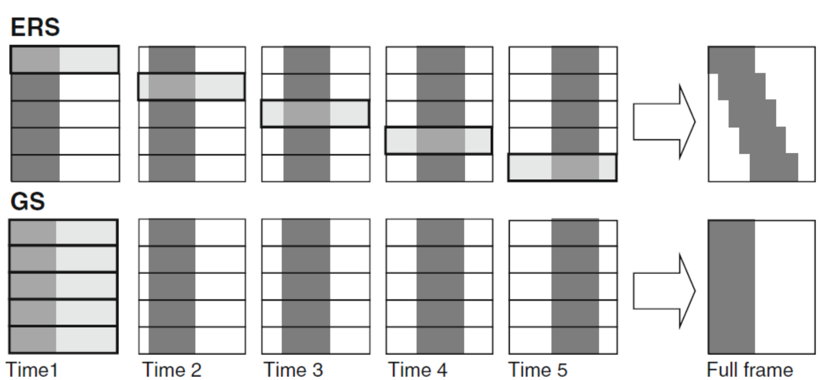
- **Driver Assistance System**

  Environment sensors:

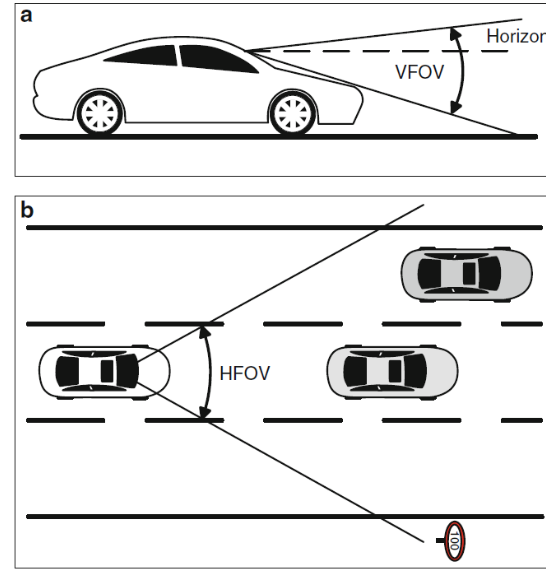
  Camera:

  - **Spatial resolution**
  - **Temporal resolution**
  - **Shutter**
  - **Dynamic range**
  - **Color reproduction**





Brief introductory



- **Driver Assistance System**

  Environment sensors:

  Camera:

  Field of view
- For a large detection range, e.g., for vehicle detection on highways, a high resolution is necessary.
- Pedestrian detection benefits from a large field of view. In general, a high sensitivity of the camera system is very important.

Brief introductory

- **Driver Assistance System**

  Environment sensors:

  Camera:

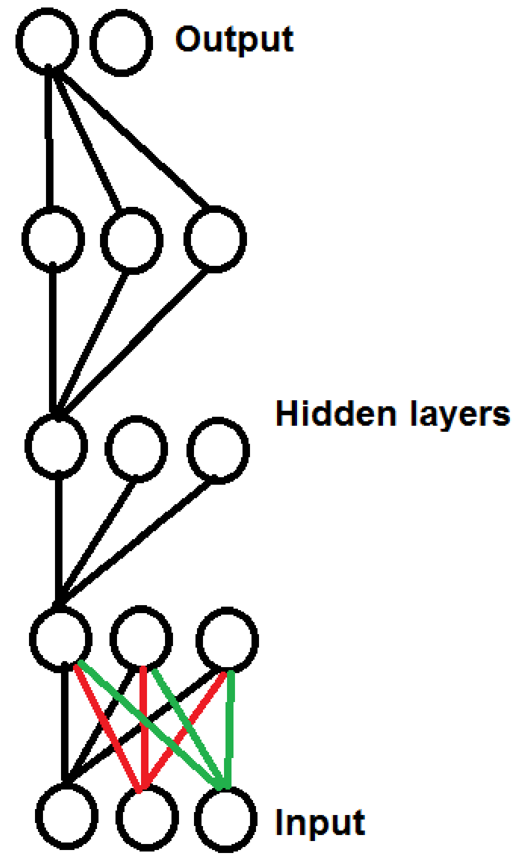
  Alignment

  – Surround view systems are equipped with four or more cameras around the vehicle.

  – Camera modules for such systems are usually equipped with the so-called fish-eye lenses, which allow a horizontal field of view of more than 180 deg

Brief introductory

- **Deep neural networks**
- **Bottom layer**
  - Referred to as independent variables in traditional modelling literature
- **Hidden layer**
  - Each neuron in this layer attains an activation which is computed by a weighted sum of its inputs plus a bias, and then applying a nonlinear function
  - Each hidden neuron in this layer will have a different set of input weights
- **Output layer**
  - Topmost layer is the output

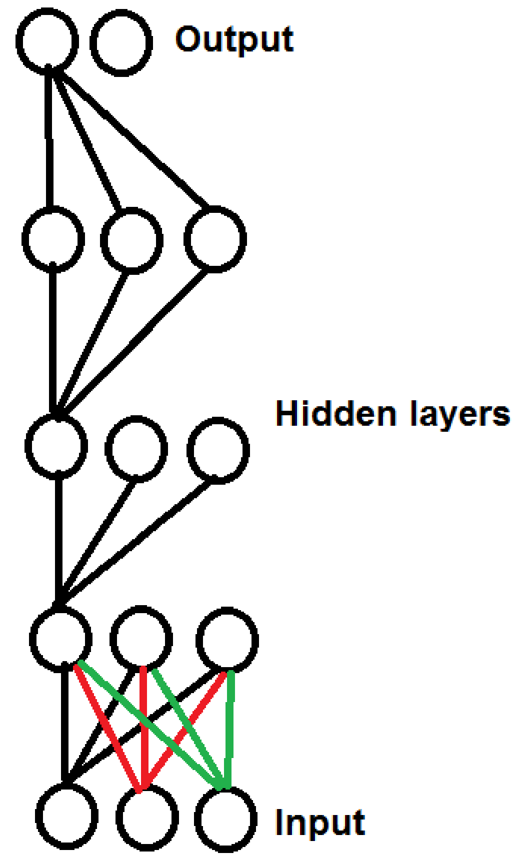


Brief introductory

- **Deep neural networks**
- **To be more specific**

$\mathbf{X} = \{x_1, x_2, \ldots, x_K\}$ is the vector of prior-layer activations

$\mathbf{W} = \{w_1, w_2, \ldots, w_K\}$ is the vector of associated weights and **b** is the bias term



$$a = f(b + \sum_{k=1}^{K} w_k x_k)$$

$$a = f(b + WX)$$

The nonlinear activation function is applied element-wise to the vector

Brief introductory

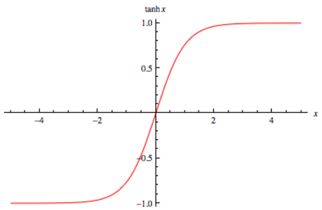
- **Deep neural networks**
- **To be more specific**
  - Bias term can be absorbed into the weight matrix **W** by appending it as one more column to the right side
  - We then augment the vector **X** by:

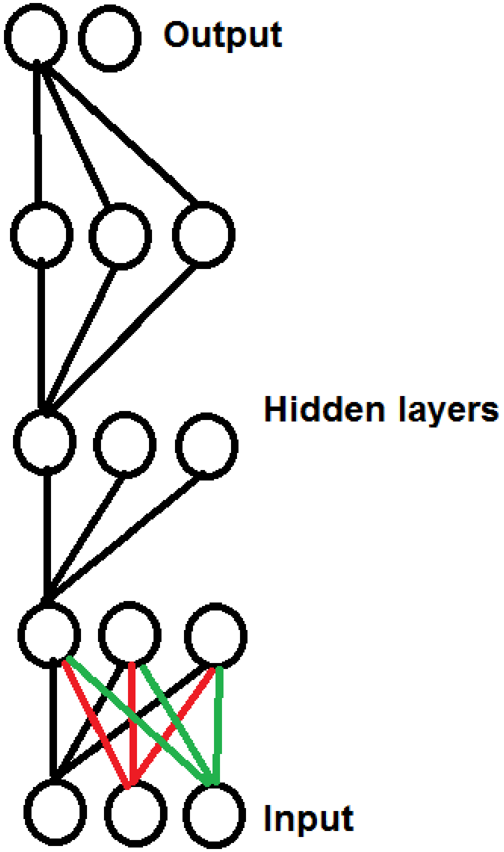
    $\mathbf{X} = \{x_1, x_2, \ldots, x_{K,} 1\}$
  - This simplifies to the activation function on a simple matrix/vector multiplication:

$$a = f(WX)$$

  - Activation function eg. Hyperbolic tangent:

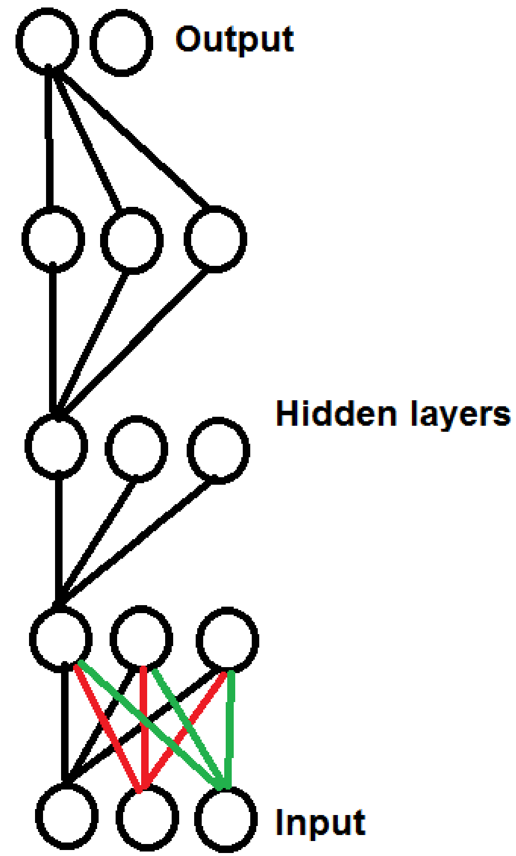
$$\tanh(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}}$$





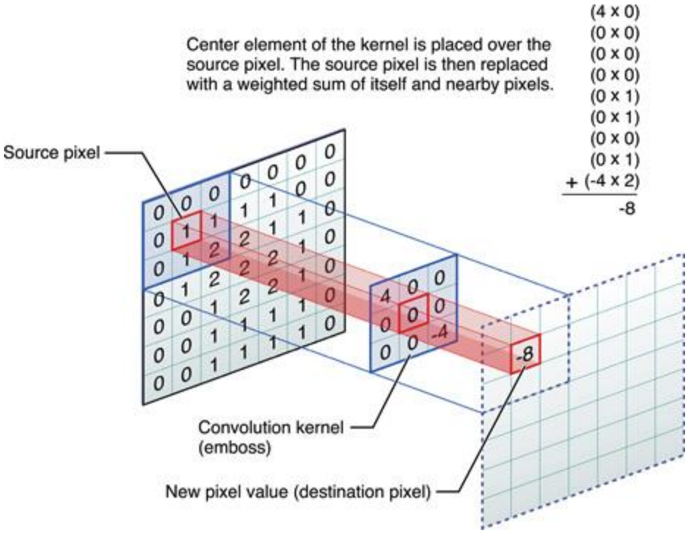
Brief introductory

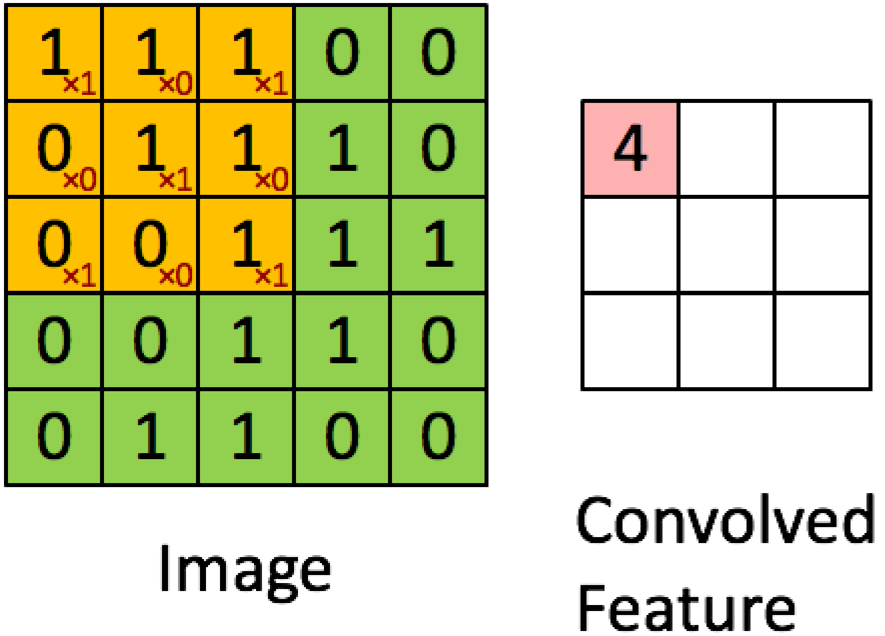
- **Deep neural networks**
- Shallow VS Deep network
- Theorems were proved showing that is very broad classes of problems, one or two hidden layers were sufficient to solve the problem
- Theoretical results began appearing which showed that for many important classes of problems, a network composed of numerous narrow layers would be more powerful than a wider, shallow network having the same number of neurons



Brief introductory

- **Deep neural networks**
- Convolutional layers



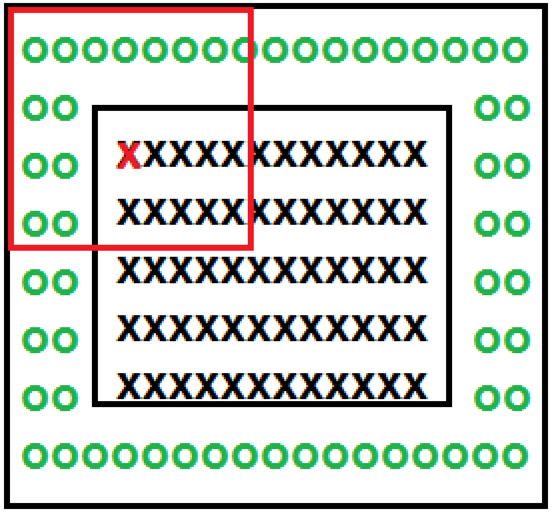


Brief introductory

- **Deep neural networks**
- Convolutional layers
- Padding: filter extending past the edge of the prior layer into undefined nothingness
- Denote the vertical and horizontal half-widths of the filter: W_v, W_h
- The number of weights involved in computing the activation of a single neuron:

$$Priorweights = N_s(2W_h + 1)(2W_h + 1)+1$$

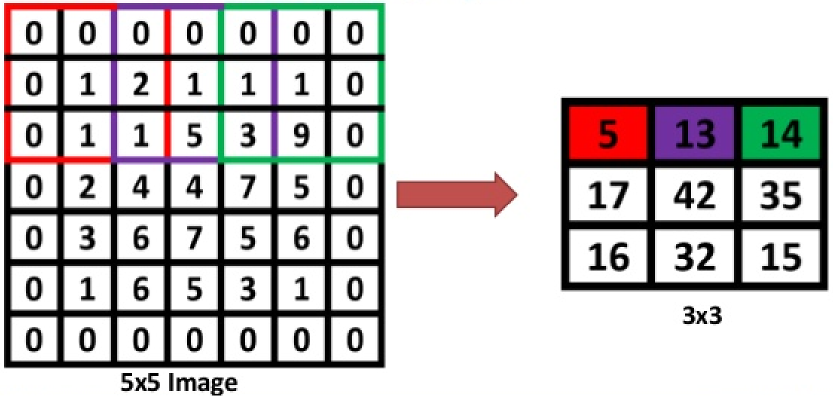
- N_s number of slices in the prior layer
- +1 is the bias term



Brief introductory

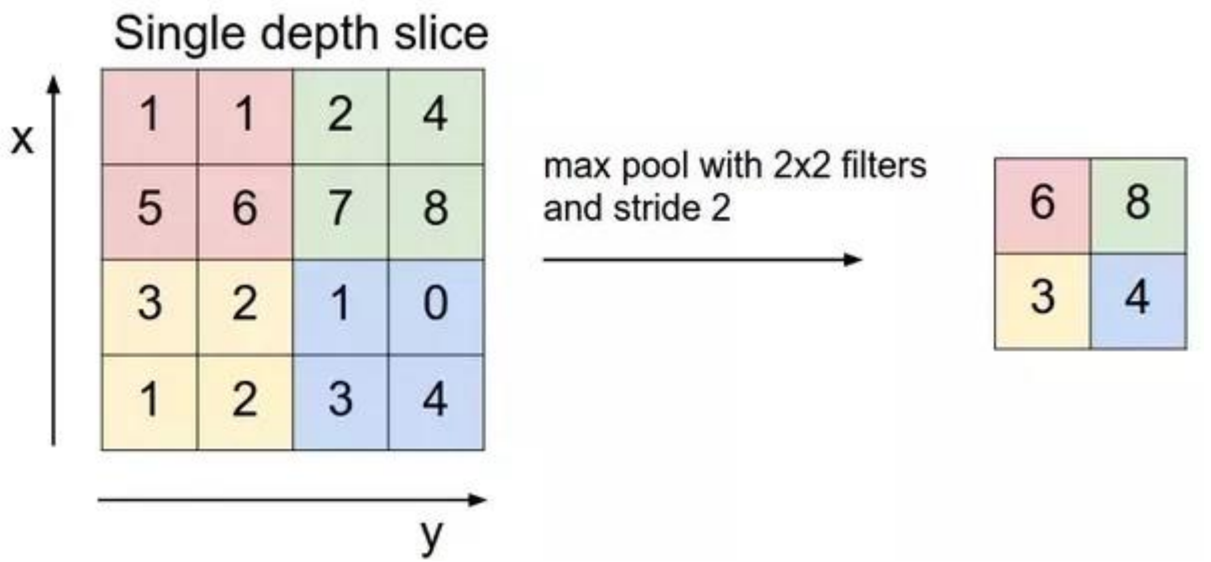
- **Deep neural networks**
- Convolutional layers
- Striding:
- Principle of NN design is that the size of hidden layers decreases from input towards to the output
- Effective information compression is obtained by decreasing the size of the visual field in successive layers

❑ Conv 3x3 with **stride=2,padding=1**



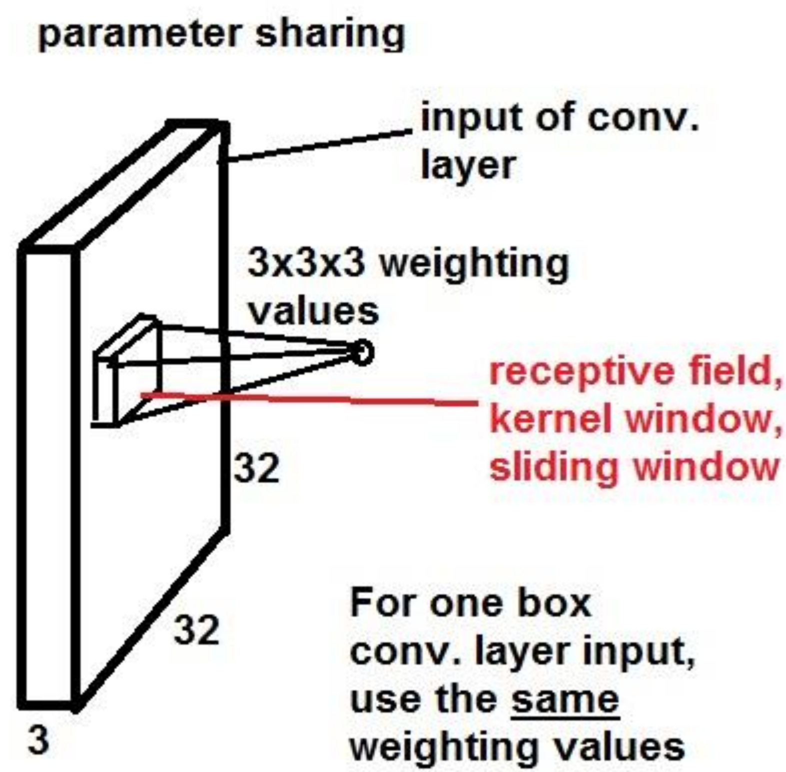
Brief introductory

- **Deep neural networks**
- Convolutional layers
- Pooling:
- Another tendency using pooling to reduce the visual field
- Different pooling strategies are existing e.g: Max pooling



Brief introductory

- **Deep neural networks**
- Convolutional layers
- Parameter sharing
  - Image feature reuse
  - Reducing complexity of NN



Brief introductory

- **Deep neural networks**
- Convolutional layers
- Formula for the number of rows/columns in the current layer:
  - **W** width/height of the prior layer
  - **F** width/height of the filter (two times half-width plus one)
  - **P** padding appended to each edges
  - **S** stride
  - **C** width/height of the current layer

$$C = \frac{(W - F + 2P)}{S} + 1$$

If the division is not exact, the alignment of the current layer with the prior layer will not be symmetric

Brief introductory

- **Deep neural networks**
- Output layer
- Making the predicted outputs resemble probabilities
- Make the model far more robust against outliers in the training/testing data
- For the output neuron we drop the nonlinear function and speak only of the weighted average (+bias)
- This quantity is called logit of the neuron

$$logit_k = b_k + \sum_i w_{ki} x_{ki}$$

Softmax output values:

$$p(y = k) = \frac{e^{logit_k}}{\sum_{i=1}^{K} e^{logit_i}}$$