

Machine Learning for Human Vision and Language

Lecture 2: Deep learning in biological neurons and networks

Ben Harvey

1

In today's class, we will move away from artificial networks for machine learning to look at the structures that inspired them: networks of biological neurons learning about the world to guide our behaviour

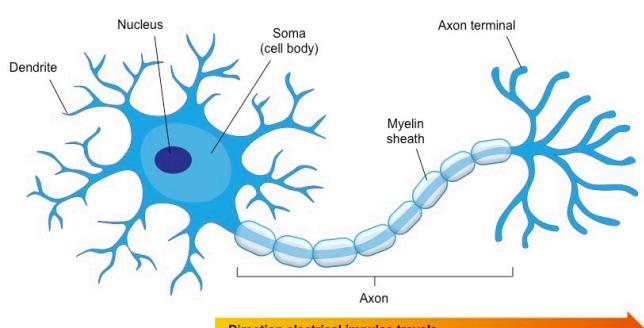
Why study biological networks?

- Inspiration for many AI networks, including DCNNs
- AI networks simplify processes involved for efficiency
 - AI deep networks not yet as advanced as human brain
 - Where the goal is to imitate human behaviour, following the neural mechanisms more closely may help
- Major link between AI and other sciences
 - Links AI to biological sciences, not just math/CS
 - Potential to link social sciences to biological sciences
- Leading model of neural computations

2

- Social sciences are only loosely connected to natural sciences
- DCNN are among the leading methods to model how the brain interprets reality

A biological neuron



3

- Tree of dendrites is very similar to a convolutional filter

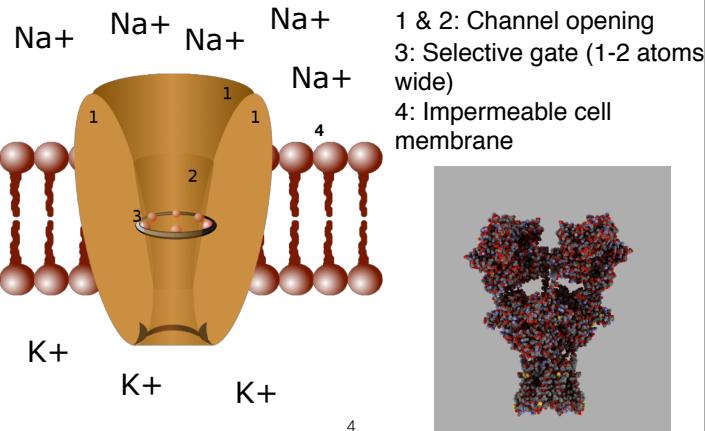
A neuron is a specialised type of animal cell. Most aspects work like other cells and are not important for performing computations. But the cell membrane in neurons is highly unusual. It is specialised for performing simple computations using electrical activity.

First, the dendrites integrating electrical signals coming in from other neurons than it is connected to.

These connections, or synapses, between dendrites and other neurons vary in strength, and change strength depending on past activity, so synaptic strengths are the biological equivalent of artificial neural network weights. A tree of dendrites working together to sample information from other neurons is therefore the biological equivalent of a convolutional filter.

If these incoming electrical signals are strong enough to pass a THRESHOLD, the neuron will then send an electrical impulse down the axon. On reaching the axon terminals, this is passed on to other cells, in the next LAYER of neurons.

The ion channel



The basic computational component of a biological neuron is a protein called an ion channel.

-This can allow charged atoms, or ions, of sodium and potassium (natrium and kalium) to pass into and out of a neuron, but they are normally closed, blocking ion passage.

-Ion channels are imbedded in a cell membrane, which surrounds the neuron.

-This also does not allow ions to pass through.

-As ions can't pass through the membrane if the ion channel is closed, the cell membrane has an imbalance of ions on inside and out: far more sodium ions outside the cell, a few more potassium ions inside.

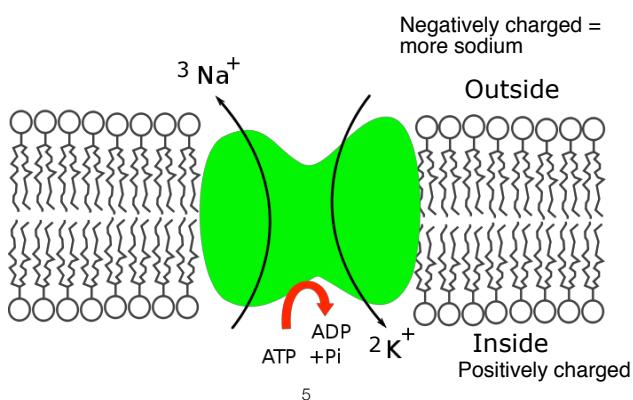
-This imbalance of electrically-charged ions causes a voltage difference across the cell membrane.

-But the ion channel protein can change shape, allowing a particular type of ion to pass through: there are different ion channels for sodium and potassium ions.

-When these ions cross the cell membrane, this changes the membrane voltage, or membrane potential. (or voltage)

-Note that the ion channel is a passive mechanism, it does not pump the ions against their concentration gradient, it only opens to allow ions to diffuse down the concentration gradient or closes to prevent this.

The sodium/potassium pump: establishing the resting potential



Because the ion channel is passive, it only allows ions to move towards lower concentrations or to overcome voltage differences

-This is very fast, much faster than pumping ions around to change voltages

But the ion imbalance must be maintained by an active mechanism (requiring energy).

-This pump is another protein imbedded in the cell membrane, the sodium-potassium pump, or Na^+/K^+ ATPase. ATP is the source of the energy used here, and is broken down by the process

-Because of more sodium being transported out, the inside of the cell is negatively charged compared to the outside.

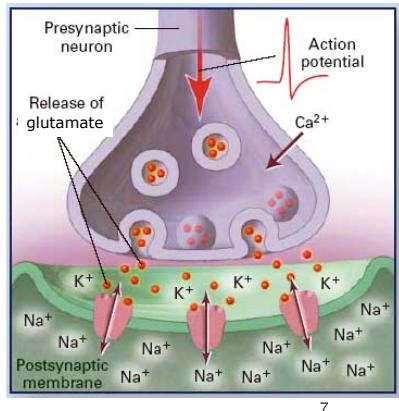
-In the cell membrane's rest state, there is imbalance of ions & charges across the cell membrane, making a voltage of -70 mV.

-This is known as a resting potential (this is the max voltage)

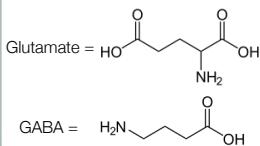
So...

- The cell membranes of neurons are specialised for electrical computations.
- Ions are atoms with electrical charges
- Cell membranes do not allow ions through
- Sodium/potassium pumps establish an imbalance of ions across the membrane
 - Much more sodium outside
 - A little more potassium inside
- Opening ion channels allows these ions to pass through, in the opposite direction, towards balance

Opening the ion channel



7



important similarity

So opening sodium ion channels causes sodium ions to enter the cell through the cell membrane, changing the membrane voltage or membrane potential.

-This may lead the cell to activate (or 'fire'), and the amount of this activation is what is simulated by the activation values in the feature maps of an artificial convolutional neural network.

-For our purposes, there are two important ways an ion channel can open.

-First, it can open because a neurotransmitter binds to the ion channel. In that case, the ion channel is a 'receptor' for the neurotransmitter.

-When the neurotransmitter binds, the ion channel protein changes shape, opening the ion channel for ions to cross the membrane.

-Another neuron firing causes the neurotransmitter release into a synapse, the gap between two linked neurons.

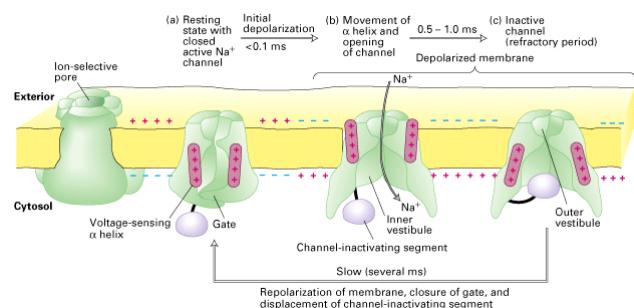
-So the neurotransmitter is a signal released by activation of one layer of a neural network (the presynaptic neuron) and causes activation of the next layer of the neural network (the postsynaptic neuron)

-Here we will look at excitation of the postsynaptic neuron by binding of glutamate, and inhibition by binding of gamma-amino butyric acid (GABA).

-Using both excitation and inhibition allows connections with both positive and negative weights, as convolutional filters have.

-Other neurotransmission works by similar mechanisms, but glutamate and GABA are by far the most common.

Opening the ion channel



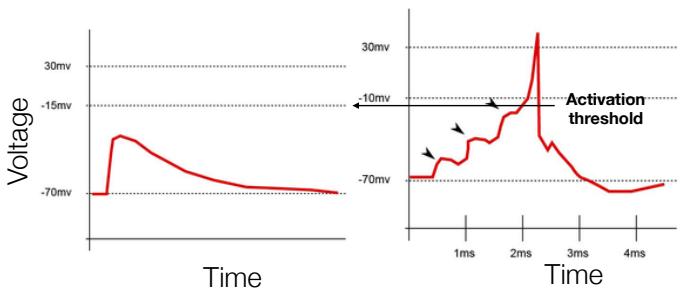
8

The second way an ion channel can open is due to a change in membrane voltage. This is particularly important because opening the ion channel also changes the membrane voltage.

-So changing membrane voltage can lead to further changes in membrane voltage.

-In this way, the voltage-gated ion channel acts very much like an electronic transistor, that is an electrical switch that opens and closes because of an electrical input. Miniaturised transistor circuits are the basis of all computer processors.

Excitatory post-synaptic potentials (EPSPs)



9

-When glutamate binds to a post-synaptic receptor, it causes the receptors ion channel to open, and sodium to enter the cell.

-This causes a change in the membrane voltage, called an excitatory post-synaptic potential.

-The size of the voltage difference across the cell membrane decreases, so we say the cell membrane is depolarised.

-But over time, this depolarisation decreases as sodium and potassium are pumped around to return the membrane to its resting potential.

-So this one neurotransmitter molecule binding causes no activity on the postsynaptic neuron, because it doesn't reach a threshold that activates the voltage-gated ion channels. Then, the neuron doesn't pass this event on, and it has no effect.

-However, if several neurotransmitters molecules bind over a short period, this can reach the threshold voltage for the voltage-gated sodium channels.

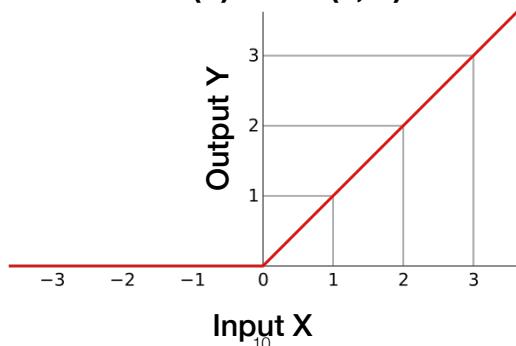
-This causes a much more extreme depolarization, a spike of voltage or action potential. This will be passed down the axon to the next layer of neurons.

-So this threshold for activation of voltage-gated sodium channels is the biological equivalent of the threshold/rectification operation we saw in artificial networks.

The threshold/rectification operation

an activation function using a rectified linear unit (ReLU)

$$Y = f(X) = \max(0, X)$$



-Until we reach a certain threshold voltage, there will be no response.

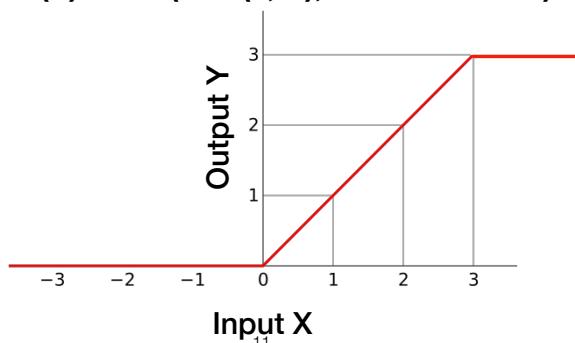
-This threshold is not zero in biological neurons, it's around -15mV because an imbalance of ions is maintained at rest to speed up neuron responses.

-Then the firing rate (output) of the neuron will increase as the strength of the inputs increases.

The threshold/rectification operation

an activation function using a rectified linear unit (ReLU)

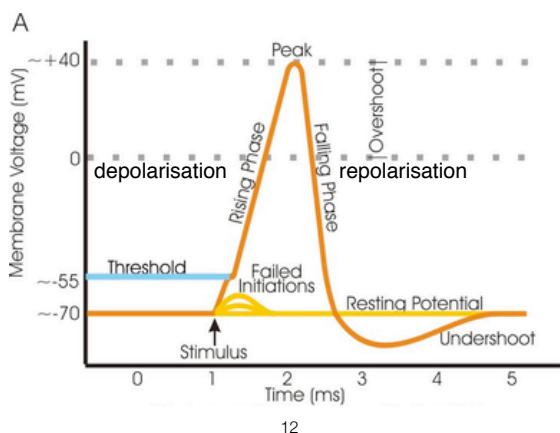
$$Y = f(X) = \min(\max(0, X), \text{MaxActivation})$$



And at some point this output will reach a maximum firing rate

At which the cell cannot fire anymore as the signal is passed to the next neuron

Action potentials



The reason for this maximum firing rate is that an action potential takes time, although not very much time.

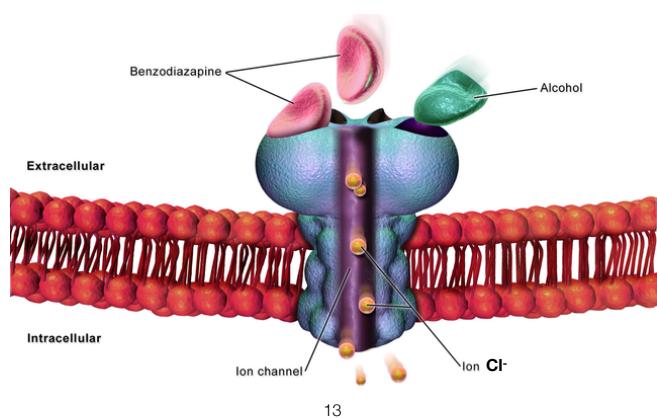
-At the peak, Na^+ concentration is almost equal across the cell membrane

-The high voltage then causes the Na^+ channels to close. This is a different closed state than before, as the channel protein is a different shape.

-Here, the Na^+ channel cannot open again until the cell is repolarised.

-The initial repolarisation (falling phase) relies on K^+ channels opening, which repolarises faster than pumping, but the Na^+ concentration is too high inside for further diffusion, and the K^+ concentration is too low inside. So opening the Na^+ channels again will do nothing. The Na^+/K^+ pump needs to return the cell to resting potential.

Inhibitory post-synaptic potentials (IPSPs)

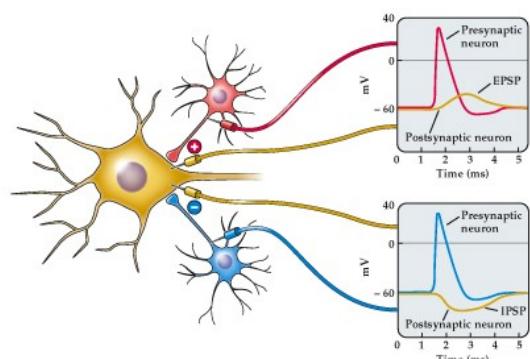


Activity in a presynaptic neuron can also inhibit the postsynaptic cell firing, by causing the membrane to polarise further, making the membrane voltage more negative, which means below -70.

-The GABA receptor is also an ion channel, but for negative chloride (Cl^-) ions. When GABA binds, Cl^- enters and the cell membrane potential becomes more negative than the resting potential.

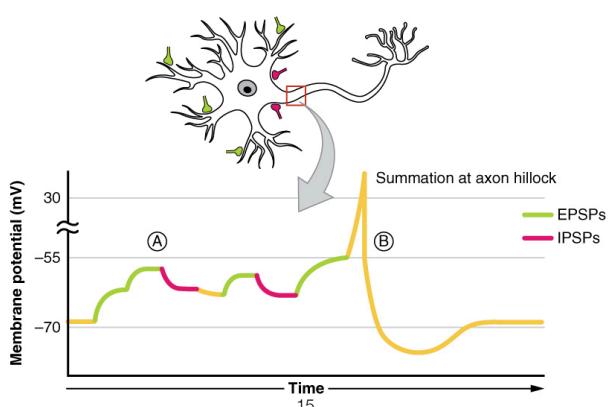
-This pushes the cell further away from activating.

Inhibitory post-synaptic potentials (IPSPs)



So we have both positive and negative inputs onto the same postsynaptic neuron

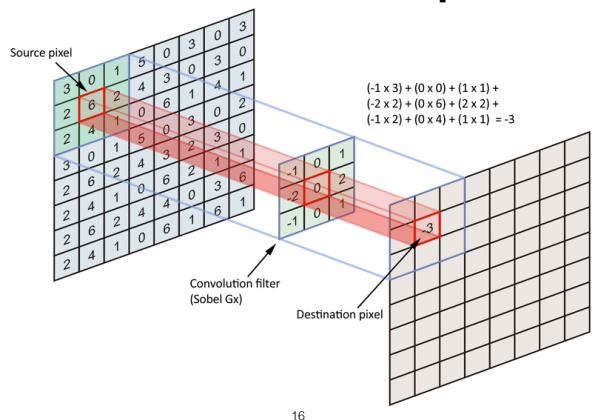
Inhibitory post-synaptic potentials (IPSPs)



These can be from a range of different places, and from a number of different neurons.

However, there is a limited spatial distribution of inputs because the dendritic tree has a limited size.
Which may sound very familiar...

The filter/convolve operation



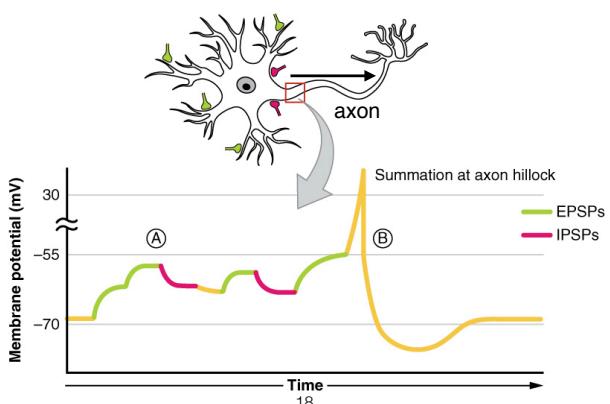
Because convolutional network filters imitate this structure: they have positive and negative filters, which are multiplied by the activity on a group of presynaptic units to give the activity of the postsynaptic unit

So...

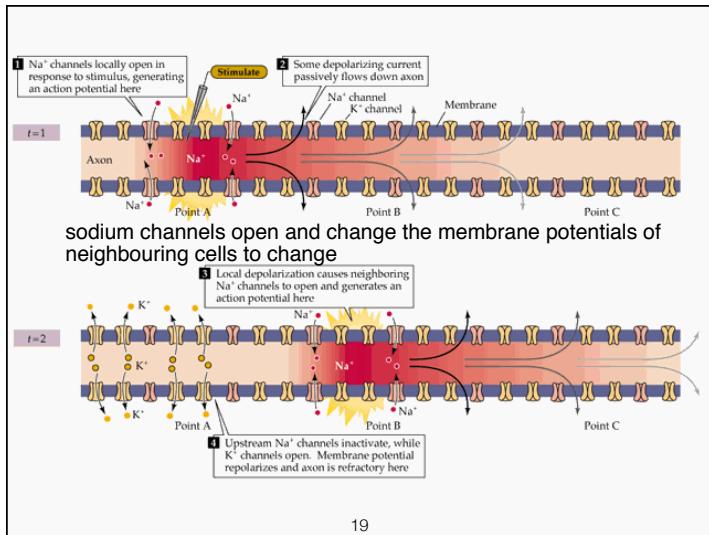
- Neurotransmitters (e.g. glutamate and GABA) released from a pre-synaptic cell can excite (depolarise) or inhibit (hyperpolarise) activity in the post-synaptic neuron
 - This relies on ligand-gated (i.e. neurotransmitter activated) ion channels
- If membrane polarisation reaches a threshold, voltage-gated ion channels open
 - Results from many excitatory inputs and few inhibitory ones
 - Strongly depolarises the neuron: Action potential

17

Action potentials



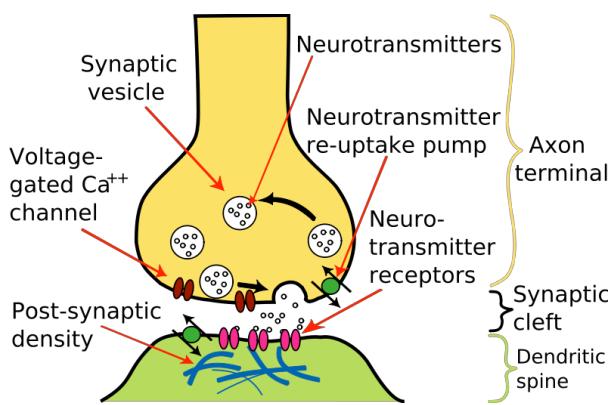
Once an action potential has been triggered, it travels from the input end of the neuron, the postsynaptic dendrites, along the length of the the neuron's main fibre, the axon, to provide inputs to the next layer, which is often in a different brain area some distance away.



19

- This process also relies on voltage-gated sodium channels.
- The spiking depolarisation at one location spreads to neighbouring locations to push their membrane potential above threshold. So the depolarisation spreads down the axon like a wire.

Neurotransmitter release



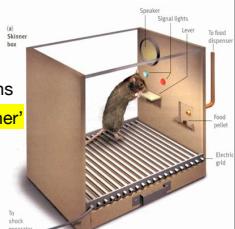
20

- On reaching the presynaptic terminal for the synapse to the next layer of processing, the action potential leads to neurotransmitter release.
- At rest, neurotransmitters are stored ready for use, in membrane bubbles called vesicles.
- The arriving action potential causes voltage gated calcium ions (Ca^{++}) channels to open, allowing calcium ions to flow into the cell.
- Ca^{++} binds to proteins in the axon terminal to change their shape, bringing the vesicles to the synaptic surface and releasing the neurotransmitter to activate the next postsynaptic neuron.

Weights in biological neurons

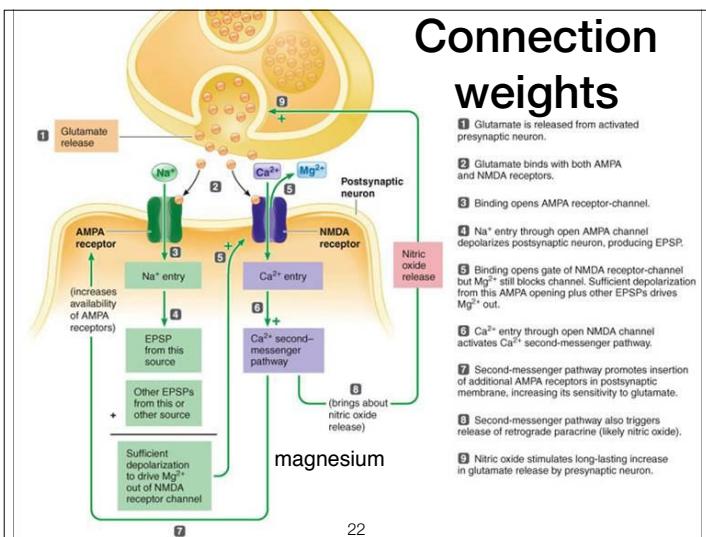
- Connection weights in biological systems don't depend on backpropagation of error
 - We are not trained by a supervised process
 - There is no 'correct' response in how we learn about the world
 - However, there can be adaptive and maladaptive responses/behaviours
- Instead, we learn mainly by unsupervised processes
 - To recognise patterns of activity we have seen before
 - To learn the statistics of the world we live in
- Long-term potentiation
 - Lasting enhancement of synaptic connections by co-activation of presynaptic and postsynaptic neurons
 - Hebb's postulate: 'Cells that fire together, wire together'
 - Also, 'cells that wire together lie together'
 - Reduce required connection length
 - Increase efficiency/speed

21



- The Skinner box is an apparatus used in animal experiments to deliver rewards (food) and punishments (electric shocks through the floor) when an animal performs (or fails to perform) certain behaviours.
- Conditioned responses like those learned from rewards and punishments are unlikely to work through backpropagation-like mechanisms affecting the whole network.
- They are more likely to work at a single decision stage, where the final decision is associated with a reward or fear response. This is enough to modify behaviour.

synapses become stronger as we release more neurotransmitters



Important to understand how this works

So, how does firing together lead to wiring together?

-Postsynaptic neural activity modifies synapse structure at both the presynaptic and postsynaptic terminal through a variety of mechanisms.

-These increase presynaptic neurotransmitter concentration and release, and the density of (active) postsynaptic receptors for the neurotransmitter.

Here are some example mechanisms.

AT END: Another important mechanism is an increase in the transcription of receptor genes in the cell nucleus, as cell polarisation affects the activity of gene transcription enzymes.

-Mechanisms like this seem very important in learning and memory, and seem to occur at all synapses.

-Several common drugs interfere with these processes. Alcohol and benzodiazepines like valium reduce spiking activity by activating GABA receptors and more reducing the probability of depolarisation. This depolarisation reduces memory formation, so these sedative drugs have a tendency to cause amnesia.

-Ketamine blocks the NMDA glutamate receptor, a vital step in strengthening weights. Ketamine is therefore very effective at inducing amnesia.

The story so far

- Network layers: layers of neurons at different levels of synapses
- Feature map: ?
- Filter: integration of EPSPs and IPSPs across the dendrite tree
- Threshold: activation of voltage-gated Na⁺ channels by above-threshold depolarisation by EPSPs
- Pool: ?
- Normalise: ?
- Learning mechanism:
 - Unsupervised: 'cells that fire together, wire together' (Hebbian learning)
 - Activity-dependent changes in synapse structure/strength (weights)

23

LOW LEVEL MODELS

AT END: So the weights on filters result primarily from learning.

-However, some filters appear hard-wired in early stages of vision. These seem to arise from evolution rather than learning, suggesting they provide a useful structure that subsequent layers can build on.

-In artificial deep networks for vision, the filters are the first level are often hard-wired edge detectors.

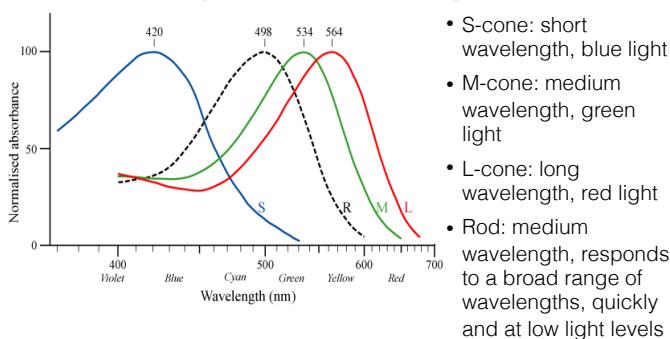
-Understanding these early filters gives an idea of which feature maps arise in biological vision, and how they are used by subsequent stages.

-Understanding this is important because the pooling and normalisation stages depend on the structure of these feature maps.

-So, let's look at some biological feature maps, and how those pool and normalise information.

-All of these structures and operations don't rely on a single neuron, but a larger group of neurons. So now we will look at larger structures containing large numbers of neurons.

Input feature maps: The photoreceptors



- S-cone: short wavelength, blue light
- M-cone: medium wavelength, green light
- L-cone: long wavelength, red light
- Rod: medium wavelength, responds to a broad range of wavelengths, quickly and at low light levels

The eye has four different sensors for light: three types of colour-sensitive cones sensor and one type of rod sensor.

-Rods don't carry colour information, but respond quickly and under low light conditions. Their output gives good information about fast events like moving objects.

-These form the input image to the visual processing network, and are essentially already four feature maps.

The photoreceptors

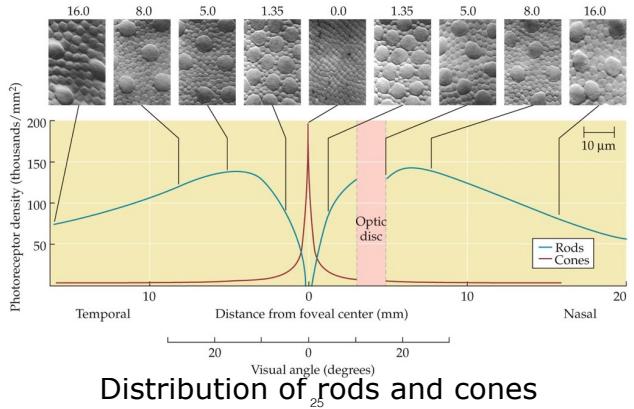
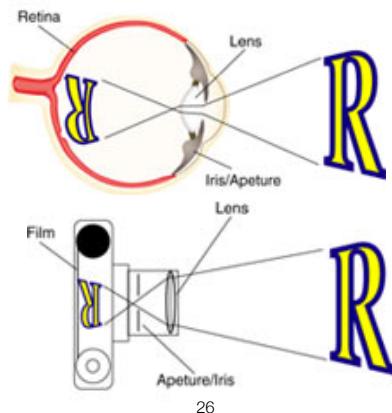
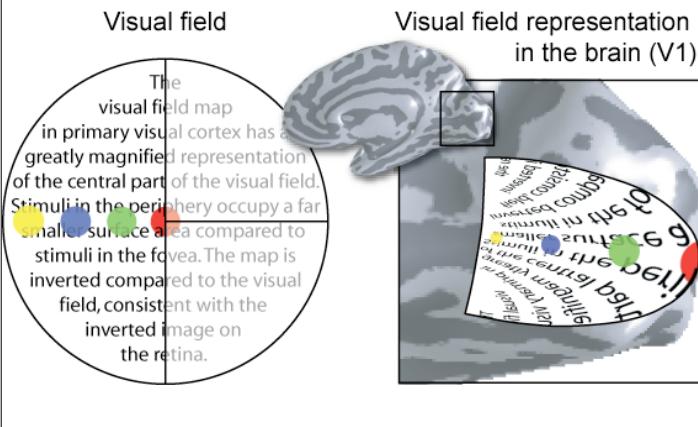


Image inversion



26

Cortical visual field representation



27

Unlike in a camera that usually provides the input to an artificial network, these light sensors are not evenly distributed across the image.

- Cone density is much higher in central vision ('where we are looking').
- Rods are much denser away from central vision. Because rods are physically large, we can cram more cones in.
- Because cones are smaller, they each respond to a smaller area of the image, so give finer vision where we are looking.
- But if we add these two curves together, we see that overall the density of photoreceptors drops off with the distance from central vision. So vision strongly over-emphasises the centre, where we are looking.
- This greatly reduces the computational load on the brain, while still giving high detail in central vision.

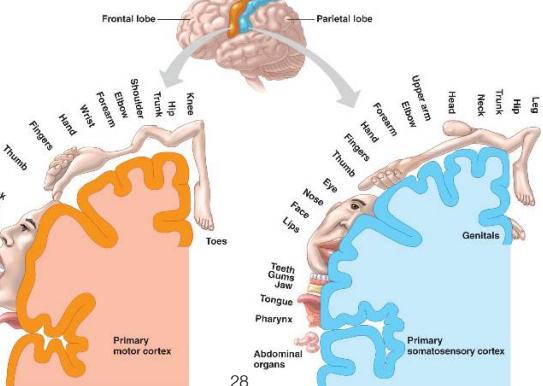
The brain's representation of visual space is also flipped up-down and left-right.

- This begins in the retinal projection of the image because rays of light are passing through a small hole at the pupil, much like an image is inverted in a camera sensor.

This over-representation of central vision and this image inversion continue into the cortex.

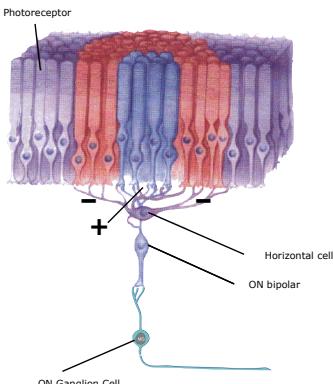
- You may think that this strange representation of vision would cause problems for perception.
- However, the location of information on the retina or cortex is not important when we think of the visual cortex as a set of feature maps, neural representations of image transformations, rather than thinking of this as an image.
- But it is very important that neighbouring locations are represented next to each other so that the spatial extent of a filter represents a continuous piece of the input image.
- This is particularly important when we consider that the dendrites of the next neuron will sample from this layer, and have a limited extent too. This is the main reason why the brain maintains spatial relationships at each level of processing.

Over-representation of important inputs



28

Spatial comparison filters: surround suppression



29

0	-1	-1	-1	-1	0
-1	-1	-1	-1	-1	-1
-1	-1	9	9	-1	-1
-1	-1	9	9	-1	-1
-1	-1	-1	-1	-1	-1
0	-1	-1	-1	-1	0

- All of the brain's sensory inputs similarly over-represent the most important parts of the input
- Here we see how the sensory and motor areas of the brain have larger, more detailed representations of our hands, faces and tongues: sensitive body parts that perform detailed movements.

-Similarly, vocal frequencies are represented in more detail than other frequencies in auditory areas.

-Artificial sensory systems generally don't use such distorted inputs, aiming instead to process everything in great detail.

-This has advantages and disadvantages. The main disadvantage is that processing the whole image input in similar detail is computationally intensive.

-The main advantage is that the cameras that provide the to artificial networks don't need to move to sample the important parts of the image in detail. Normally, the inputs are static images, so this isn't an option anyway.

-It will take considerable advances in artificial deep networks before they can process input images immediately, decide where the important details are, and move the sensors to sample those details.

The first filtering stages in the eye analyses the relationships between nearby locations, and simultaneously the relationship between the different colour maps.

-Effectively these are filters with a spatial spread and also a spread across the four colour feature maps.

-Let's look at the spatial aspect of the filter first. The unit has no real space but uses the short

- The retina has several stages between the photoreceptor and the ganglion cells of the optic nerve to the brain.
- The first of these compares the responses of a group of

The first of these compares the responses of a group of neighbouring photoreceptor cells, all linked in to its tree of dendrites.

These in the centre of the tree produce EPSPs, while those at

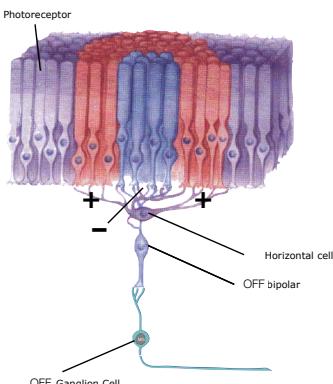
-Those in the centre of the tree produce EPSPs, while those at the edges of the tree produce IPSPs. So activation at the edges inhibits responses to activation at the centre.

-As a result, a point of light at the centre produces a stronger response than a field of light covering both the centre and the surround

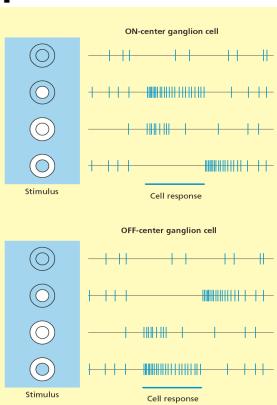
CLICK-Expressed as a convolutional filter, it might look something like this. If we have a bar of equal intensity running through the centre of the filter, we get a positive response (+28). A point of light will at the centre only will produce a larger positive response (+36). A field of light everywhere will produce zero response

-Just like a convolutional filter, there are overlapping copies of this cell throughout the retina. So a photoreceptor that falls in the negative zone of one horizontal cell filter, also falls in the positive zone of another.

Spatial comparison filters: surround suppression



30



- A complementary filter is operating in parallel, responding to darkness at the centre and suppressed by darkness in the surround.
- This is called an 'off' response, a response to no light. This might sound unnecessary, but features are often darker than the background, like the text on this page.

The corresponding convolutional filter might look like this, if we express darkness as negative, so darkness -1 multiplied by filter value -9 gives $+9$ activation. Darkness -1 multiplied by filter $+1$ gives -1 activation.

- Here we can see how these cells respond to different patterns of light and darkness throughout their responsive area.
- The set of on-centre and off-centre responses effectively form**

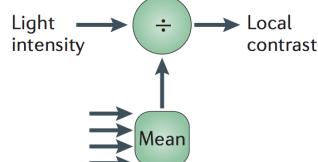
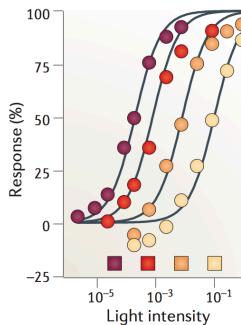
- The set of on-centre and off-centre responses effectively form separate feature maps for light contrast and dark contrast.
- Contrast is necessary in this description because a full field of light or darkness produces no response. More generally, the visual system

or darkness produces no response. More generally, the visual system responds to changes rather than constant inputs. These spatially-specific 'changes' can be thought of as 'features'. The area producing a positive response is called the receptive field.

- The area producing a positive response is called the **receptive field** of the cell, while the area suppressing that response is called the **suppressive surround**. These are technically parts of the retina, though they are typically thought of as parts of the image on the retina.
- In an artificial neural network, the term 'receptive field' is often used for the spatial extent of a filter. Properly, this should refer to its

-In biological systems, ‘receptive field’ always refers to the extent in the retinal image, partly because it’s really hard to determine the spatial spread of a cell’s inputs. We’ll look at some methods to do this in the next class.

Surround suppression and normalisation



31

As a result of this surround suppression, the response to light intensity inside the receptive field is normalised by the light intensity outside the receptive field.

-The light level in the surround is used to normalise the light level in the receptive field, yielding local contrast.

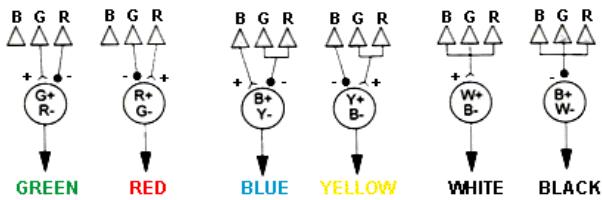
-This closely resembles the normalisation operation of deep networks, and effectively converts light level into contrast level at the first stage.

-However, this mean light level used here is probably not taken over the whole image: there is a limitation to the distance that the neuron can be connected to.

-Furthermore, the mean light level is probably restricted to a single feature map.

-It's unclear whether these differences are better or worse than global normalisation. They allow more complex patterns of normalisation, which might be useful if an image has light parts and dark parts, but are limited in spatial extent.

Colour map comparison filters: colour opponency



32

Together with this spatial comparison, there is also a comparison over the different colour cones.

-This forms four further feature maps, one each for greenness, redness, blueness, and yellowness.

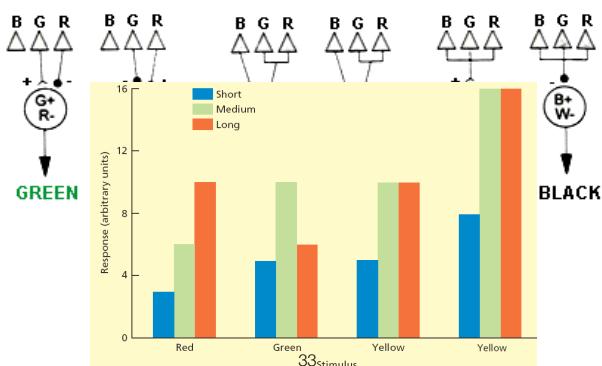
-There are relatively few blue cones, so redness and greenness are each determined by the difference between their intensity and the intensity of the other. Where both red and green are present at equal intensity, for example in white light, there is no response in these channels.

-Blueness is determined by the difference between the intensity of blue cone activation and the intensity of other cone's activations.

-Yellow light activates both green and red cones, but not blue. So yellowness is the average of red and green cone activations minus blue cone activation. Again, white light will produce no response.

-To detect white and black, all three cones activate or inhibit the response.

Colour map comparison filters: colour opponency



This is approximately optimal for representing colour information

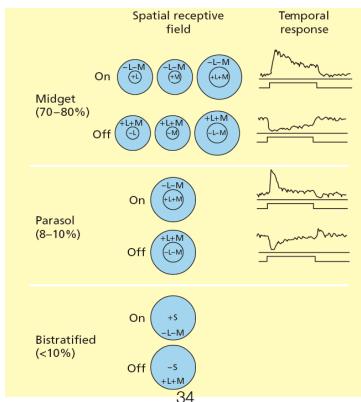
-Beyond the photoreceptor, we have three opponent channels for the ratios of R vs G, Y vs B and light vs dark

-Using these ratios, each individual colour representation gives an accurate description of the colour of light falling on the retina, not its intensity.

-To get this colour-based description, there would always need to be a stage to COMPARE the ratio of activation of the three cone types

-Doing the comparison early (in the retina) produces that representation as early as possible, so it can be used for all later processing.

Colour and position comparison together



So far, we have treated the spatial comparisons and the colour comparisons separately, but actually they are together, so the filters cross both colour and space: they cover a limited spatial extent differently for different feature map.

-In the case of the bistratified retinal ganglion cell, there is no suppressive surround around the activating centre, but these filters do have a limited spatial spread.

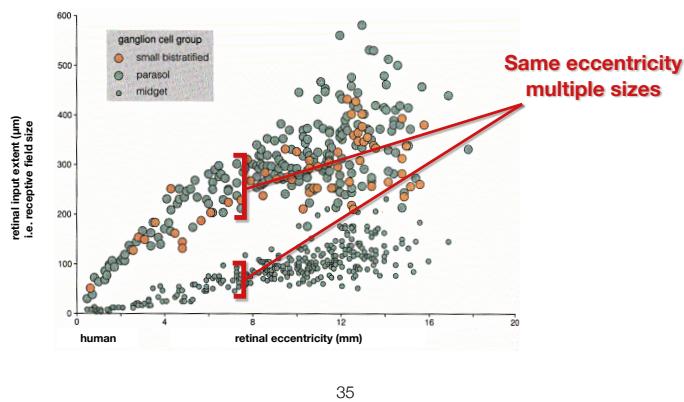
-The midget ganglion cells have small receptive fields and small suppressive surrounds, with different colour inputs to the receptive field and the surround.

-Some of these also respond to any colour in the centre receptive field and also the surround. Because blue cones (S-cones) are relatively low density, these are not used here. This forms the black-white opponent filter.

-Parasol cells have a similar structure, with no colour selectivity but a larger spatial extent.

-These also stop responding quickly after the stimulus turns on, even if the stimulus stays on the retina, so they are useful important for signalling temporal change.

Receptive field extent and spatial frequency



As we go from central to peripheral vision, the spatial extent of these integrating filters, the horizontal cells and retinal ganglion cells, increases.

-So unlike in artificial deep networks, filters are not the same over the whole visual field

-Even at the same retinal location, we see two distinct classes of retinal ganglion cells integrating spatial information over different extents, the midget and parasol cells.

-And even with these cell populations, there is a range of spatial extents.

To produce the maximum response, we must give an input with the right spatial frequency, the right spacing between light and dark areas.

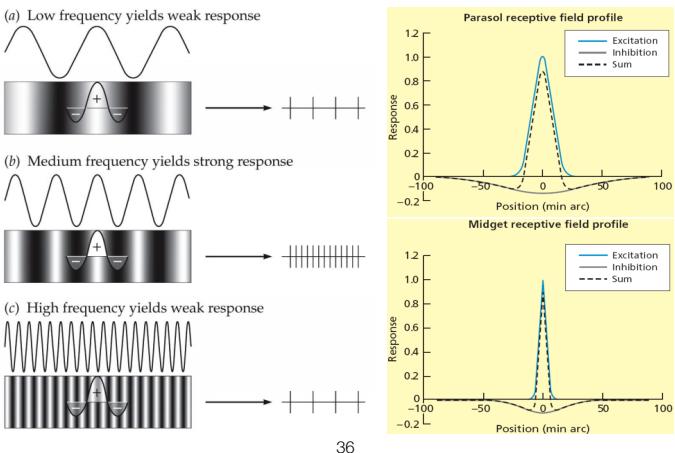
DESCRIBE EXAMPLES

-Receptive fields with this arrangement have a range of sizes.

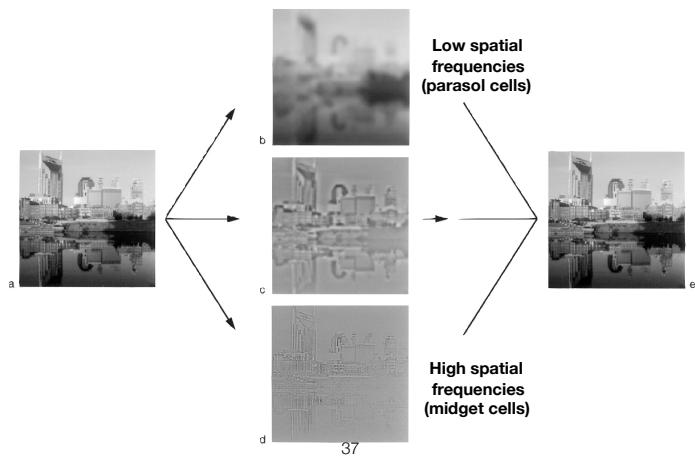
-Parasol cells have large receptive fields, responding to larger areas of the visual field.

-Midget cells have small receptive fields, spatially localised responses.

Receptive fields and sine gratings



Spatial frequency components of an image



37

Using these differently sized receptive fields, effectively different filter sizes, the retina can break down any image into different components at different spatial frequencies.

-These have different properties. At low frequencies, we miss changes that are at high frequency, so we can see big blocks of light and dark with no edges

-At high frequencies, we can only see the edges, with no lower-frequency changes in color because these would stimulate both the center and the surround.

-But if we keep these different parts, we can add them back together to recreate the original image

-So these representations of different spatial frequencies also form different feature maps, even at the first stages of processing.

-This is quite different from the process in artificial deep networks, where each layer is convolved with filters of a single spatial size.

-In artificial networks, larger-scale features can only be detected by filters at later stages, after intermediate stages of spatial integration and pooling.

Retinal feature maps

- In the input image from the photoreceptor:
 - Red light intensity, green light intensity, blue light intensity (cones), overall light intensity (rods)
- By the retinal ganglion cell:
 - Redness, greenness, blueness, yellowness, lightness, darkness
 - Many spatial frequency ranges
 - Many combinations (not all)

38

So the filters in the retina perform a range of comparison operations on their inputs

Normalisation and pooling

- Normalisation
 - Inhibition of activity by average nearby activity
- Pooling
 - No distinct problem of computational load
 - But progressive layers do have smaller representations, and more ‘feature maps’
 - No distinct pooling stage: just use fewer neurons to sample previous layer

Although features can be already identified at earlier stages vs. later stages in an artificial neural network

39

Shared weights?

- Weights at each synapse develop independently from other neurons
- No fixed filter extent at the retina
 - Receptive field (filter) sizes vary with eccentricity
 - Multiple receptive field sizes at same location
- However, if a filter structure develops at one part of the feature map to compute something useful, similar filter structures are likely to develop elsewhere.

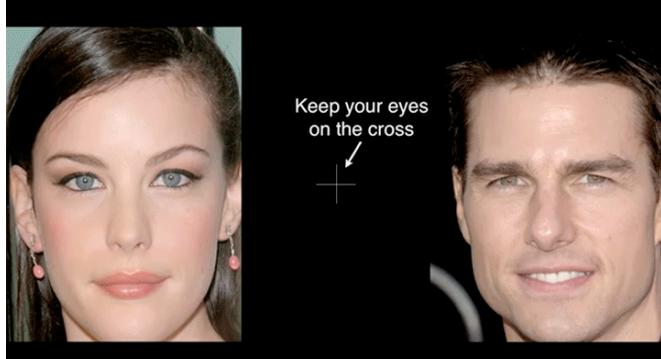
40

Remember that a major benefit of shared weights is that convolution is a fast matrix operation, but relies on a single filter. This limitation does not apply if filters are all made from distinct biological elements. There is no convolution operation, and no single processor core that needs to process the operations of all units.

Another benefit of shared weights was that the target for machine learning is a very small number of filters if the same filter applies across the whole feature map.

Again, this is not important if filters are all made from distinct biological elements. Hebbian learning, unlike back propagation, operates on every synapse independently.

Differences in perception across the visual input



41

-Indeed, differences in object recognition abilities across the visual image are clear in perception

-Try identifying these faces using your peripheral vision only

-If we move our gaze between the cross and the faces, we can see how faces appear with and without high spatial frequencies

-This shows that even at higher levels like object recognition, the brain's filters for visual information are not the same over the whole visual image

Deep learning in biological neurons and networks

- Shares most operations with artificial networks (except pooling), but typically with some differences
- Filter/convolve:
 - Dendritic tree synapses with many neurons
 - Different synapses have different strengths (weights)
 - Filters cross multiple feature maps, even at earliest stages
- Threshold:
 - Activation of voltage-gated ion channels
 - Following summation on EPSPs and IPSPs
 - Leads to action potentials
- Normalise:
 - Inhibition of activity by local mean activity
- Learning
 - Hebbian learning: 'cells that fire together wire together'
 - Strengthens responses to patterns of activity previously causing with responses

42

Major differences to artificial deep networks

- Feature maps do not represent the whole image with the same detail: strong bias towards sensitive/important parts
 - Sensitive areas (hands, eyes) can move to sample task-relevant information
- Filters have a limited spatial extent, but not a fixed extent
 - Multiple spatial scales analysed at the same layer
 - Image layout is maintained in later stages, allowing analysis of spatial relationships at many levels
- Filter weights not shared across each layer, but similar filters Only if they are biologically useful!
likely to emerge in many places
- Maximum response is always part of the threshold
- Normalise of activity by local (not global) mean activity
- Learning is generally unsupervised

43

DCNN are not necessary for doing this, but we are only using these because of computational cost!
