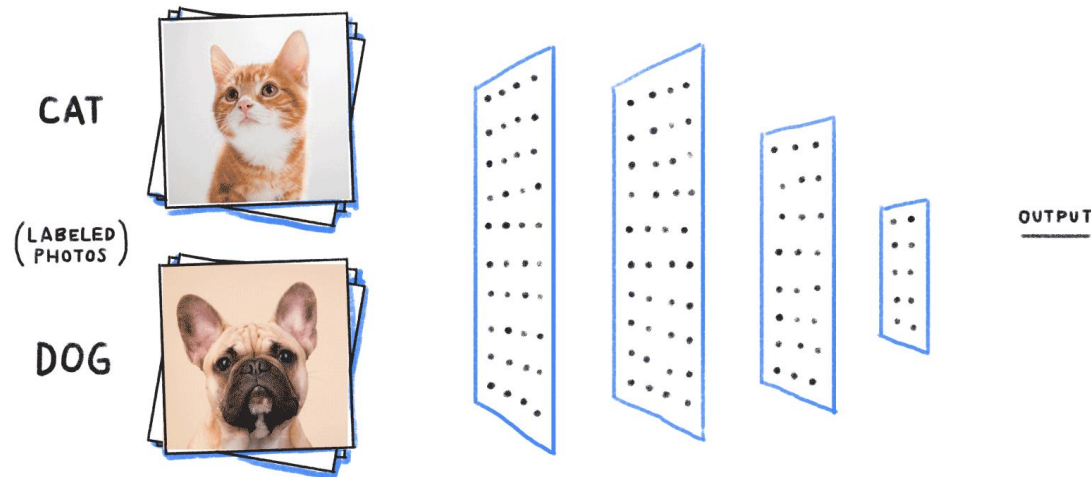


Deep Learning Fundamentals 2 – Convolution Neural Networks (25t4)



Michelle Yu

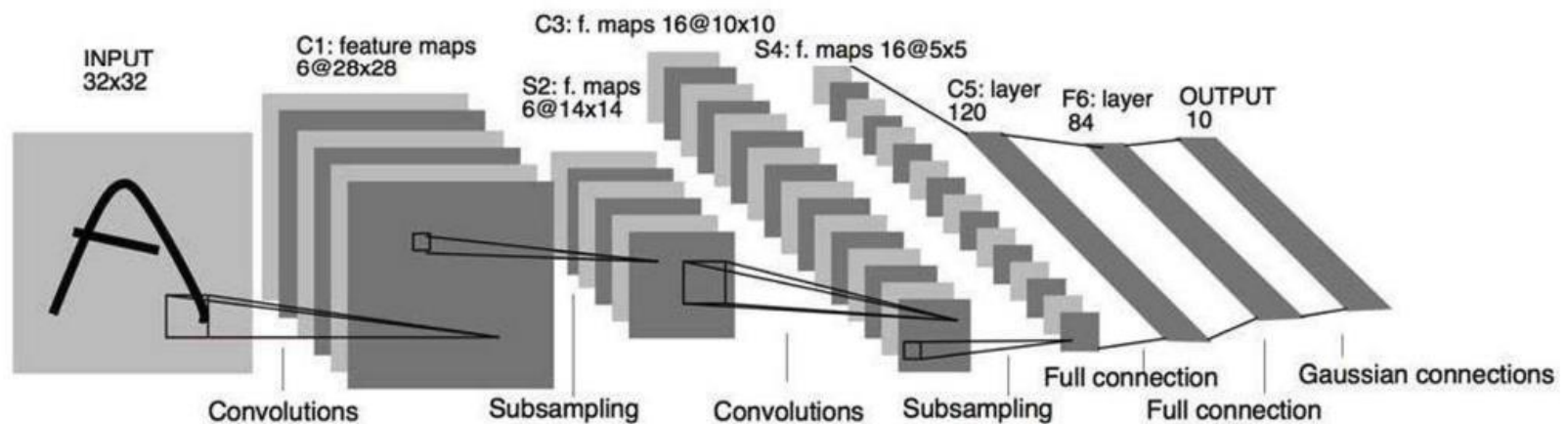
CBB 752 Spring 2025

Overview

- Convolutional operations in image classification
- Application of CNNs in regulatory genomics
 - Learn meaningful representations from DNA sequences (Basset)
 - Predict and precisely locate functional genomic regions (DECODE)
 - Map DNA sequences to enhancer activity and uncover TF motif syntax rules (DeepSTARR)

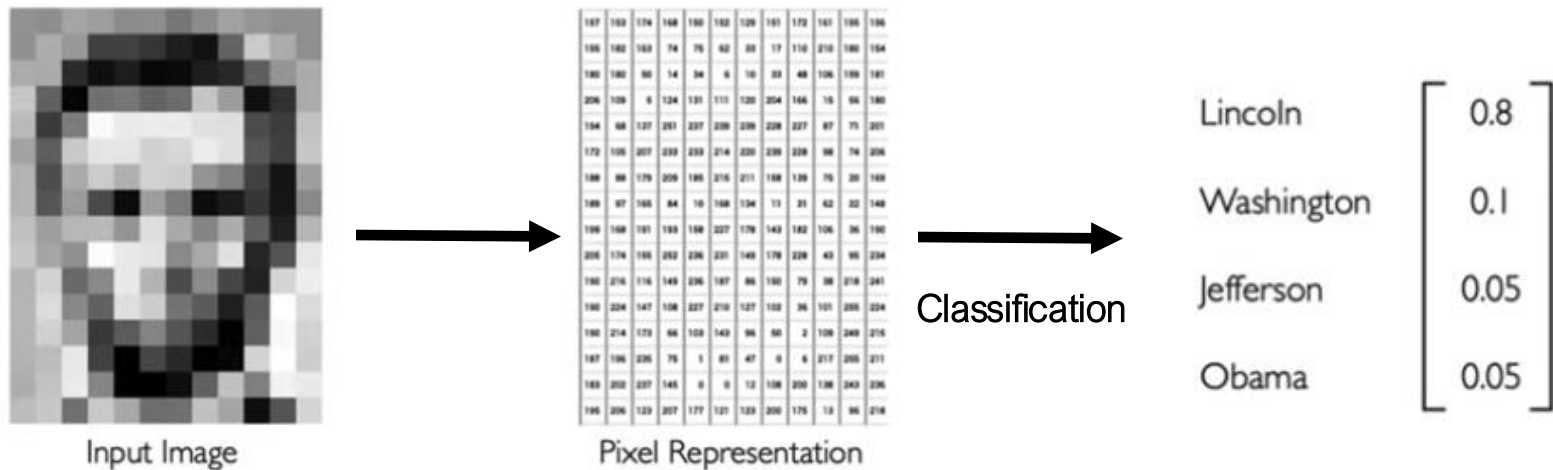
Convolutional Neural Network: LeNet (1998)

Convolutional Neural Network: LeNet (1998)



LeCun et al., 1998

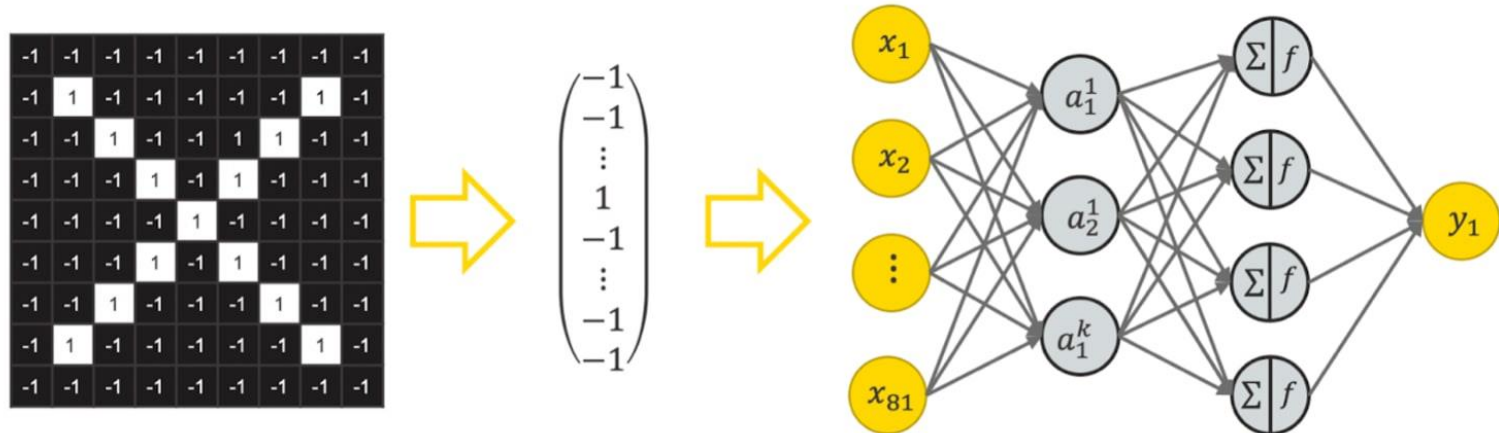
Classification in computer vision



Adaptative
feature learning
w NN

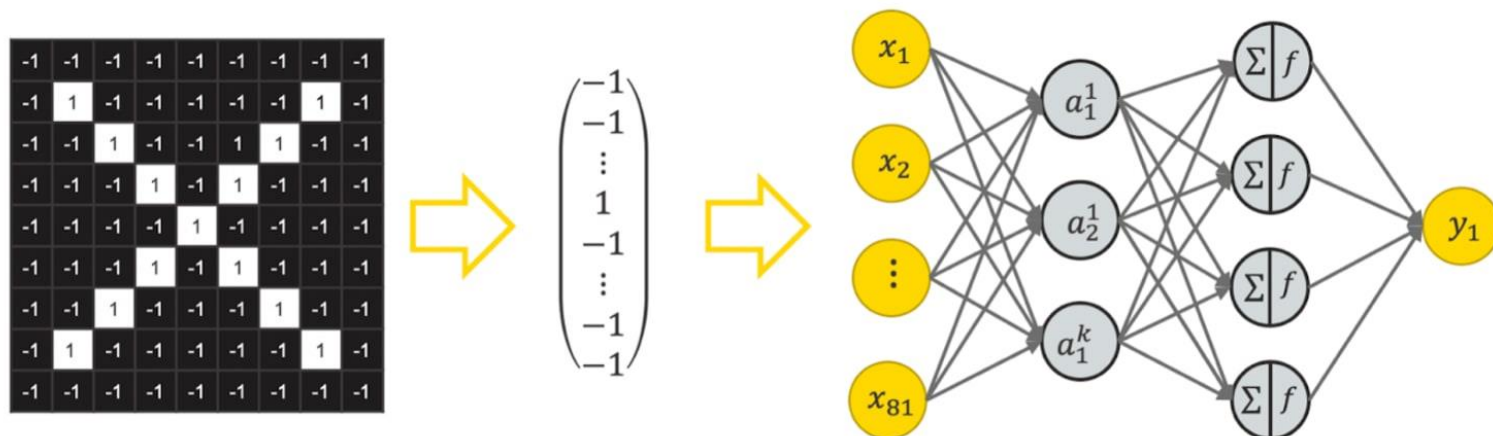
How do we effectively extract and detect features from image data?

Why CNN over Feed-Forward Neural Nets?



- Represent an image with a flattened tensor
- Problems of using fully connected layers for image classification:
 - Too many weights
 - $224 \times 224 \times 3$
 - Loss of spatial structure!

Why CNN over Feed-Forward Neural Nets?



DNN works much worse than a shallow CNN even on MNIST!

Error rate: $\sim 1.0\%$ vs. $\sim 0.60\%$



MNIST
Task

Classification in computer vision



Input Image



187	163	174	168	180	182	129	181	172	141	185	186
185	180	163	74	76	62	88	17	110	210	188	184
180	180	88	14	84	6	10	88	48	106	188	181
206	109	6	124	131	111	120	254	166	16	98	180
184	88	127	281	237	288	288	227	87	71	201	
172	105	257	239	233	214	220	238	228	88	74	206
188	88	179	209	185	215	211	188	138	76	28	188
188	87	185	84	10	188	134	11	31	42	22	148
188	168	181	183	188	227	178	143	182	106	38	180
208	178	188	282	236	231	148	178	228	43	98	234
180	216	176	148	236	187	88	180	78	38	218	281
180	224	147	158	227	218	127	188	38	181	208	224
180	214	178	88	108	148	88	88	2	108	248	216
187	186	236	78	5	81	47	0	6	217	288	211
188	202	227	148	0	0	12	108	202	138	243	236
188	206	122	227	177	121	138	288	178	18	88	218

Pixel Representation



Classification

Lincoln

Washington

Jefferson

Obama

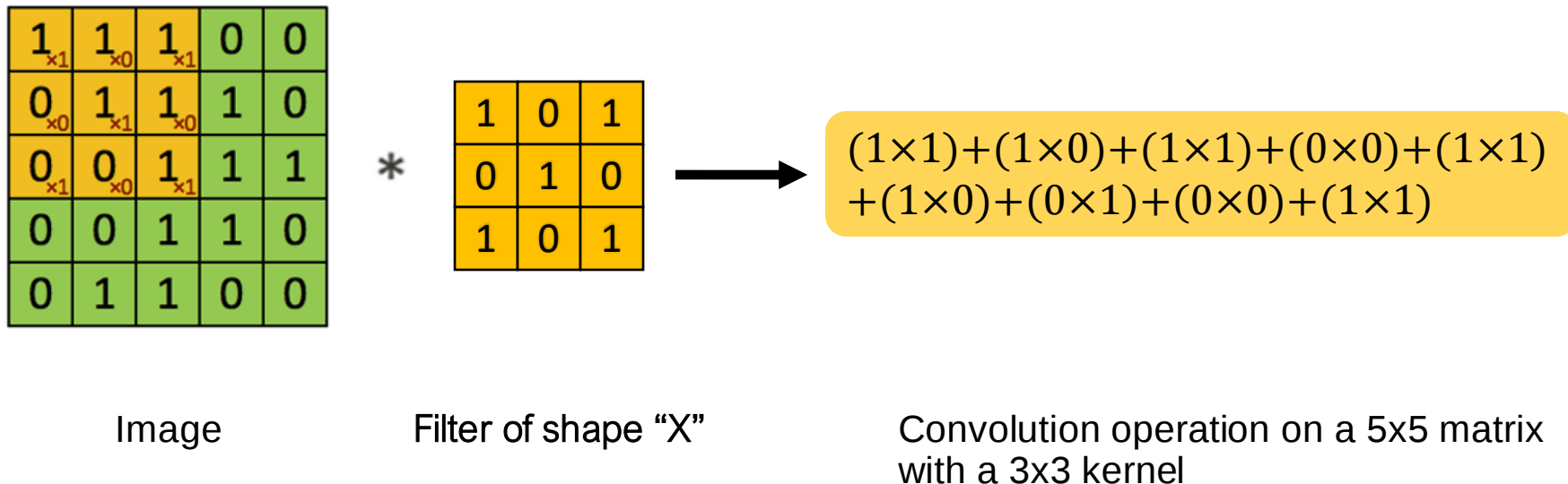
$$\begin{bmatrix} 0.8 \\ 0.1 \\ 0.05 \\ 0.05 \end{bmatrix}$$

Adaptive
feature learning
w NN

How do we effectively extract and
detect features from image data?

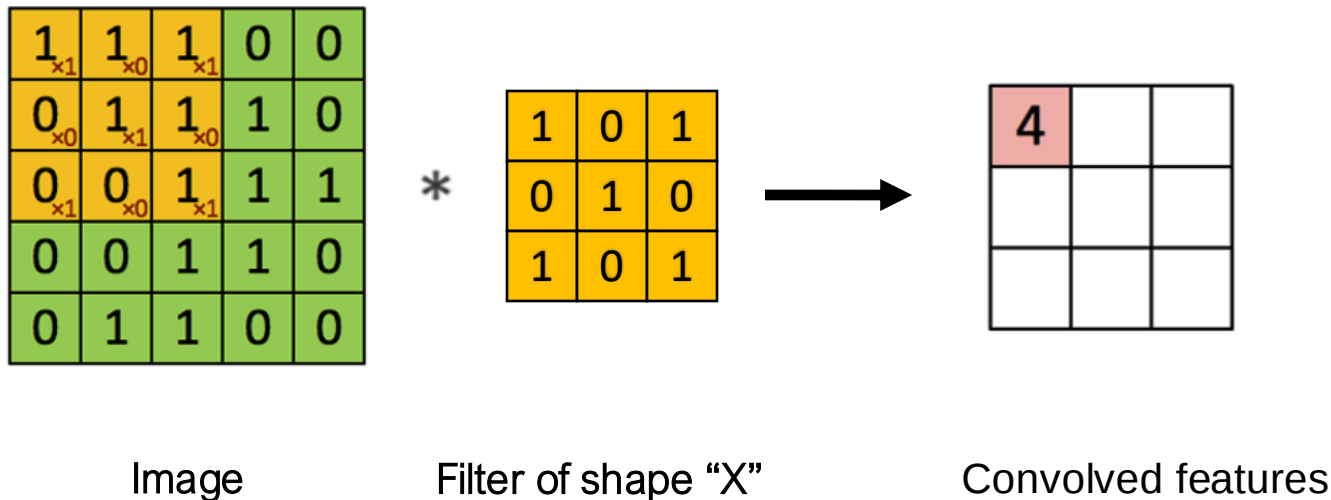
Use convolutional filters

Detect features in an image via **Kernel / Filter**



- A **filter** (weight matrix) detects features in the input.
 - They are initialized randomly and updated through backpropagation during training.
- **Convolution** identifies where these features appear in the image.
 - **Element-wise multiplication** between filter and image region.
 - **Sum** the results.

Detect features in an image via **Kernel / Filter**



- The **filter** is applied to **3x3 image patches** to detect patterns or features.
- Convolution produces **higher values** when the filter matches the pattern in the image; **Lower values** indicate weaker matches or opposing patterns.
- The resulting **feature map** highlights **where the feature appears** in the image.

Detect features in an image via **Kernel / Filter**

1 x1	1 x0	1 x1	0	0
0 x0	1 x1	1 x0	1	0
0 x1	0 x0	1 x1	1	1
0	0	1	1	0
0	1	1	0	0

Image

4		

Convolved
Feature

- The **filter** is applied to **3x3 image patches** to detect patterns or features.
- Convolution produces **higher values** when the filter matches the pattern in the image; **Lower values** indicate weaker matches or opposing patterns.
- The resulting **feature map** highlights **where the feature appears** in the image.

Detect features in an image via **Kernel / Filter**

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

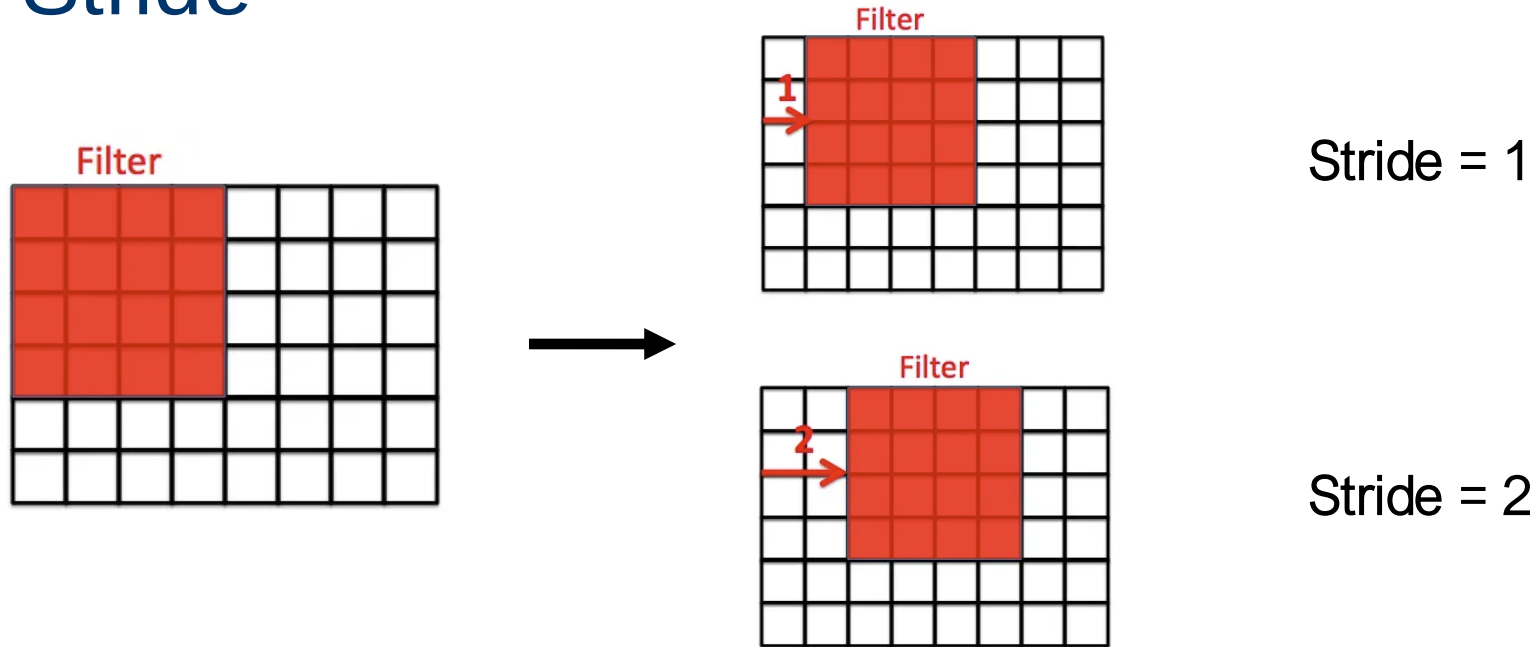
Image

4		

Convolved
Feature

- **Parameter sharing.** In convolutional layers, the **same weights / same filters** (parameters) are applied across different spatial positions of the input.
 - For each feature map, neurons share weights, resulting in far fewer parameters compared to fully connected layers.
- **Depth.** Each convolution operation applies **multiple filters** (kernels) to the input, producing **multiple feature layers** (feature maps).
 - Each filter detects a different pattern, such as edges, textures, or shapes.
 - By stacking these feature maps, the network can capture a wide variety of features in the input image.

Stride



- **Stride** determines how far the filter moves with each step (e.g., 1 pixel, 2 pixels).

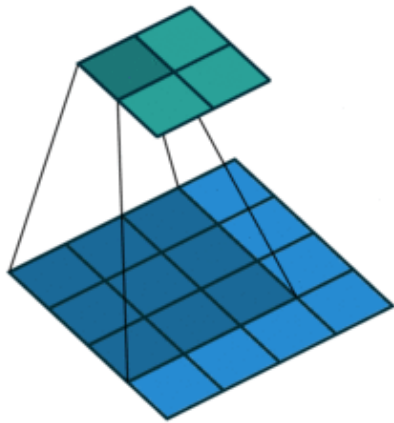
Zero-Padding

- **Padding** adds extra pixels (commonly zeros) around the image boundary to prevent loss of information at the edges.
- Without padding, pixels at the edges are used less in convolution, leading to loss of information and reduced output size.

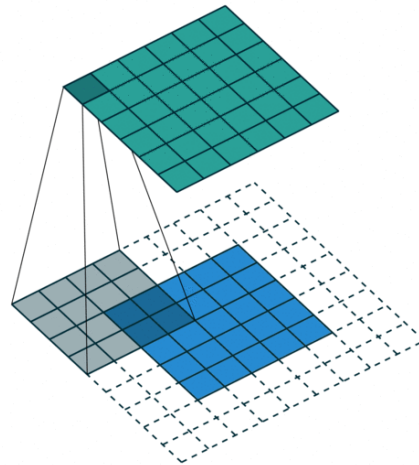
Image

0	0	0	0	0	0	0
0						0
0						0
0						0
0						0
0						0
0	0	0	0	0	0	0

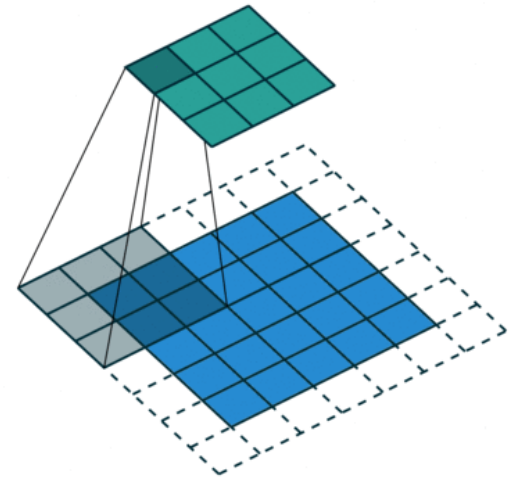
2D convolution examples



No padding,
Stride = 1

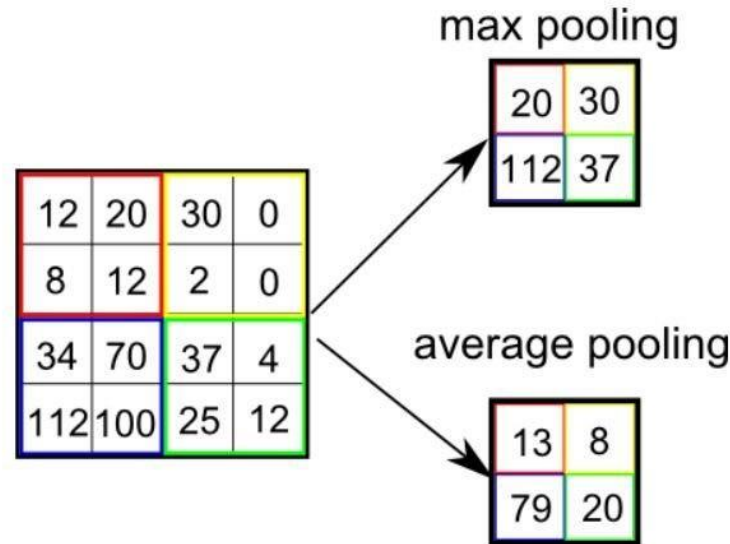


Padding = 2,
Stride = 1



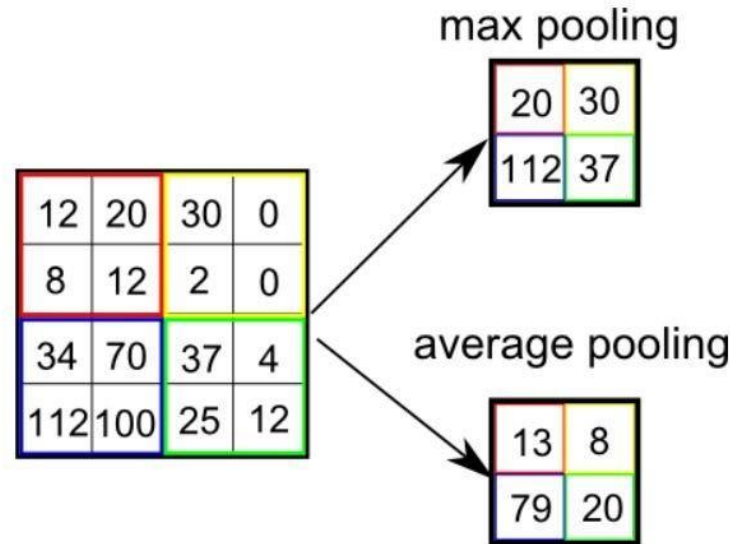
Padding = 1,
Stride = 2

Pooling



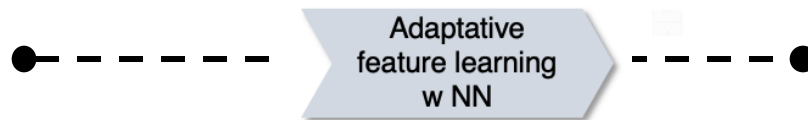
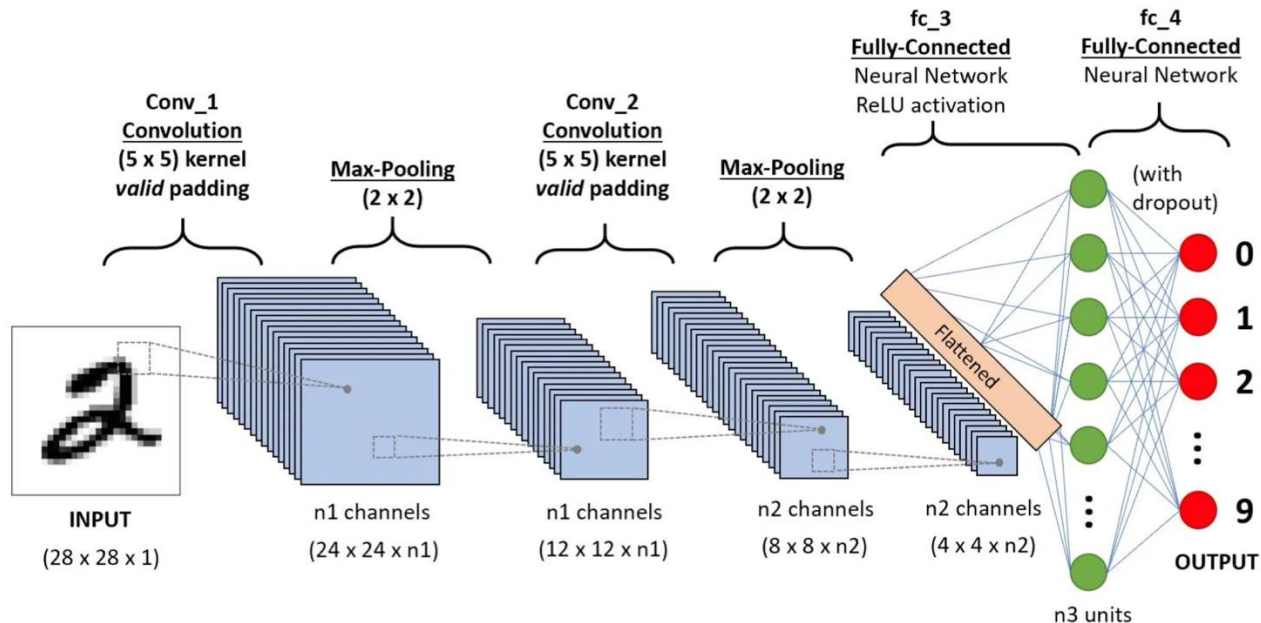
- Pooling is often applied after convolution operation to simplify the output data.
- **Max Pooling:**
 - Selects the **maximum value** within each region covered by the kernel.
- **Average Pooling:**
 - Computes the **average value** within each region covered by the kernel.

Pooling



- **Max Pooling** also acts as a **noise suppressant** by discarding noisy activations.
- Pooling simplifies data and improves computational efficiency.
- Convolution performs a **linear operation**; non-linearity is often introduced through subsequent **ReLU activations** and **pooling layers**.

CNN learns features directly from the data

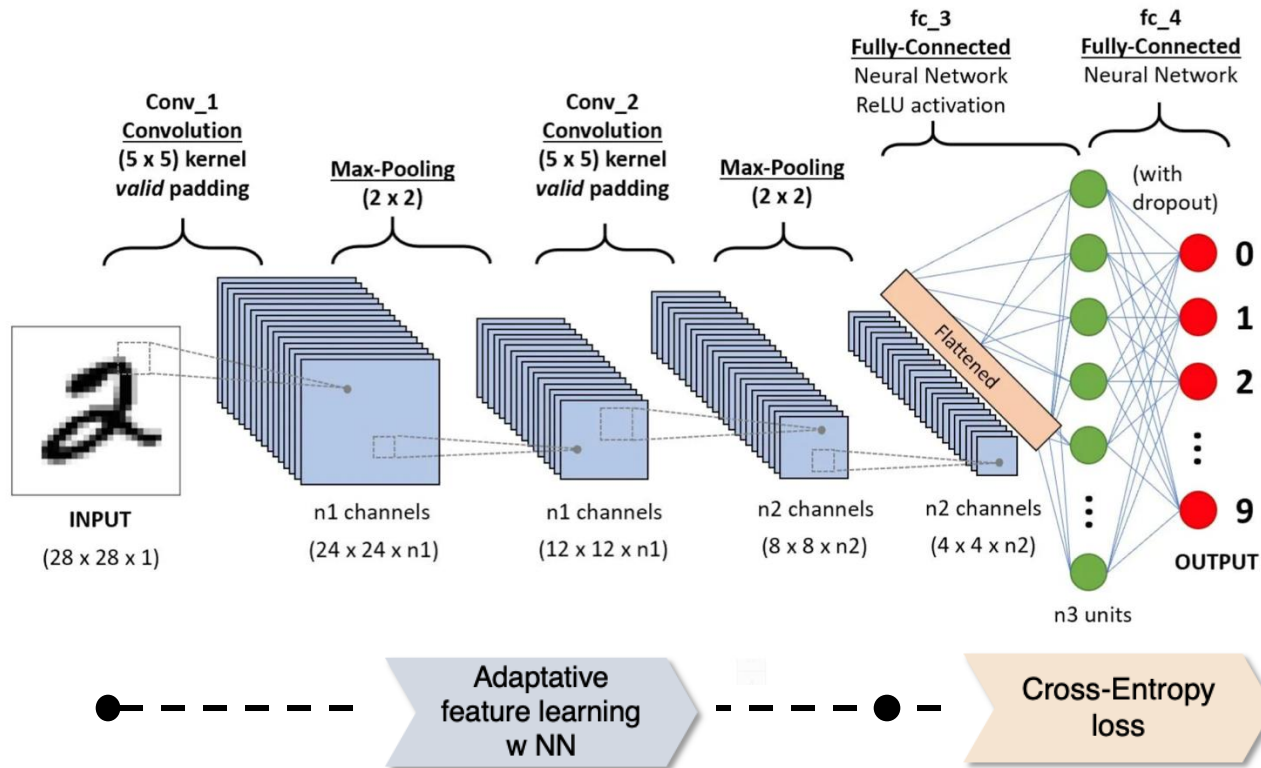


The classification task drives the feature extraction in CNN

1	0	1
0	1	0
1	0	1

Again, a **convolutional filter** is a matrix of **weights**, initialized randomly and updated through backpropagation during training.

CNN learns hierarchy of features



Low-level features:

Pixel intensity changes

Edge detection

Local texture variations

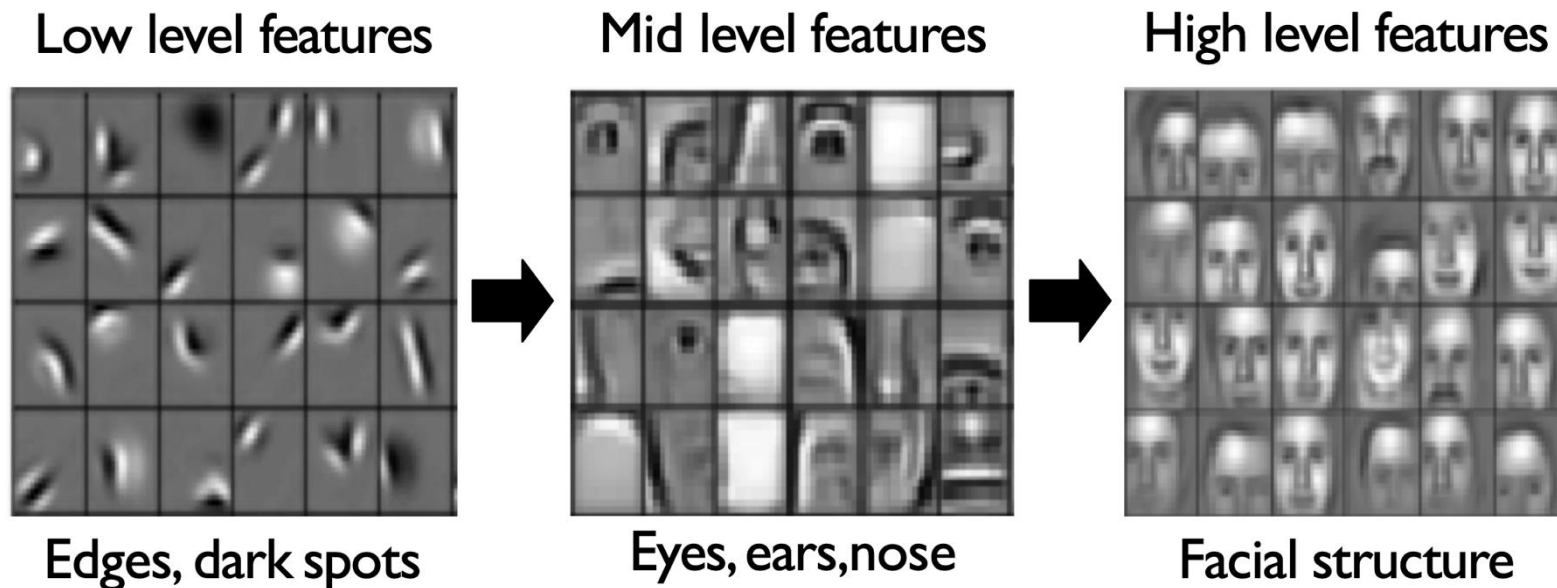
High-level features: (later layers of CNN)

Digit shapes (e.g., circular for "0", vertical line for "1")

Stroke thickness and curvature

Spatial relationships between different parts of a digit

Hierarchical representation learning



- Learn hierarchy of features directly from data (rather than hand-engineering them)

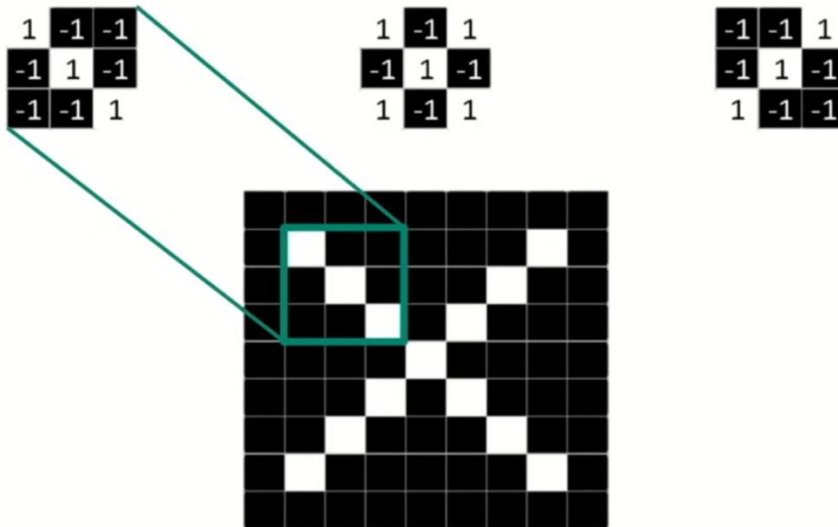
The above images are **not** direct filter weights (this visualization is for demonstration purposes; they are filters projected back into pixel space, to show what features each layer is detecting).

In practice, CNN filters are often less visually interpretable.

Filters to detect X features

Filters to Detect X Features

filters

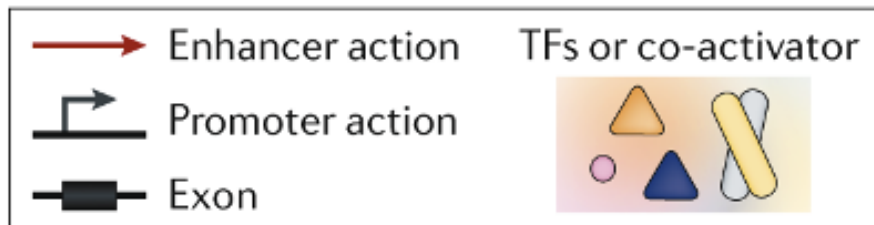
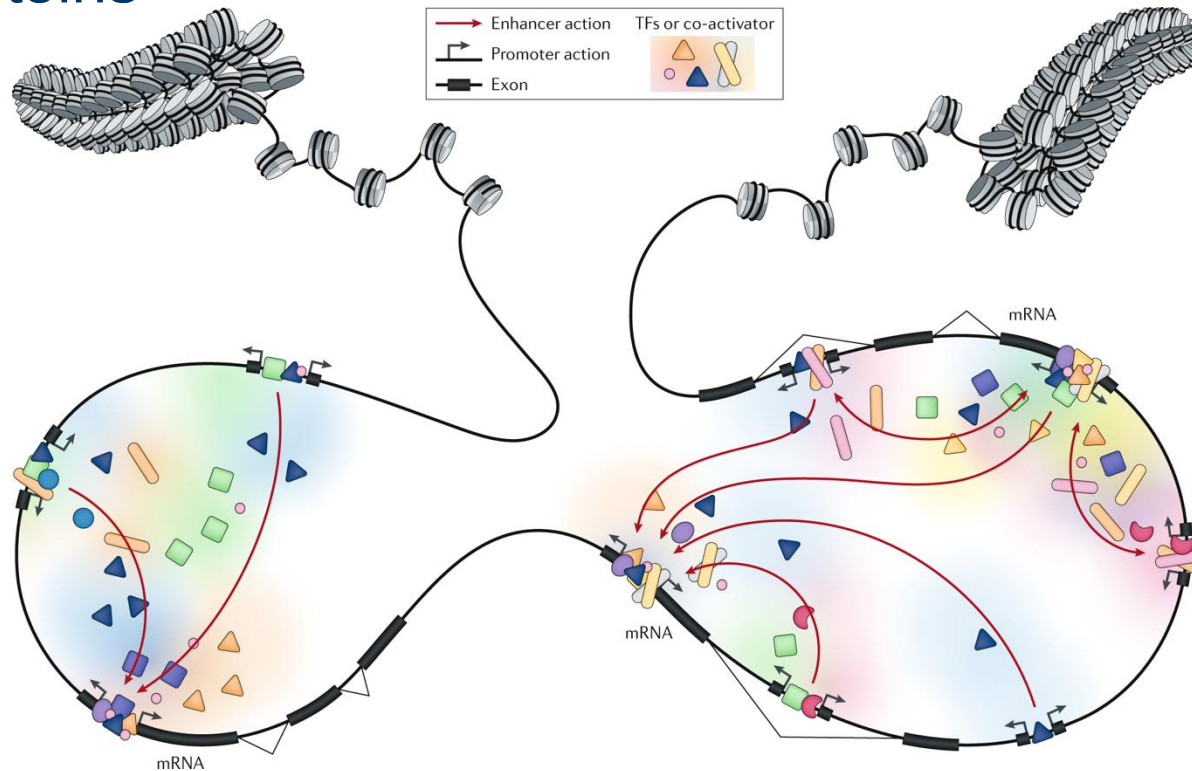


- CNN filters are **learnable weight matrices** with **defined sizes** (e.g., 3×3, 5×5) and learned features from data during training.
- Filters stack hierarchically:
 - lower layers: simple features (e.g., edges)
 - deeper layers: complex patterns (e.g., “X”).
- Modern CNNs can go very deep—e.g., ResNet: 34 layers (uses skip connections to avoid information loss).

Overview

- CNNs in image classification
- Application of CNNs in regulatory genomics
 - Learn meaningful representations from DNA sequences ([Basset](#))
 - Predict and precisely locate functional genomic regions ([DECODE](#))
 - Map DNA sequences to enhancer activity and uncover TF motif syntax rules ([DeepSTARR](#))

Transcriptional regulation requires the complex coordination of many proteins



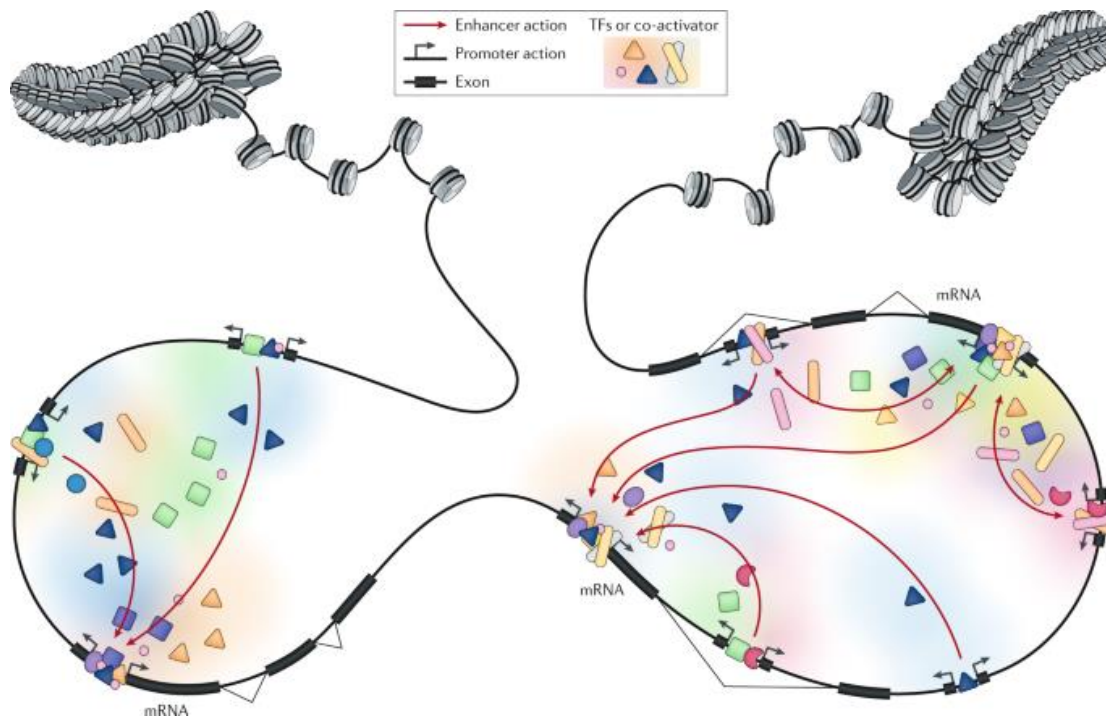
CREs – cis-regulatory elements

Promoters: Initiate transcription near the gene

Enhancers: Regulate gene expression from a distance

Motifs: Short DNA patterns for TF binding

Overview of gene regulation



- ✓ Where are the regulatory regions and protein binding motifs located?
- ✓ How strongly do transcription factors bind to DNA?
- ✓ How do these motifs interact with each other (within CREs or across different CREs)?
 - e.g. Competitive binding, collaborative binding.

How DNA Sequences Drive Transcription?

TCCAAATCAAAACAGTTGTATTATTAGAACTGAGGGCTAAAACTGTGCACATACACAGACACACATATTATTTAATATAGATTTTCAATAATTGGTCTAGGATAAGG/
 TAAGCAAGAAGAAAAACAAAGACTGTTACTATGGAAAAATGAAATAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTCTTCACTTCTTA/
 CAATAAATCACATTAATTCCTTATCTCATGTGAAATTTTCATATTATGATTGATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTT
 CCAGTATTATGTTCTAGGCATTGGGGATACCATGTTCAAGACAGACTATGATTACAGGATCAGATGTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTA/
 GTAATTGATGCTAGAAAGACAATGAAACAGAGCCATGTGACCAATGAGAGAGATGAGGGTGGCAGCAGCCTGTTTATAGATAAGGTACCTGATTGGTGGGATTGGA
 ATATGCCTTAATGATATGAAAGAACCATTATGGAAGGCCTAGCATTAAAAACCGTCTAGGCAGAATGAGCAGCAAGTGCAAGGGTCTGGATAGGAATGAGCTC/
 TATGGAAAAATGAAATAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTCTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAAC
 TGAATTTTCATATTATGATTGATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTGAATAATAAGTCCCAGGCACAAGAC/
 CCATGTTCAACAAGACAGACTATGATTACAGGATCAGATGTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAGTAAATAAAGTTAATTTCAAGTTG/
 TAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTCTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACAATAAATCACATTAAT
 GATTGATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGG
 GACTATGATTACAGGATCAGATGTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAGTAAATAAAGTTAATTTCAAGTTGTAATTGATGCTATCCCAG
 ATTGGGGATACCATACCTGTCAATGTTATTAATATTTTAGGAACAATAAATCACATTAATTCACATGCAAAAGAGGAAATCTCCATATCATGCTTGTCACTCGTTT/
 CGTGTGTAACCATTCCTCAGAATTTTAAACAATAACAAATCAGGGCTGAATGTGGCCAACATGCAAGAGGAAATCTCCATCTGTCCAAATCAAACAGTTGTATTA/
 ACATACACAGACACACATATTATTTAATATAGATTTTCAATAATTGGTCTAGGATAAGGATAATATACAGAGAACATGCCAAAAGTTTAAAGCAAGAAGAAAAACAAAGA/
 TTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTCTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACAATAAATCACATTAATTCCTT/
 ATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATT/
 ATGATTACAGGATCAGATGTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAGTAAATAAAGTTAATTTCAAGTTGTAATTGATGCTAGAAAGACAA/
 GAGATGAGGGTGGCAGCAGCCTGTTTATAGATAAGGTACCTGATTGGTGGGATTGGAAGACCTCTCTGAGATTAGTGTCTTCAAGATATGCTTAAATGATATGAAAGA
 AAACCGTCTAGGCAGAATGAGCAGCAAGTGCAAGGGTCTGGATAGGAATGAGCTGGATATACCTCAAGGAAGAAAGAGAAACTATGGAAAAATGAAATAGATTTT/
 TTAATTTACTTTTCTTCTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACAATAAATCACATTAATTCCTTATCTCATGTGAAATTTTCATATTATGATTGAT
 ATTTATCATTTTTCATTCAATAAATATTTTGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGATACCATGTTCAACAAGACAGACTATGATTACAGGATCAGAT
 AATTCGACTGAGAATAAAACAGACACTAAACAAGTAAATAAAGTTAATTTCAAGTTGTAATTGATGCTATCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGATACCATACCTGTCA
 CTTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACAATAAATCACATTAATTCCTTATCTCATGTGAAATTTTCATATTATGATTGATACCTTTAAATGTCATTT
 TCAATAAATATTTTGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGATACCATGTTCAACAAGACAGACTATGATTACAGGATCAGAT
 AAACAGACACAACAAGTAAATAAAGTTAATTTCAAGTTGTAATTGATGCTATCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGATACCATACCTGTCA
 TCACATTAATTCACATGCAAAAGAGGAAATCTCCATATCATGCTTGTCACTCGTTTATCAGAGGCCAAATGTTTTCTTGTAAACGTGTGAAAAACATTCTCAGAAT
 TGTGGCCAACATGCAAAAGAGGAAATCTCCATCTGTCCAAATCAAACAGTTGATTATTAGAACTGAGGGCTAAAACTGTGCACATACACAGACACACATATTAT
 GGATAAGGATAATATACAGAGAACATGCCAAAAGTTTAAAGCAAGAAGAAAAACAAAGACTGTACTATGGAAAAATGAAATAGATTTTAAACATGTTAATTCACGTT/
 ACTTCTACCTGTCAATGTTATTAATATTTTAGGAACAATAAATCACATTAATTCCTTATCTCATGTGAAATTTTCATATTATGATTGATACCTTTAAATGTCATTTGTTG/
 AAATATTTTGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGATACCATGTTCAACAAGACAGACTATGATTACAGGATCAGATGTGG
 GACACTAAACAAGTAAATAAAGTTAATTTCAAGTTGTAATTGATGCTAGAAAGACAATGAAACAGAGCCATGTGACCAATGAGAGAGATGAGGGTGGCAGCAGCCT
 GATTGGGAAGACCTCTCTGAGATTAGTGTCTTCAAGATATGCTTAAATGATATGAAAGAACCATTATGGAAGGCCTAGCATTAAAAACCGTCTAGGCAGAATGAGC/
 TGAGCTGGATATACTCAAGGAAGAAAGAGAAACTATGAAAAATGAAATAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTCTTCACT
 AGGAACAATAAATCACATTAATTCCTTATCTCATGTGAAATTTTCATATTATGATTGATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAA/
 CAAGA/
 AAGTT/
 CATTAA/
 GTTCTA/
 TATCCC/
 TTCGTT/
 TTGTAT/
 ACAAC/
 ATTCCTTATCTCATGTGAAATTTTCATATTATGATTGATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTGAATAATAAGTCA
 IACAGA
 TCAATG
 AATAAT/
 GTAAAT
 TGCAA/
 AGAGG/
 GAGAA
 GTTATT

- How does DNA sequence encode gene regulation?
- What regulatory logic is embedded in the sequence?

How DNA Sequences Drive Transcription?

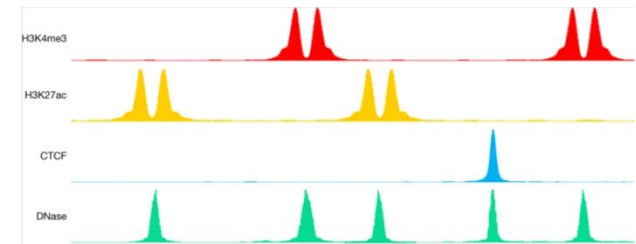
Input

```
ACAATAAATCACATTAATTCTTATCTCATGTGAAATTCATATTTATGATTG
ATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAAT
AAATATTTTTTAGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCT
AGGCATTGGGGATACCATGTTCCACAAGACAGACTATGATTACAGGATC
AGATGTGGACTCTCAAATTCGACTGAGAATAAACAGACACTAAACAAG
TAAATAAAGTTAATTTCAAGTTGTAATTGATGCTAGAAAAGACAATGAAACA
GAGCCATGTGACCAATGAGAGAGATGAGGGTGGCAGCAGCCTGTTTTA
GATAAGGTACCTGATTGGTGGGATTGGAAGACCTCTCTGAGATTAGTGT
CTTCAGATATGCCTTAATGATATGAAAGAACCATTTCATGGGAAGGCCTAG
CATTA AAAACCGCTCTAGGCAGAATGAGCAGCAAGTGCAAGGGTCCTGG
ATAGGAATGAGCTGGATATACTCAAGGAAGAAAGAGAACTATGGAAAA
ATGAAAATAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTA
CTTTTCTTTCTTCTACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACA
ATAAATCACATTAATTCCTTATCTCATGTGAAATTCATATTTATGATTGATA
CCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAA
TATTTTTTAGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGG
CATTGGGGATACCATGTTCCACAAGACAGACTATGATTACAGGATCAGAT
GTGGACTCTCAAATTCGACTGAGAATAAACAGACACTAAACAAGTAAAT
AAAGTTAATTTCAAGTTGTAATTGATGCTACTATGGAAAAATGAAAATAGA
TTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTTT
CACTTCTTACCTGTCAATGTTATTAATTTTTTAGGAACAATAAATCACATT
AATTCCTTATCTCATGTGAAATTCATATTTATGATTGATACCTTTAAATGT
CATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTTTAGA
ATAAATAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGAT
ACCATGTTCCACAAGACAGACTATGATTACAGGATCAGATGTGGACTCTC
AAATTCGACTGAGAATAAACAGACACAACAAGTAAATAAAGTTAATTT
CAAGTTGTAATTGATGCTATCCAGGCACAAGACCA....
```



ML / DL

Output



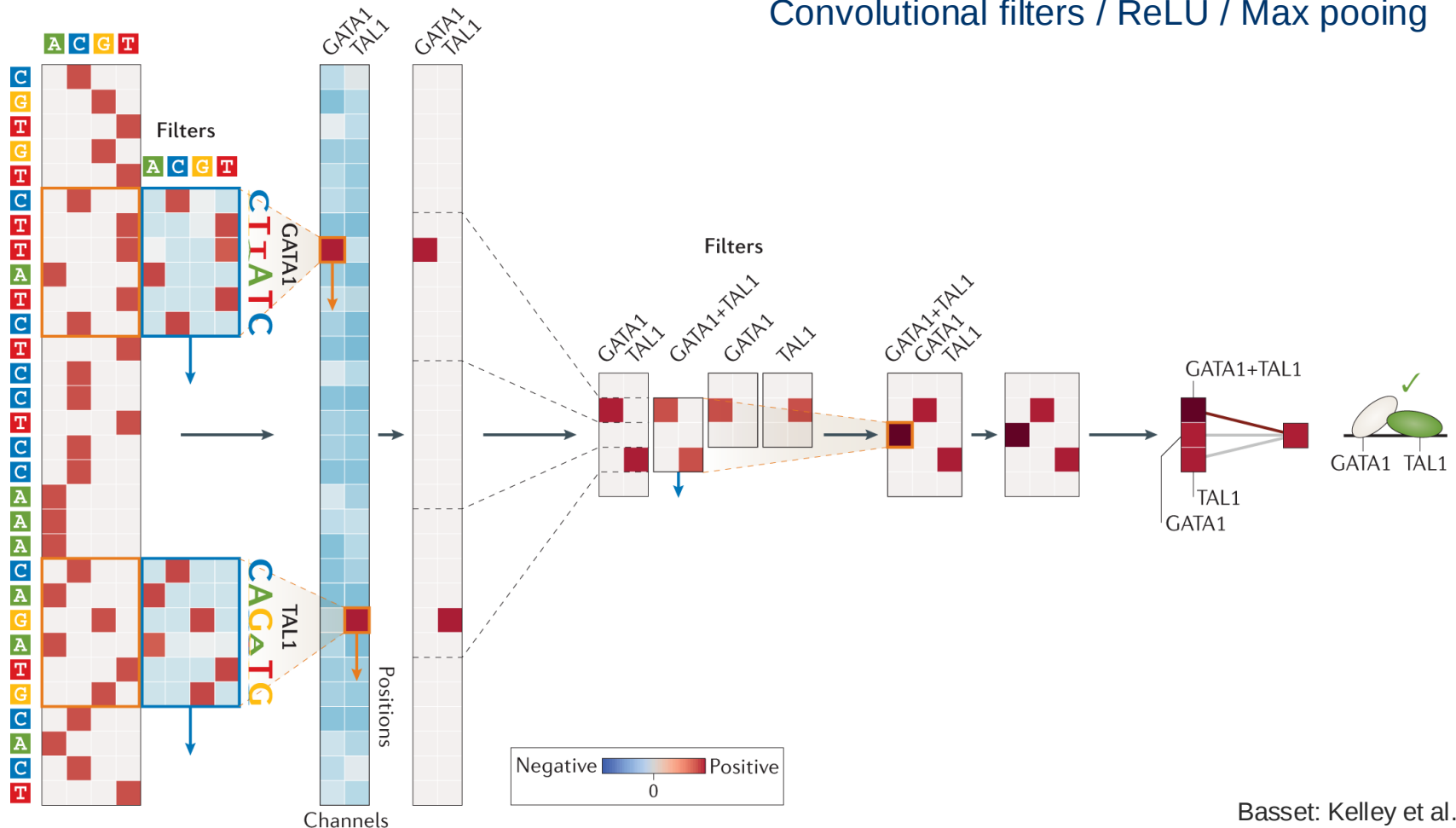
Functional assays:

1. ATAC-seq/DNase-seq
2. ChIP-seq, CUT&RUN
3. STARR-seq, MPRA
4. RNA-seq, CAGE, PRO-seq
5. Hi-C, ChIA-PET, micro-C

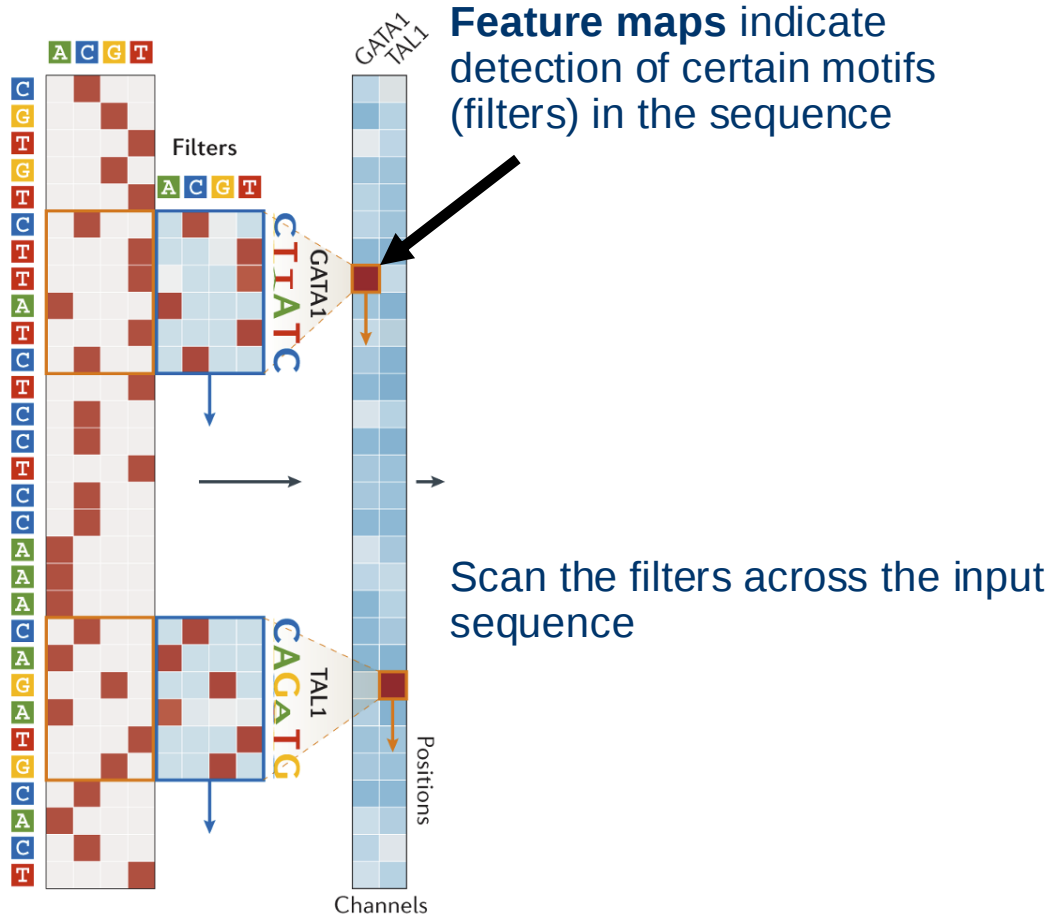
➤ Learn meaningful representations from DNA sequences (Basset).

Basset: Representation Learning of DNA Sequences Using CNNs

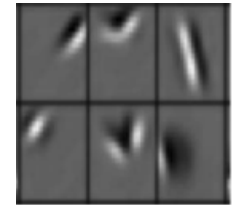
Convolutional filters / ReLU / Max pooling



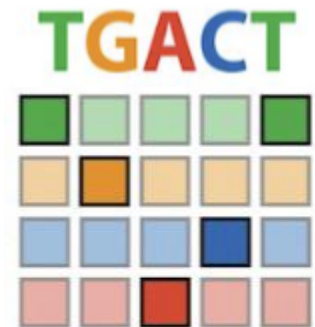
DNA convolutional filters



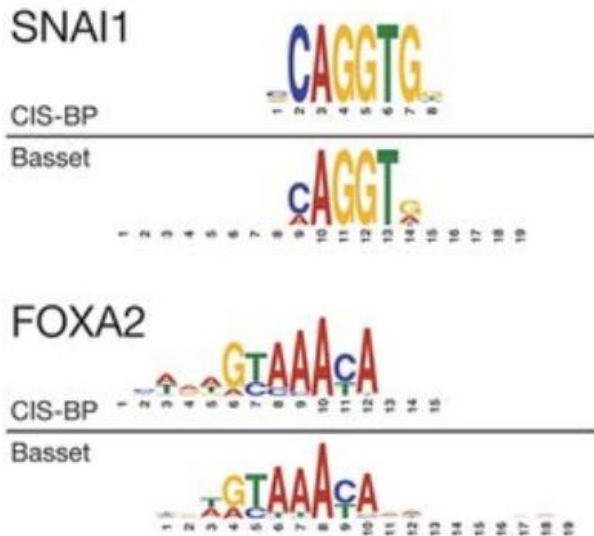
In image:



In DNA: PWM



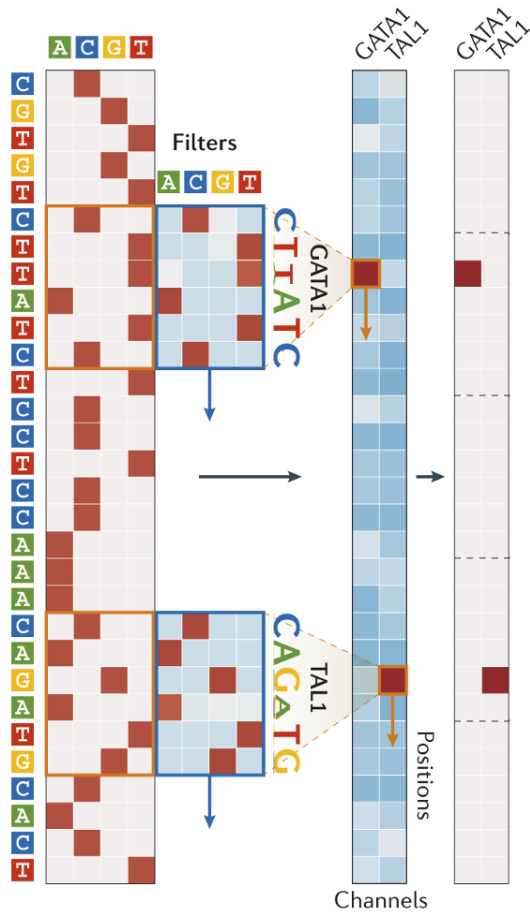
First-Layer Filters Align with Known TF Binding Motifs



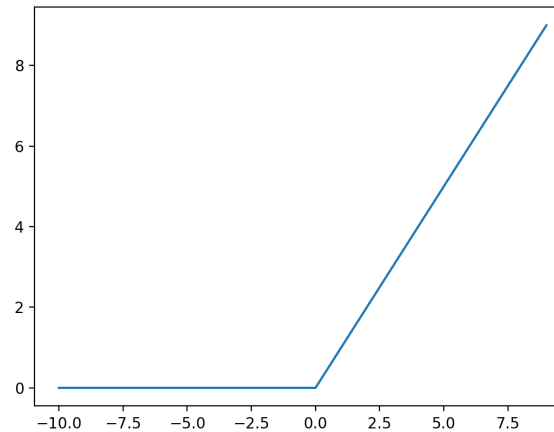
Motif information is derived by converting first-layer filter activations into probabilistic PWMs, counting nucleotide occurrences above a threshold.

Overall, 45% of filters could be annotated in the motif database

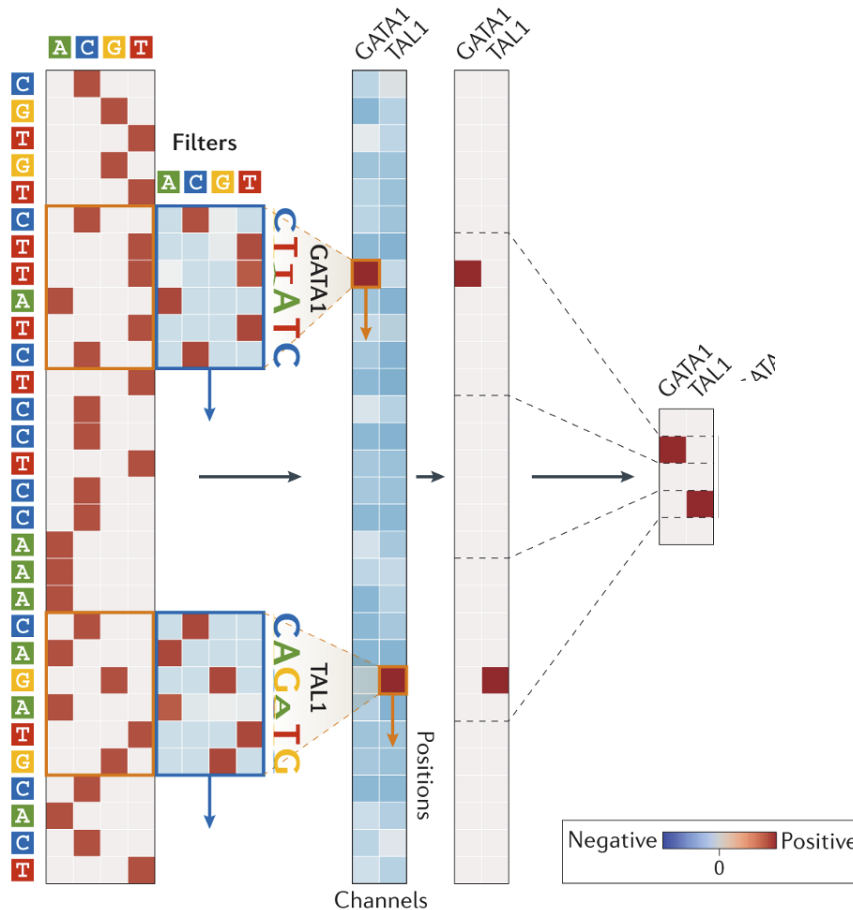
ReLU activation ignores misalignment



ReLU
 $\max(0, x)$



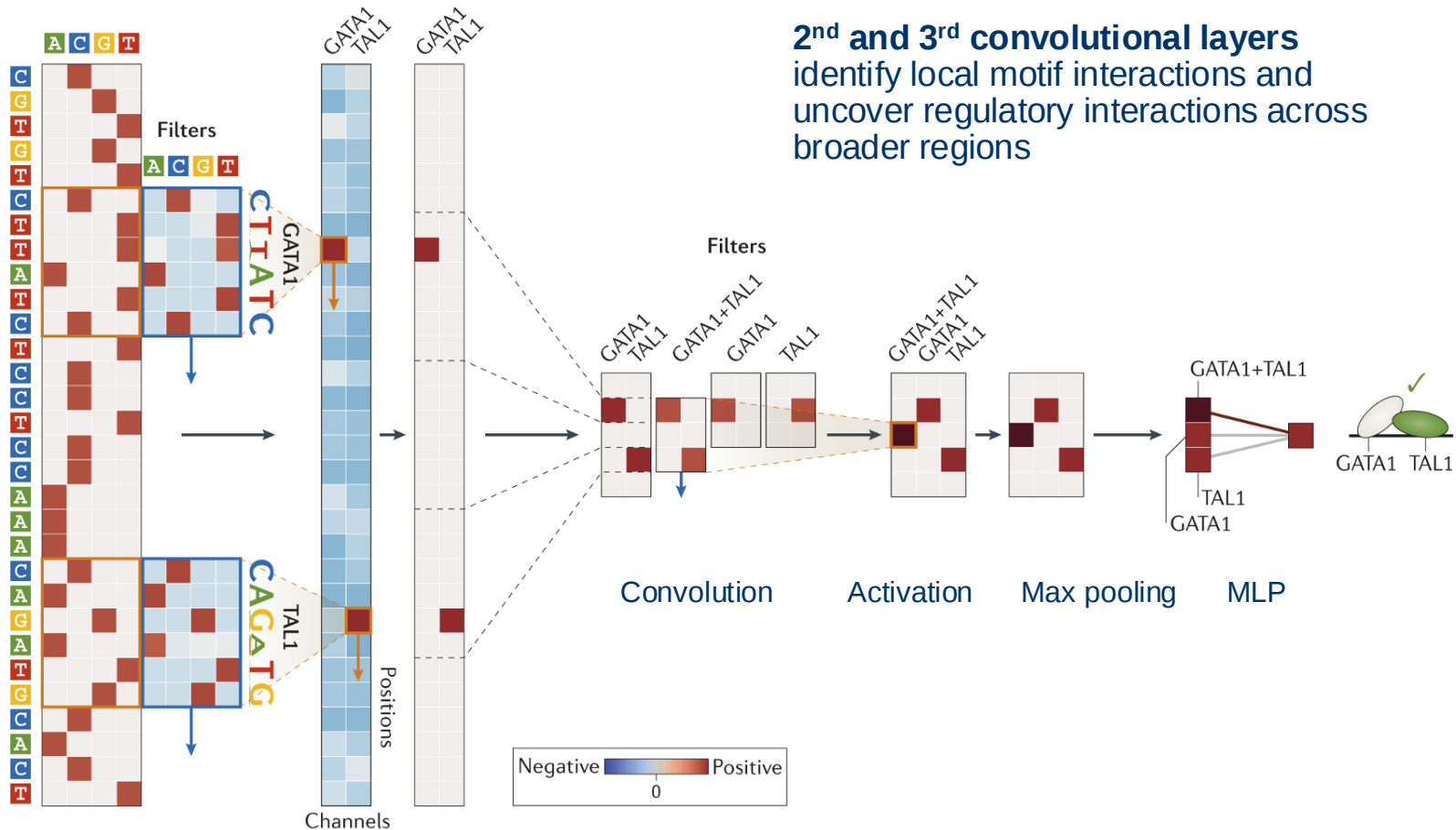
Max pooling summarizes features



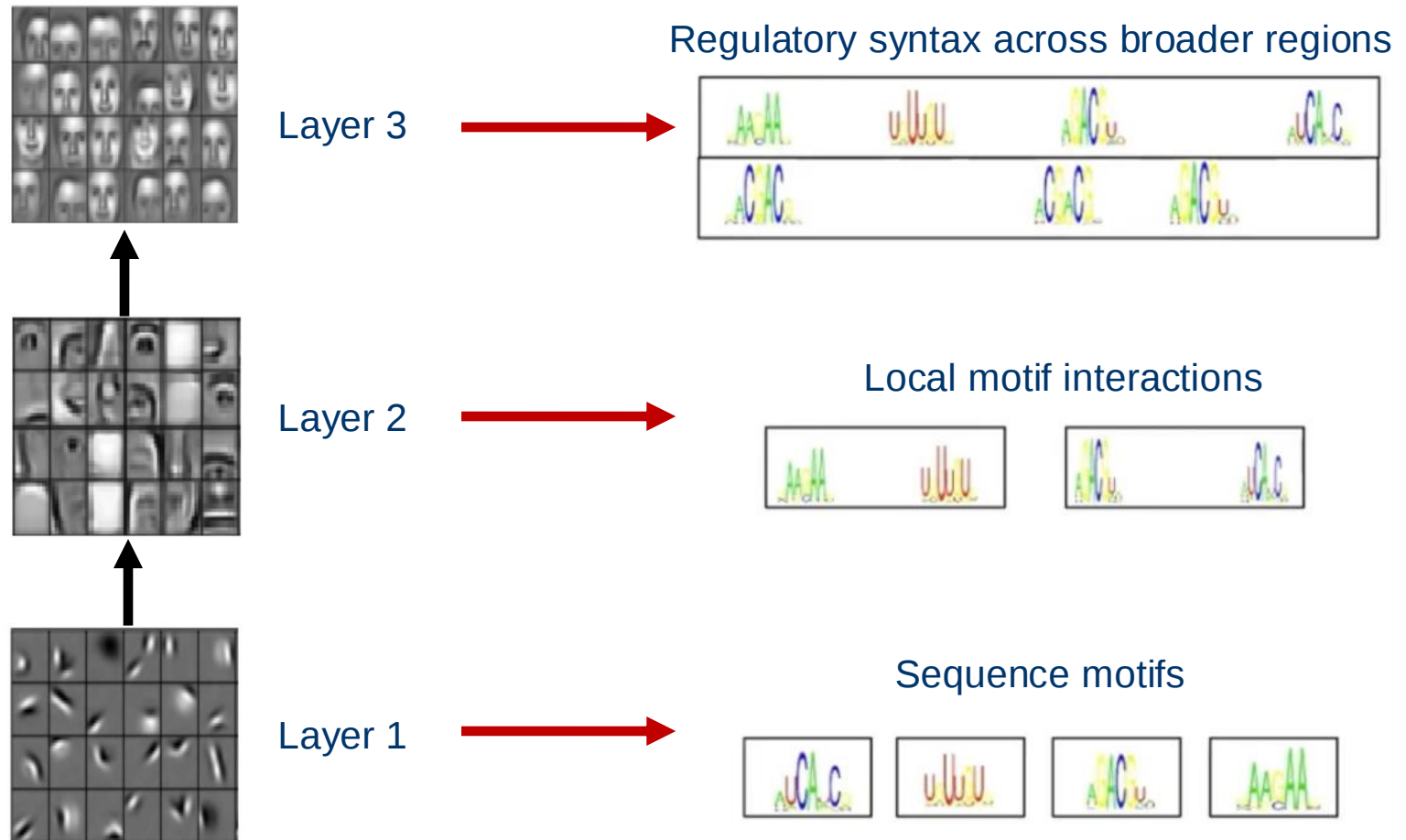
Max pooling

- Summarize / aggregate information
- Reduces the dimension of the input
- Provides invariance to small sequence shifts to the left or right.

Representation Learning of DNA Sequences Using CNNs



Hierarchical representation learning



Representation Learning of DNA Sequences Using CNNs

CNNs have inherent **limitations in capturing long-range interactions**

- Basset: 20k nt range; distal enhancers can act thousands of nts away

Sequence models (RNNs, LSTMs, and Transformers) extend the visible range of sequence interactions by modeling long-range dependencies.

✓ More on this in Martin's lecture.



How DNA Sequences Drive Transcription?

ACAATAAATCACATTAATTCTTATCTCATGTGAAATTTTCATATTTATGATTG
 ATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAAT
 AAATATTTTTTAGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCT
 AGGCATTGGGGATACCATGTTCCACAAGACAGACTATGATTACAGGATC
 AGATGTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAG
 TAAATAAAGTTAATTTCAAGTTGTAATTGATGCTAGAAAAGACAATGAAACA
 GAGCCATGTGACCAATGAGAGAGATGAGGGTGGCAGCAGCCTGTTTTA
 GATAAGGTACCTGATTGGTGGGATTGGAAGACCTCTCTGAGATTAGTGT
 CTTTCAGATATGCCTTAATGATATGAAAGAACCATTTCATGGGAAGGCCTAG
 CATTAAAAACCGTCTAGGCAGAATGAGCAGCAAGTGCAAGGGTCCTGG
 ATAGGAATGAGCTGGATATACTCAAGGAAGAAAGAGAACTATGGAAAA
 ATGAAAATAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTA
 CTTTTCTTTCTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACA
 ATAAATCACATTAATTCCTTATCTCATGTGAAATTTTCATATTTATGATTGATA
 CCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAA
 TATTTTTTAGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGG
 CATTGGGGATACCATGTTCCACAAGACAGACTATGATTACAGGATCAGAT
 GTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAGTAAAT
 AAAGTTAATTTCAAGTTGTAATTGATGCTACTATGGAAAAATGAAAAATAGA
 TTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTTCTTTCTTT
 CACTTCTTACCTGTCAATGTTATTAATTTTTTAGGAACAATAAATCACATT
 AATTCCTTATCTCATGTGAAATTTTCATATTTATGATTGATACCTTTAAATGT
 CATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTTTAGA
 ATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGAT
 ACCATGTTCCACAAGACAGACTATGATTACAGGATCAGATGTGGACTCTC
 AAATTCGACTGAGAATAAAACAGACACAACAAGTAAATAAAGTTAATTT
 CAAGTTGTAATTGATGCTATCCCAGGCACAAGACCA....

- Learn meaningful representations directly from DNA sequences (Basset)
- Predict and locate functional genomic regions (i.e. enhancers) (DECODE)
- Map DNA sequences to enhancer activity and uncover TF motif syntax rules (DeepSTARR)

STARR-seq allows direct measurement of enhancer activity

```

.....TAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTCAGGCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAGGTTGTTAAAGGGCAATCAGTCAAGGACAGGGGGA
GCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCCCTCAACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGCATCTCGGGATGAGGGCTAAGGCACAGTGCC
CATTAGTGCCAAAGACCAAGATCTGAGTCTTTGCAGAGCCTAAAAATTAAATCCTTTCTGTGCCCGGGTGGAGGGGGAGAGTTCAAGTCACTTTATAATATTTGCCCTCATATACATATCAC
CTGTGCCCTTCCACCTCCTTTTCTTGCAGTCTAATTATCCAAGTAGCATCTACAACCTCTTAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGCATAGGAGTCCCATCTTTC
TCGGAAAGAAATGACAATAAGGTTGTTAAAGGGCAATCAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCCCTCAACCACCAATTGCTCCCA
GGAATAGCACCTGGCTGGTCTGGCATCTCGGATGAGGGCTAAGGCACAGTGCCATTAGTCCCAAGAACCAAGATCTGAGTCTTTGCAGACCTAAAAATTAAATCCTTTCTGTGCCCGGG
GGTGGAGGGGGAGAGTTCAAGTCACTTTATAATATTTGCCCTCATATACATATCACCTGTGCCCTTCCACCTCCTTTTCTGTGCAGTCTAATTATCCAAGTAGCATCTACAACCTCTTAGCTACA
TCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAGGTTGTTAAAGGGCAATCAGTCAAGGACAGGGGGAGGCCAGTCCCTCTT
CCACCTCCTTTTCTTGCAGTCTAATTATCCAAGTAGCATAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAG
GTTGTTAAAGGGCAATCAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGAC
AGTGATGCAGCATTCCCTCAACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGCATCTCGGGATGAGGGCTAAGGCACAGTGCCATTAGTCCCAAGAACCAAGATCTGAGT
CTTTGCAGACCTAAAAATTAAATCCTTTCTGTGCCCGGGTGGAGGGGGAGAGTTCAAGTCACTTTATAATATTTGCTCATATACATATCACCTGTGCCCTTCCACCTCCTTTTCTTGC
GTCTAATTATCCAAGTAGCATCTACAACCTCTTAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAGGTTGTTAA
AGGGCAATCAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCCCTCAACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGCATC
TCGGGATGAGGGCTAAGGCACAGTGCCATTAGTGCCAAAGAACCAAGATCTGAGTCTTTGCAGACCTAAAAATTAAATCCTTTCTGTGCCCGGGTGGAGGGGGAGAGTTCAAGTCACT
TTATAATATTTGCCCTCATATACATATCACCTGTGCCCTTCCACCTCCTTTTCTTGCAGTCTAATTATCCAAGTAGCATAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGC
ATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAGGTTGTTAAAGGGCAATCAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCCCTC
AACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGCATCTCGGGATGAGGGCTAAGGCACAGTGCCATTAGTGCCAAAGAACCAAGATCTGAGTCTTTGCAGACCTAAAAATTAA
AATCCTTTCTGTGCCCGGGTGGAGGGGGAGAGTTCAAGTCACTTTATAATATTTGCCCTCATATACATATCACCTGTGCCCTTCCACCTCCTTTTCTTGCAGTCTAATTATCCAAGTAGCA
TCTACAACCTCTTAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAGGTTGTTAAAGGGCAATCAGTCAAGGACA
GGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCCCTCAACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGCATCTCGGGATGAGGGCTAAGGCACAGTG
ACAGTGCCATTAGTGCCAAAGGTCTGGCATCTCGGGATGAGGGCTAAGGCACAGTGCCATTAGTGCTCGGCATCTCGGGATGAGGGCTAAGGCACAGTG
TAAAGGGCAATCAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCCCTCAACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGC
ATCTCGGGATGAGGGCTAAGGCACAGTGCCATTAGTGCCAAAGAACCAAGATCTGAGTCTTTGTCAGACCTAAAAATTAAATCCTTTCTGTGCCCGGGTGGAGTTCAAGTCA
ACTTTATAATATTTGCCCTCATATACATATCACCTGTGCCCTTCCACCTCCTTTTCTTGCAGTCTAATTATCCAAGTAGCATAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAG
GCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAATAAGGTTGTTAAAGGGCAATCAGTCAAGGACAGGGGGAGGCCACTCAAAATCAGCACATGTGGCATGACAGTGATGCAGCATTCC
CTCAACCACCAATTGCTCCAGGAATAGCACCTGGCTGGTCTGGCATCTCGGGATGAGGGCTAAGGCACAGTGCCATTAGTGCCAAAGAACCAAGATCTGAGTCTTTGCAGAC
CTTAAATCCTTTCTGTGCCCGGGTGGAGGGGGAGAGTTCAAGTCACTTTATAATATTTGCCCTCATATACATATCACCTGTGCCCTTCCACCTCCTTTTCTTGCAGTCTAATTATCCAAGTA
GCATCTACAACCTCTTAGCTACATCAGTTGAAGATGGGGTTGCTGGGTAGCTGTGAGGCATAGGAGTCCCATCTTTCTCGGAAAGAATGACAACATTAGT.....

```

Localize enhancer regions

- **STARR-seq** provides direct readouts of enhancer activity
 - STARR-seq is massively parallel reporter assay that measures enhancer activity from arbitrary sources of DNA.
 - Genomic fragments are transfected into target cells in front of a luciferase gene, and enhancer activity is quantified by measuring luciferase expression.
- **Limitation:** STARR-seq lacks the resolution needed for precise enhancer localization.

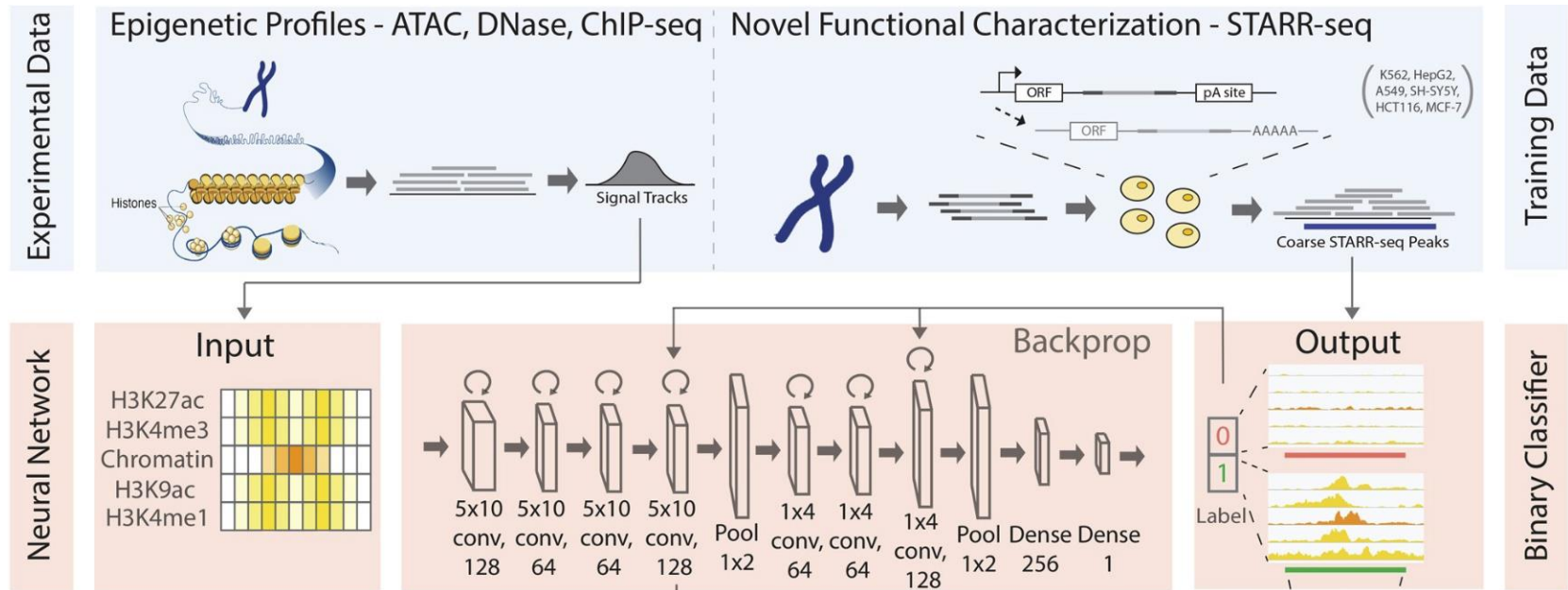
Epigenetic Features as Inputs → STARR-seq as Output

Labels

Cell Type	STARR-	ATAC-	DNase-	H3K27ac	H3K4me3	H3K4me1	H3K9ac
	seq	seq	seq	ChIP-seq	ChIP-seq	ChIP-seq	ChIP-seq
K562	✓	✓	✓	✓	✓	✓	✓
HepG2	✓	✓	✓	✓	✓	✓	✓
A549	✓		✓	✓	✓	✓	✓
HCT116	✓		✓	✓	✓	✓	✓
MCF-7	✓		✓	✓	✓	✓	✓

Training datasets. Histone marks and ATAC-seq across cell types.

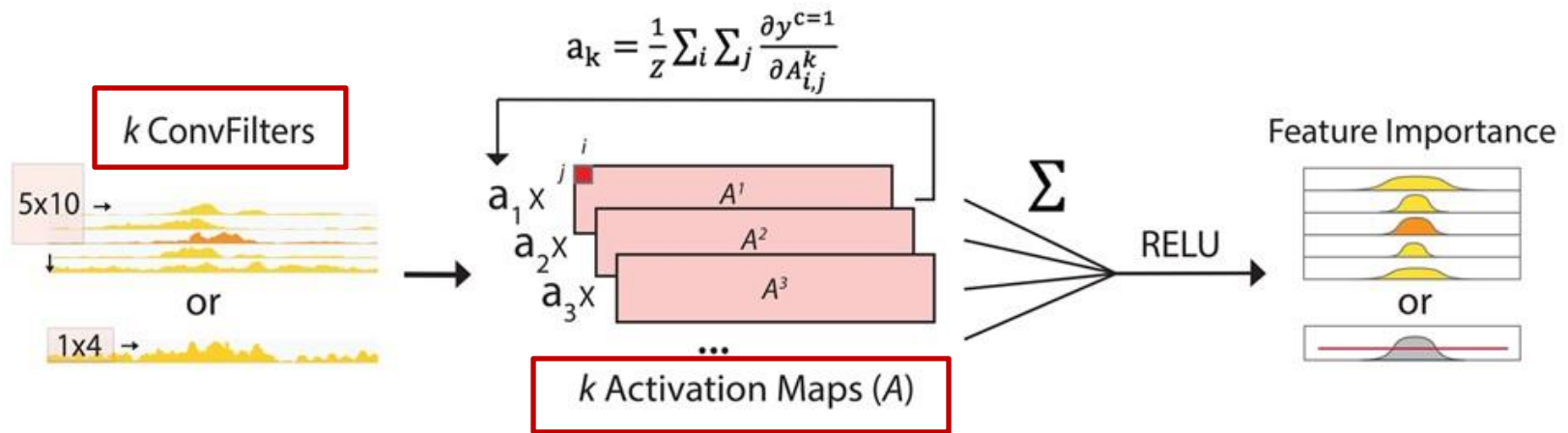
Predicting and Localizing Regulatory Regions with CNNs



Overall Workflow

- **Input:** High-resolution epigenetic features across a **4kb** genomic window.
- **Model:** A CNN-based binary classifier predicts the presence of enhancers.
- **Target:** STARR-seq peaks, representing enhancer activity regions.
- **Grad-CAM** is further applied to generate a **feature importance score**, justifying key features and precisely localizing enhancer regions.

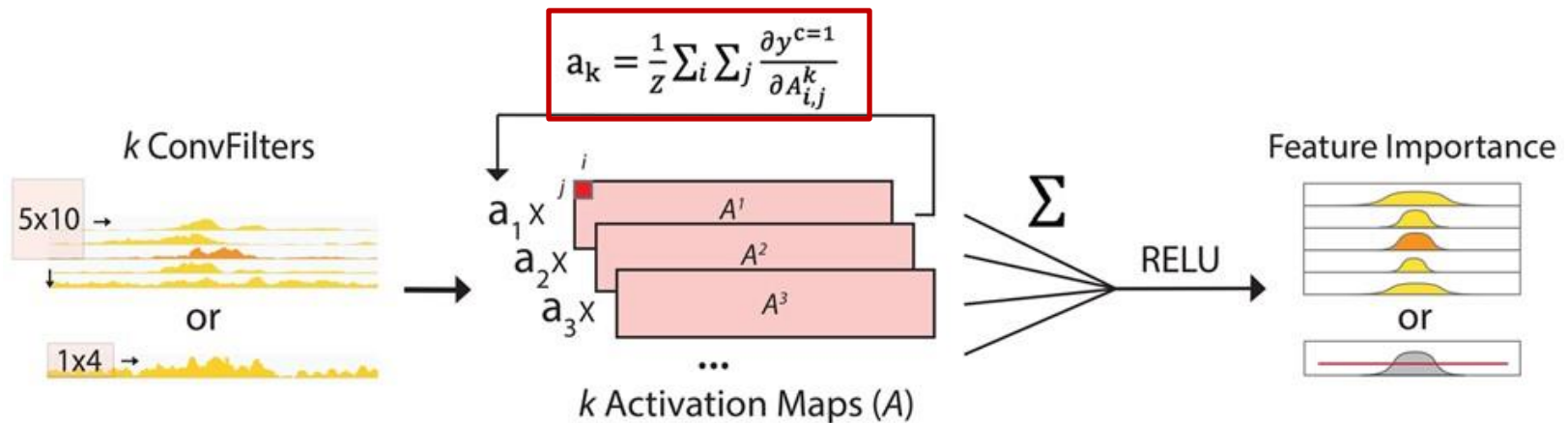
Leverage Feature Maps to Localize Enhancer Regions



Gradient-weighted Class Activation Mapping (Grad-CAM)

- Problem: Enhancer localization?
- **Convolutional Filters:** Learn features critical for enhancer prediction, outputting activation maps.
- **Feature Maps:** Highlight genomic regions containing these essential features

Leverage Feature Maps to Localize Enhancer Regions

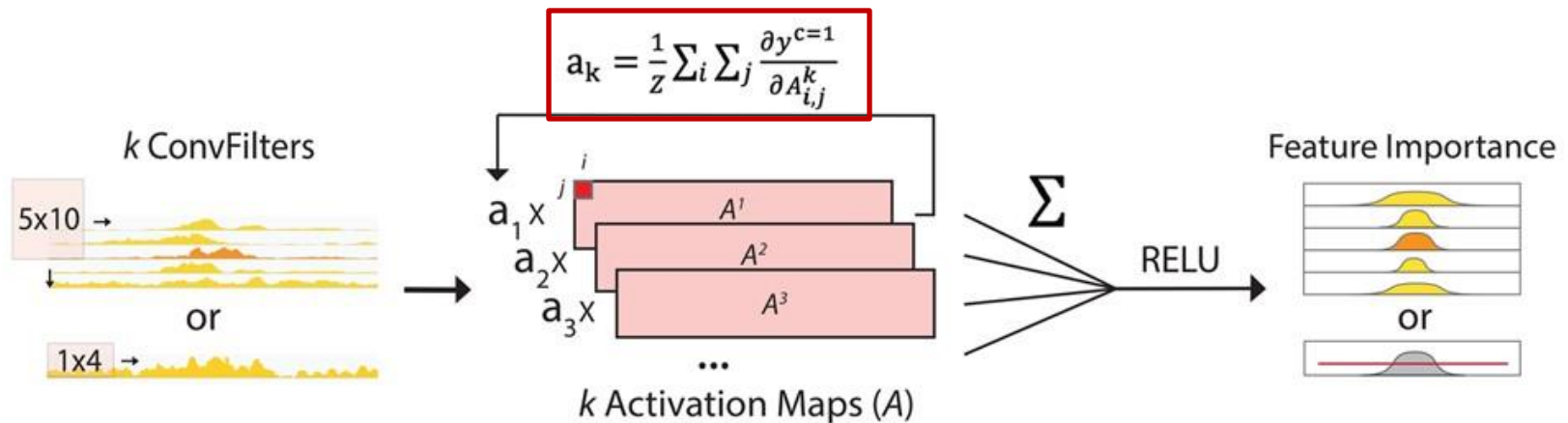


Gradient-weighted Class Activation Mapping (Grad-CAM)

$$\alpha_k^c = \overbrace{\frac{1}{Z} \sum_i \sum_j}^{\text{global average pooling}} \underbrace{\frac{\partial y^c}{\partial A_{i,j}^k}}_{\text{gradients via backprop}}$$

- Problem: Enhancer localization?
- By **superimposing activation maps (A^k) weighted by gradient-based importance scores (a_k)**, Grad-CAM highlights the most salient features for making predictions and reveals high-resolution core enhancer annotations.

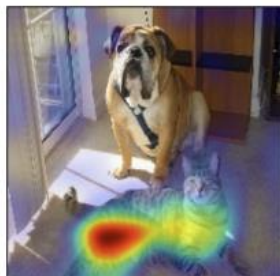
Leverage Feature Maps to Localize Enhancer Regions



Gradient-weighted Class Activation Mapping (Grad-CAM)



(a) Original Image



(c) Grad-CAM 'Cat'



(i) Grad-CAM 'Dog'

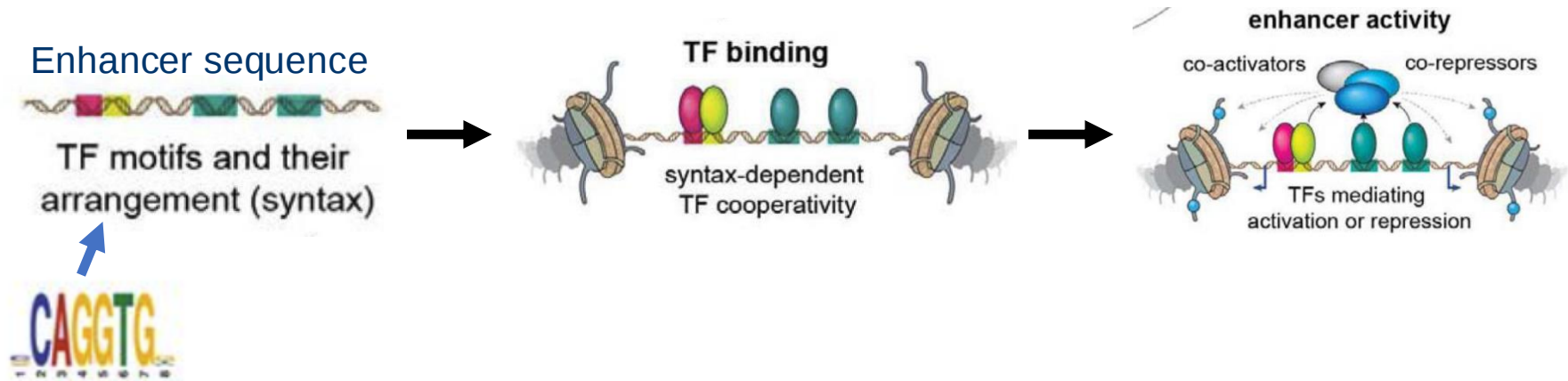
- Problem: Enhancer localization?
- Grad-CAM enables **precise enhancer localization** at a much higher resolution compared to the original 4 kb input.

How DNA Sequences Drive Transcription?

```
ACAATAAATCACATTAATTCTTATCTCATGTGAAATTTATATTTATGATTG
ATACCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAAT
AAATATTTTTTAGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCT
AGGCATTGGGGATACCATGTTCCACAAGACAGACTATGATTACAGGATC
AGATGTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAG
TAAATAAAGTTAATTTCAAGTTGTAATTGATGCTAGAAAAGACAATGAAACA
GAGCCATGTGACCAATGAGAGAGATGAGGGTGGCAGCAGCCTGTTTTA
GATAAGGTACCTGATTGGTGGGATTGGAAGACCTCTGAGATTAGTGT
CTTCAGATATGCCTTAATGATATGAAAGAACCATTTCATGGGAAGGCCTAG
CATTA AAAACCGTCTAGGCAGAATGAGCAGCAAGTGCAAGGGTCTCTGG
ATAGGAATGAGCTGGATATACTCAAGGAAGAAAGAGAACTATGGAAAA
ATGAAAATAGATTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTA
CTTTCTTTCTTTCACTTCTTACCTGTCAATGTTATTAATATTTTAGGAACA
ATAAATCACATTAATTCCTTATCTCATGTGAAATTTATATTTATGATTGATA
CCTTTAAATGTCATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAA
TATTTTTTAGAATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGG
CATTGGGGATACCATGTTCCACAAGACAGACTATGATTACAGGATCAGAT
GTGGACTCTCAAATTCGACTGAGAATAAAACAGACACTAAACAAGTAAAT
AAAGTTAATTTCAAGTTGTAATTGATGCTACTATGGAAAAATGAAAATAGA
TTTTAAACATGTTAATTCACGTTACTTTTTGTTAAATTTACTTTCTCTTT
CACTTCTTACCTGTCAATGTTATTAATTTTTAGGAACAATAAATCACATT
AATTCCTTATCTCATGTGAAATTTATATTTATGATTGATACCTTTAAATGT
CATTTGTTGAAGGAAGATTATTCATTTTTTCATTCAATAAATATTTTTTAGA
ATAATAAGTCCCAGGCACAAGACCAGTATTATGTTCTAGGCATTGGGGAT
ACCATGTTCCACAAGACAGACTATGATTACAGGATCAGATGTGGACTCTC
AAATTCGACTGAGAATAAAACAGACACAACAAGTAAATAAAGTTAATTT
CAAGTTGTAATTGATGCTATCCAGGCACAAGACCA....
```

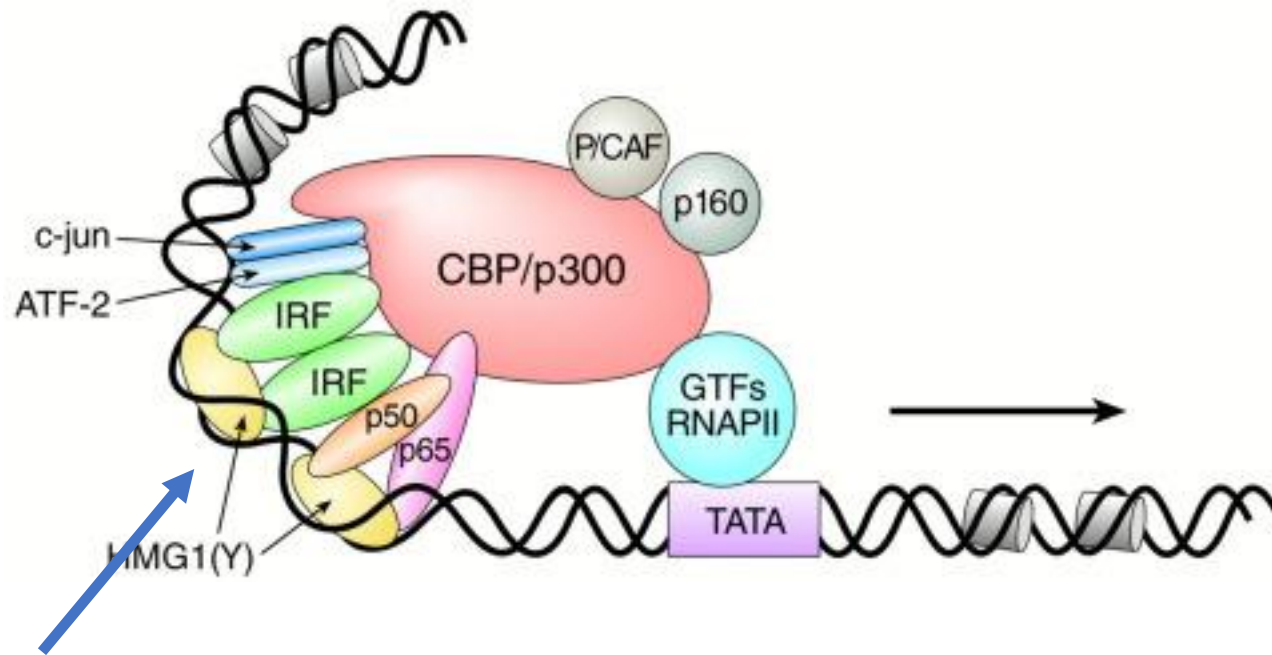
- Learn meaningful representations directly from DNA sequences (Basset)
- Predict and locate functional genomic regions (DECODE)
- Map DNA sequences to enhancer activity and uncover TF motif syntax rules (DeepSTARR)

The cis-regulatory code that regulates enhancer activity



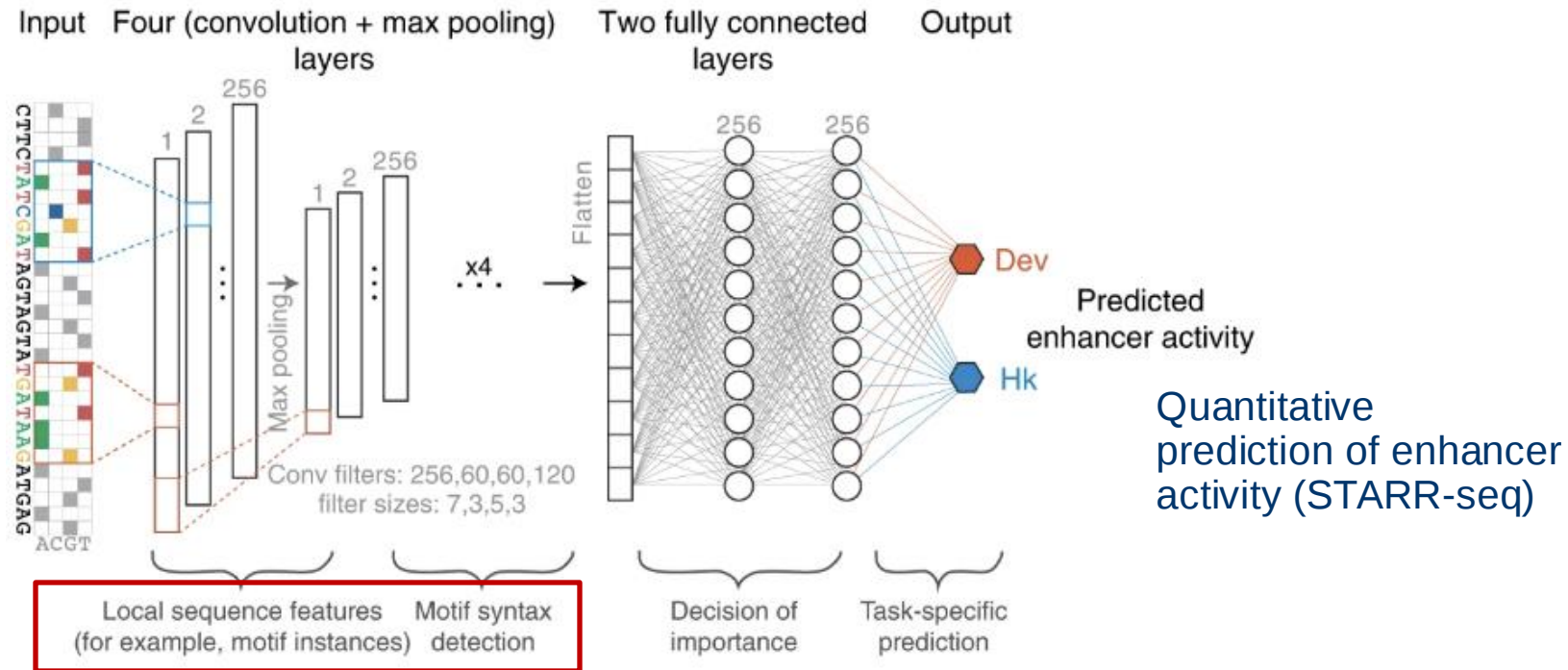
TF motif arrangements/syntax
(number, order, orientation, flanks, and spacing)

The spatial arrangement of TF binding is critical



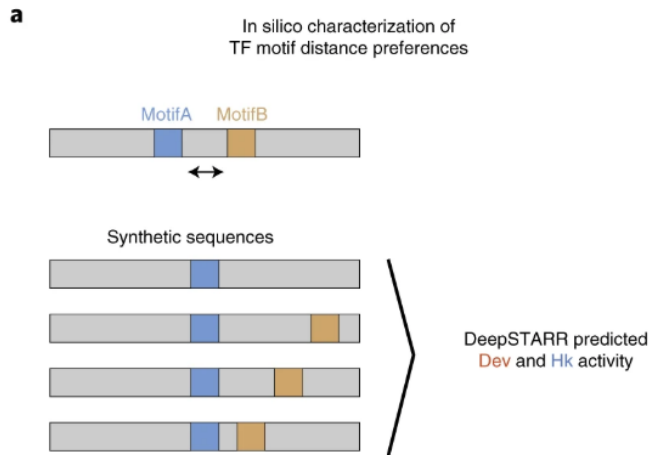
Enhancer sequence

Quantitatively predicting enhancer activity from DNA sequence



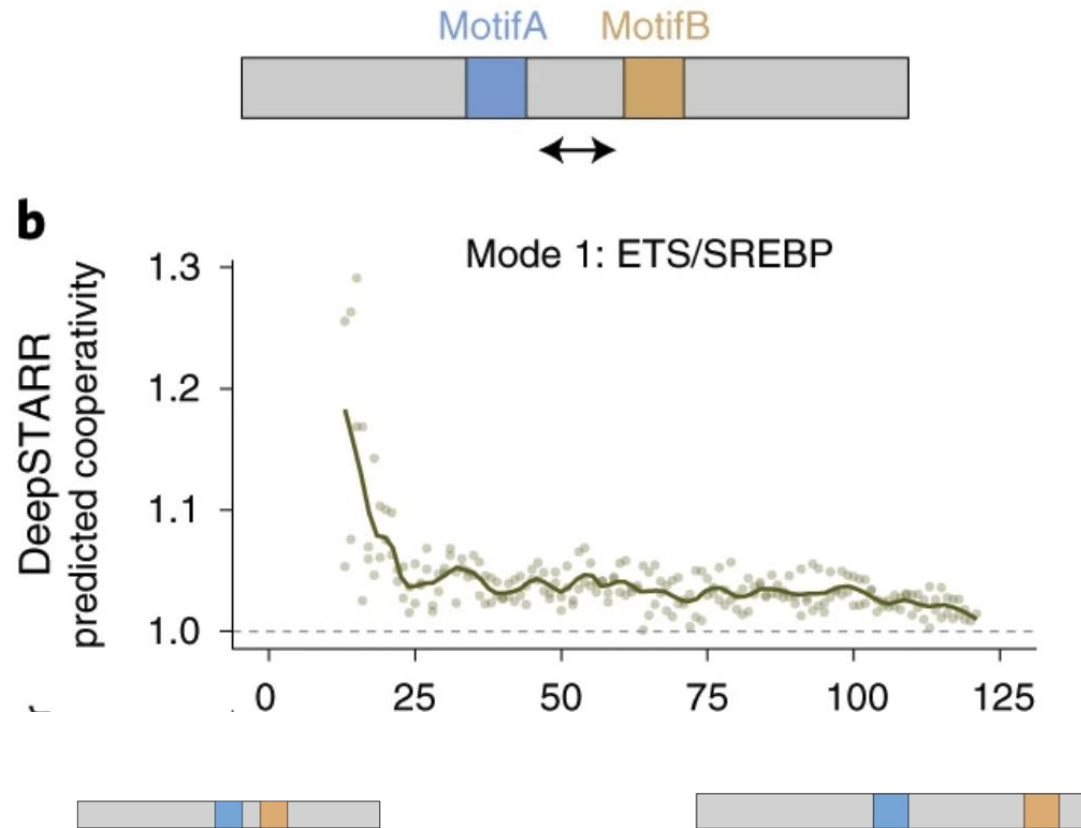
- **Input:** 1-hot encoding of 249-bp DNA sequences.
- **Architecture:** 4-layer CNN with max-pooling after each layer, followed by two fully connected layers.
- **Convolutional filters:**
 - Early convolutional layers identify TF motifs.
 - Later layers capture more complex patterns like motif syntax and local motif interactions.
- **Output:** Simultaneous prediction of quantitative enhancer activities for developmental (Dev) and housekeeping (Hk) enhancers.

In silico analysis of motif distances

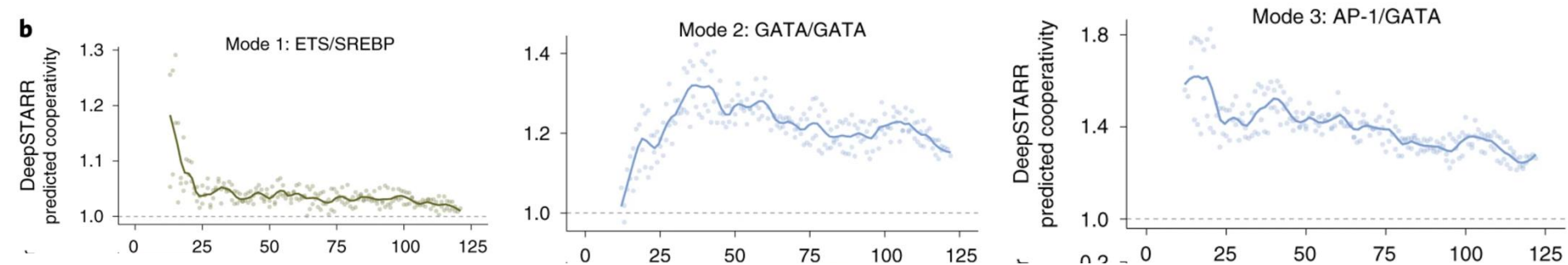


DeepSTARR

Quantitative enhancer
activity prediction



In silico analysis reveals distinct modes of motif cooperativity



Cooperativity mode (1/2/3) depends on the TF and motif pair

- **ETS**: always mode 1
- **AP-1**: always mode 3
- **GATA**: modes 1/2/3 depending on partner TF
- Also validated with experimental data

Summary

- Convolutional operations in image classification
 - Use **convolutional filters** to learn features in the image (as well as genome)
 - Introduce non-linearity through **ReLU activation**
 - **Pooling** is commonly use to summarize information and preserve spatial invariance
- Application of CNNs in regulatory genomics
 - Learn meaningful representations from DNA sequences (Basset)
 - Predict and precisely locate functional genomic regions (DECODE)
 - Map DNA sequences to enhancer activity and uncover TF motif syntax rules (DeepSTARR)

Suggested Readings

- **ISLR [v2](#)**: Chapter 10.3 Convolutional Neural Networks
- **Greener et al., 2022** – A guide to machine learning for biologists, Nat Rev Mol Cell Biol.
- **Kelley et al., 2016** – Basset: learning the regulatory code with CNNs, Genome Res.
- **Chen et al., 2021** – DECODE: deep-learning framework for enhancer analysis, Bioinformatics.
- **de Almeida et al., 2022** – DeepSTARR: predicting enhancer activity from DNA, Nat Genet.

(Optional) Suggested Readings

- **Kelley et al., 2016** – Basset: learning the regulatory code with CNNs, Genome Res.
- **Kelley et al., 2018** – Dilated CNNs for sequential regulatory activity prediction, Genome Res.
- **Avsec et al., 2021** – Enformer: gene expression prediction with long-range interactions, Nat Methods.
- **Linder et al., 2025** – Borzoi: predicting RNA-seq coverage from DNA sequence, Nat Genet.