Artificial Intelligence: *Modeling Human Intelligence with Networks*

Jeová Farias Sales Rocha Neto
jeova_farias@brown.edu

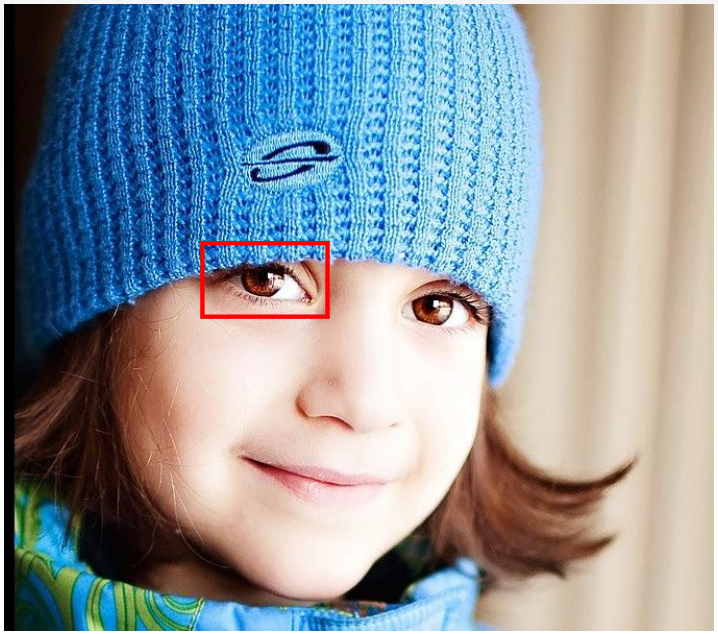
Convolutions and Max Pooling

Images

- What is an image?

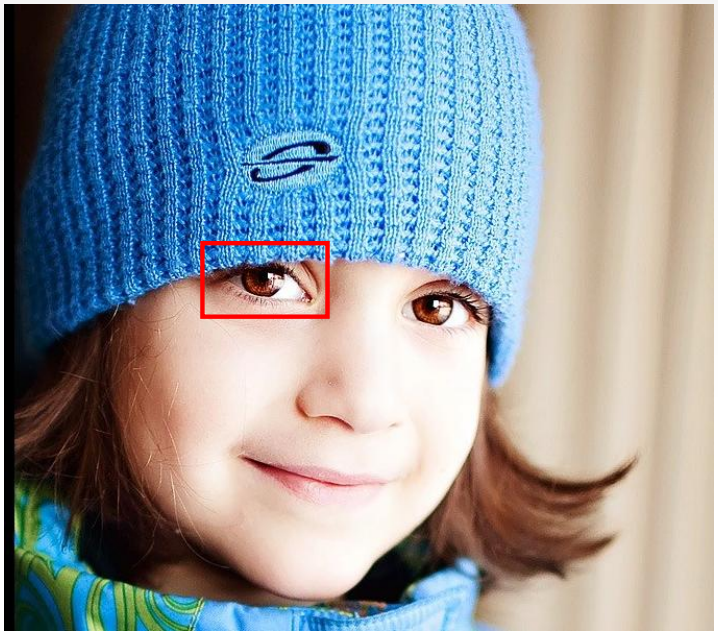
Images

- What is an image?



Images

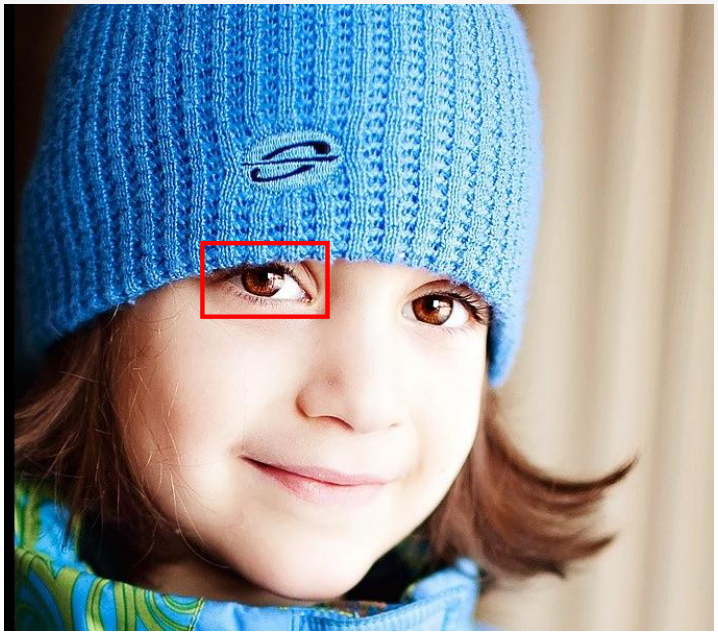
- What is an image?



- What is a pixel?

Images

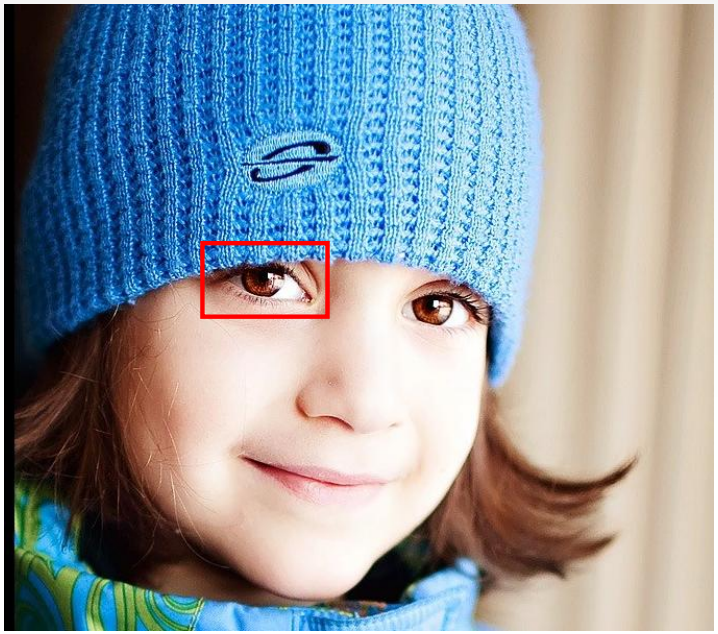
- What is an image?



- What is a pixel?
- A pixel is a set of three numbers between 0 and 255 that together represent a particular color.

Images

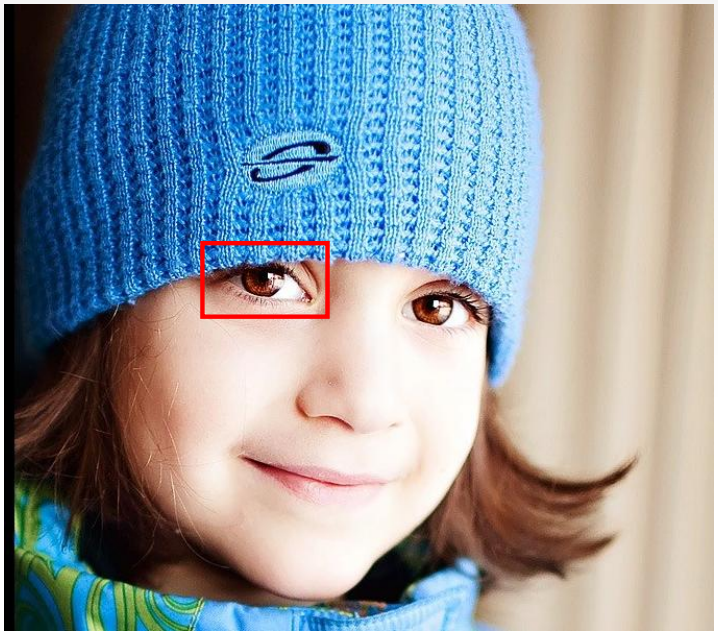
- What is an image?



- What is a pixel?
- A pixel is a set of three numbers between 0 and 255 that together represent a particular color.
- The values of the set are R (for red), G (for green) and B (for blue)

Images

- What is an image?



- What is a pixel?
- A pixel is a set of three numbers between 0 and 255 that together represent a particular color.
- The values of the set are R (for red), G (for green) and B (for blue)

R 255 G 255 B 102	R 255 G 0 B 204	R 51 G 204 B 255
R 51 G 51 B 0	R 51 G 51 B 153	R 255 G 153 B 153

Grayscale image

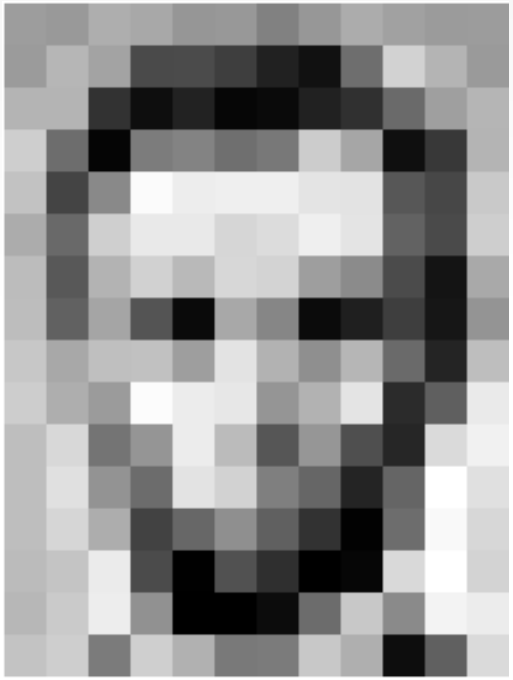
- Now, what is a grayscale image?

Grayscale image

- Now, what is a grayscale image?
- Instead of 3 numbers per pixel, just 1.

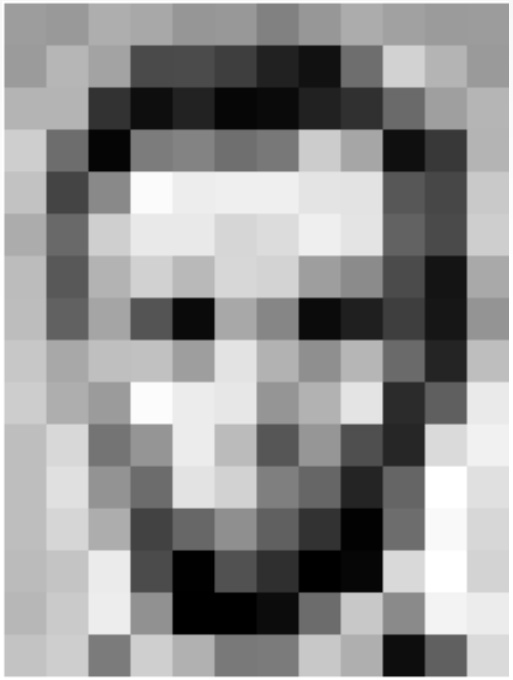
Grayscale image

- Now, what is a grayscale image?
- Instead of 3 numbers per pixel, just 1.



Grayscale image

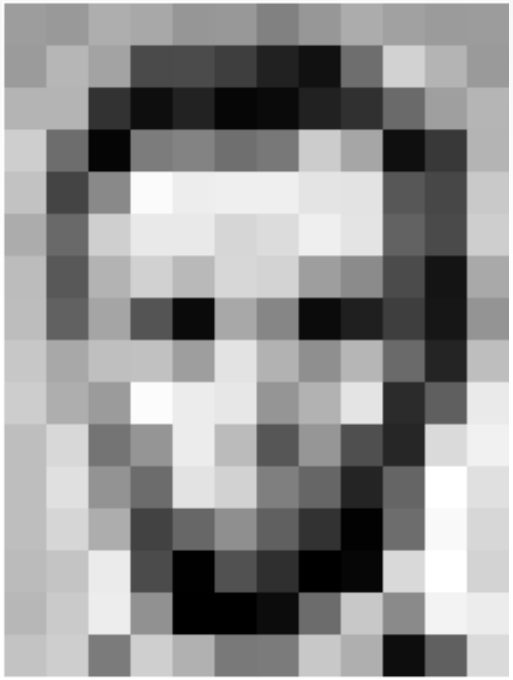
- Now, what is a grayscale image?
- Instead of 3 numbers per pixel, just 1.



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Grayscale image

- Now, what is a grayscale image?
- Instead of 3 numbers per pixel, just 1.



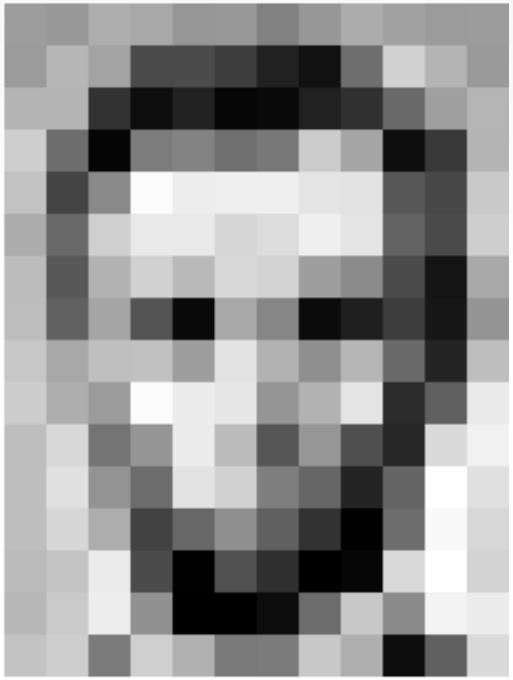
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

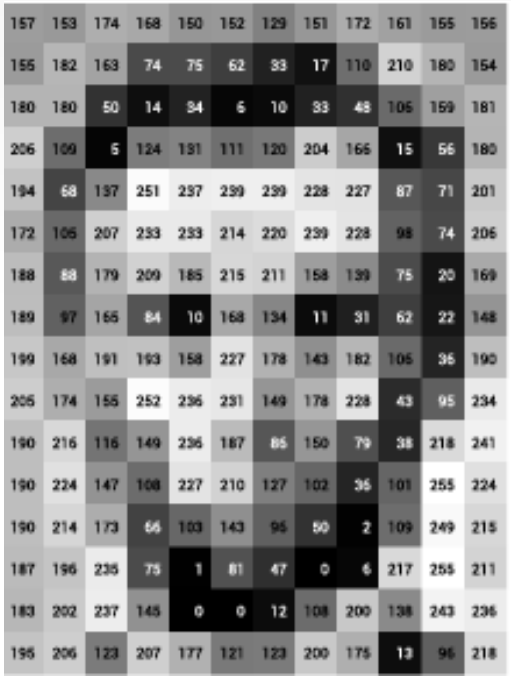
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

- What does the last figure represent, mathematically?

Grayscale image

- Now, what is a grayscale image?
- Instead of 3 numbers per pixel, just 1.





157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

- What does the last figure represent, mathematically?
- A matrix!

Convolutions

- Convolutions involve the concept of a sliding window.

Convolutions

- Convolutions involve the concept of a sliding window.
- It simply sweeps a window over the image

Convolutions

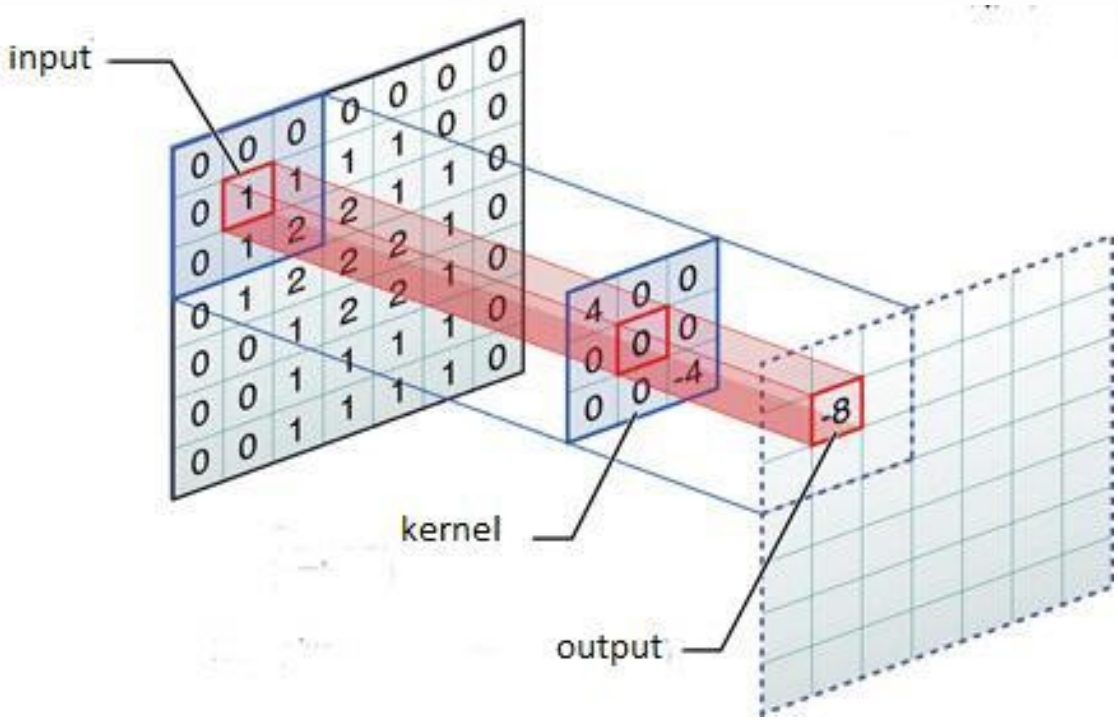
- Convolutions involve the concept of a sliding window.
- It simply sweeps a window over the image
- And compute the dot product between that window and a given matrix f.

Convolutions

- Convolutions involve the concept of a sliding window.
- It simply sweeps a window over the image
- And compute the dot product between that window and a given matrix f.
- We will call this matrix a filter (some people call it kernel).

Convolutions

- Convolutions involve the concept of a sliding window.
- It simply sweeps a window over the image
- And compute the dot product between that window and a given matrix f.
- We will call this matrix a filter (some people call it kernel).



Convolutions

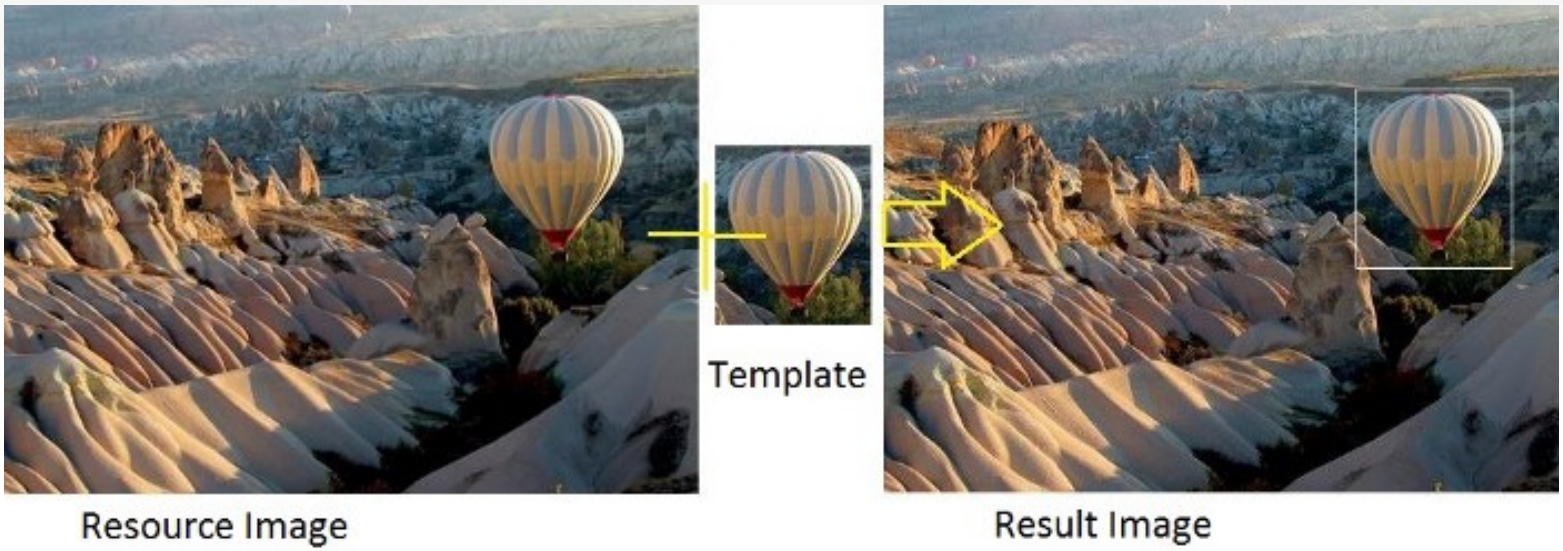
Algorithm 1 *Convolution*

Input: An image I and a filter $f \in \mathbb{R}^{n \times n}$, both matrices
Output: A matrix M.

1: Compute n_row and n_col as the number of rows and columns of I
2: Compute $d = (n-1)/2$
3: **for** i from d to $n\_row - d$ **do**
4: **for** j from d to $n\_col - d$ **do**
5: $M[i,j] = I[i-d : i+d+1, j-d : j+d+1] \cdot f$
6: **end for**
7: **end for**

Template Matching

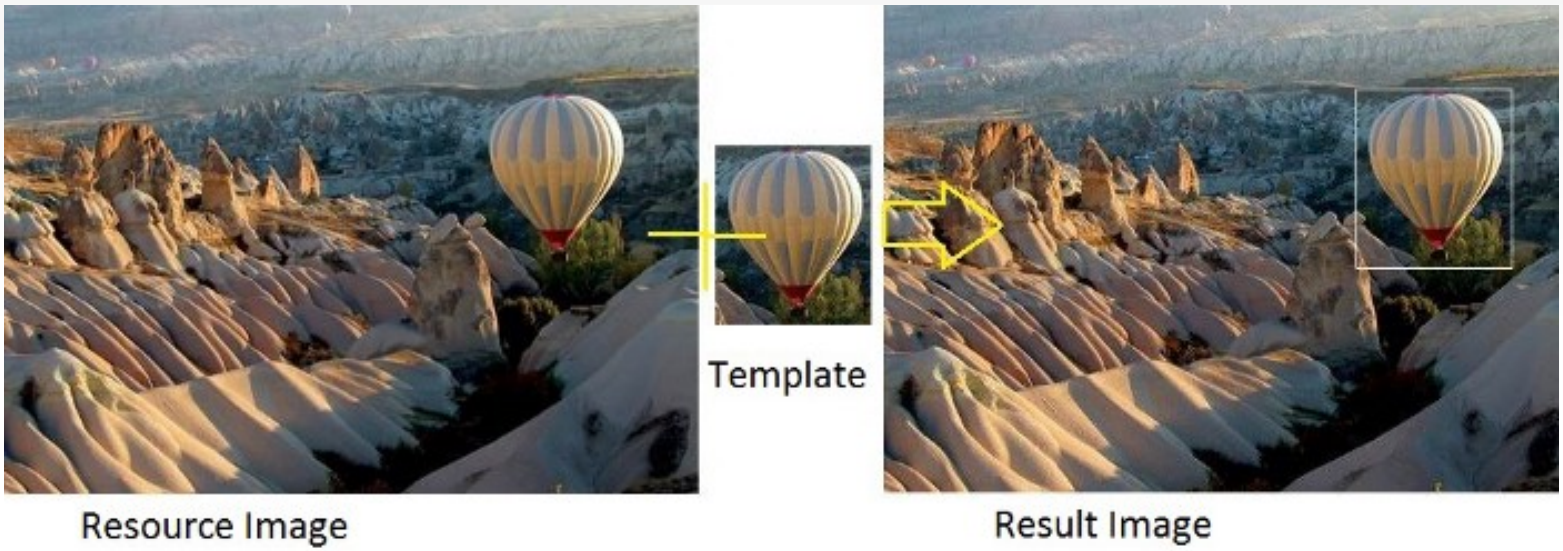
- Convolutions can also be used to do cool things with images, like template matching!



- How can we use the sliding window idea to do that?

Template Matching

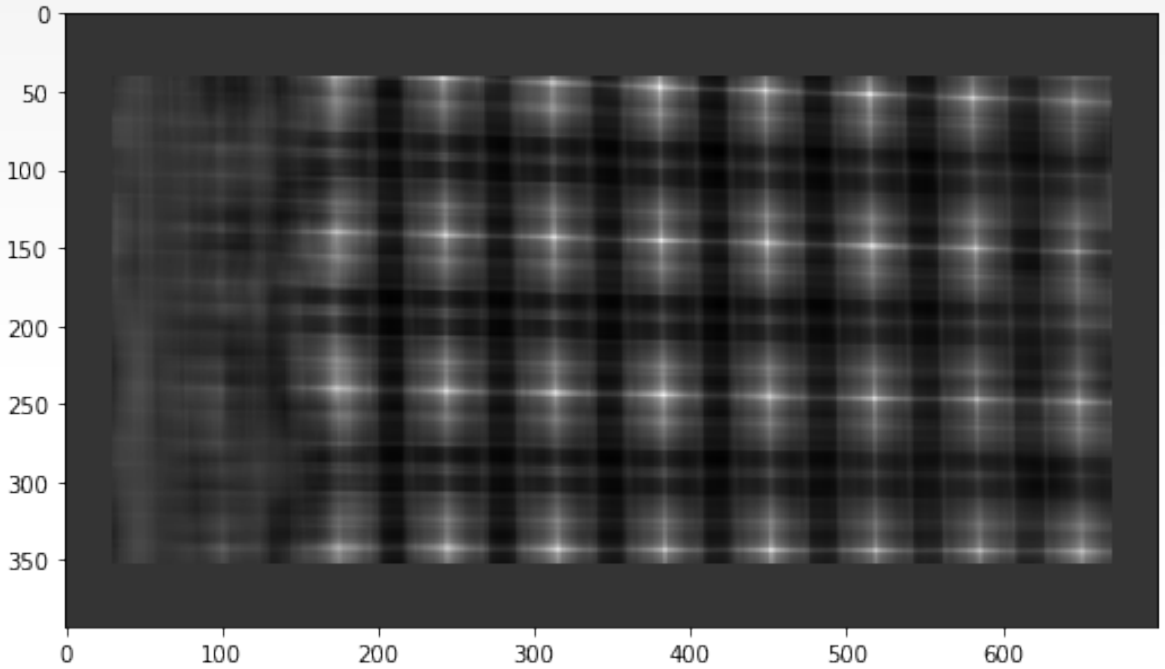
- Convolutions can also be used to do cool things with images, like template matching!



- How can we use the sliding window idea to do that?
- Let's try to do that in python!

Non-maxima Suppression

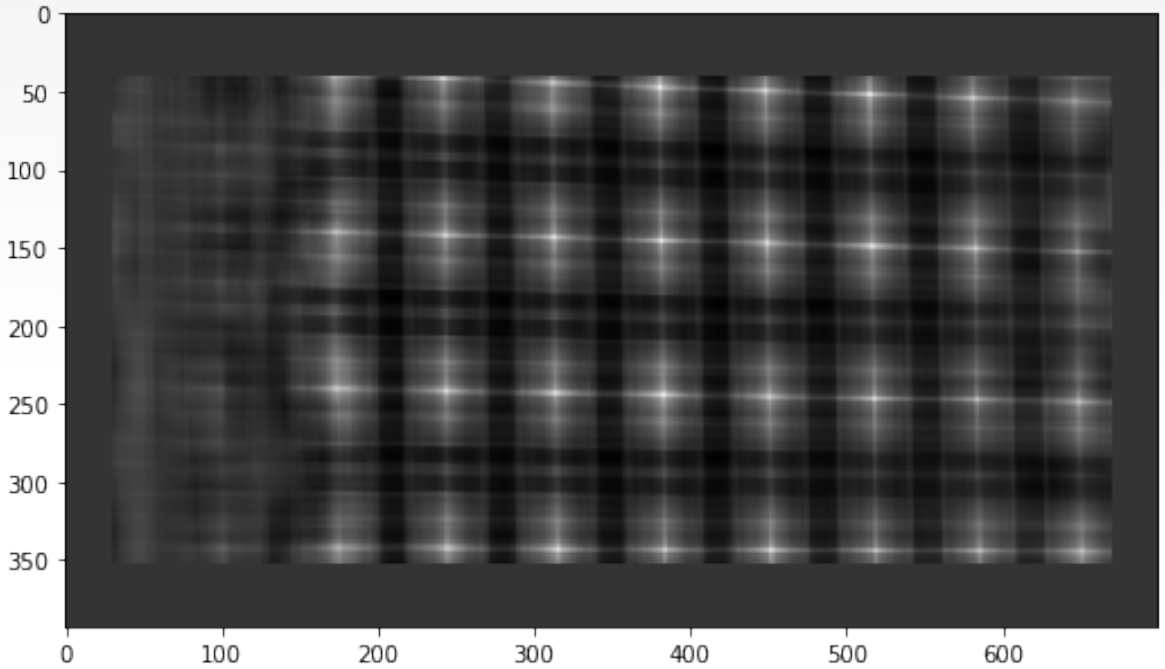
- This is probably what you got from the exercise:



- Where do you thing your templates are?

Non-maxima Suppression

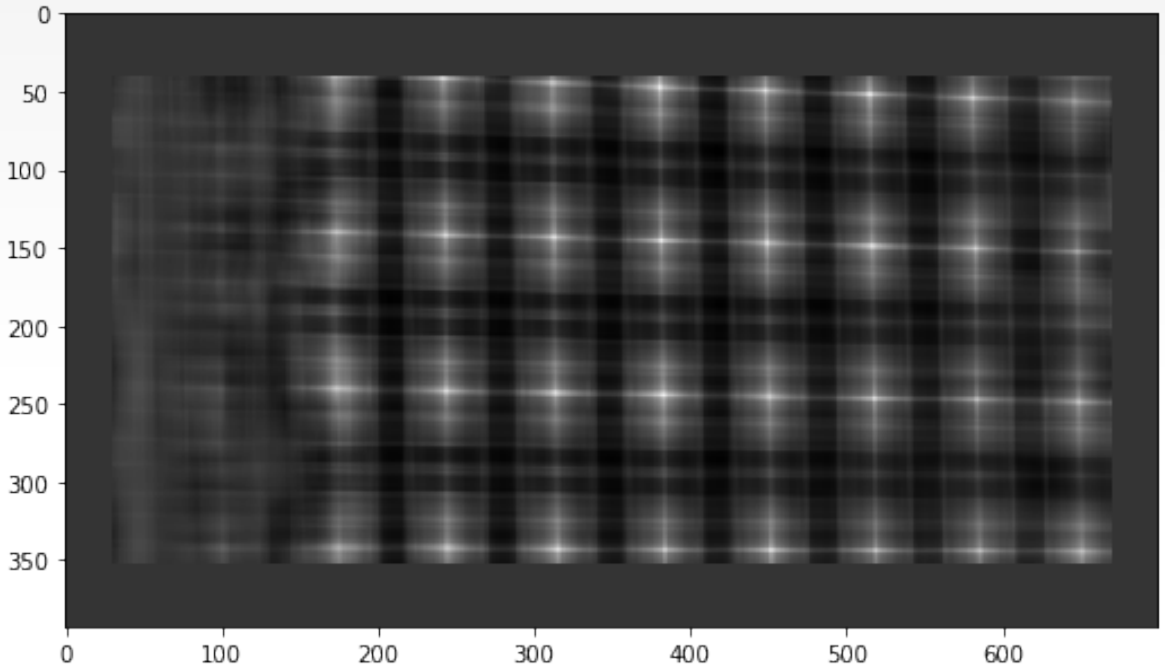
- This is probably what you got from the exercise:



- Where do you thing your templates are?
- What values do these spots have compared to the others?

Non-maxima Suppression

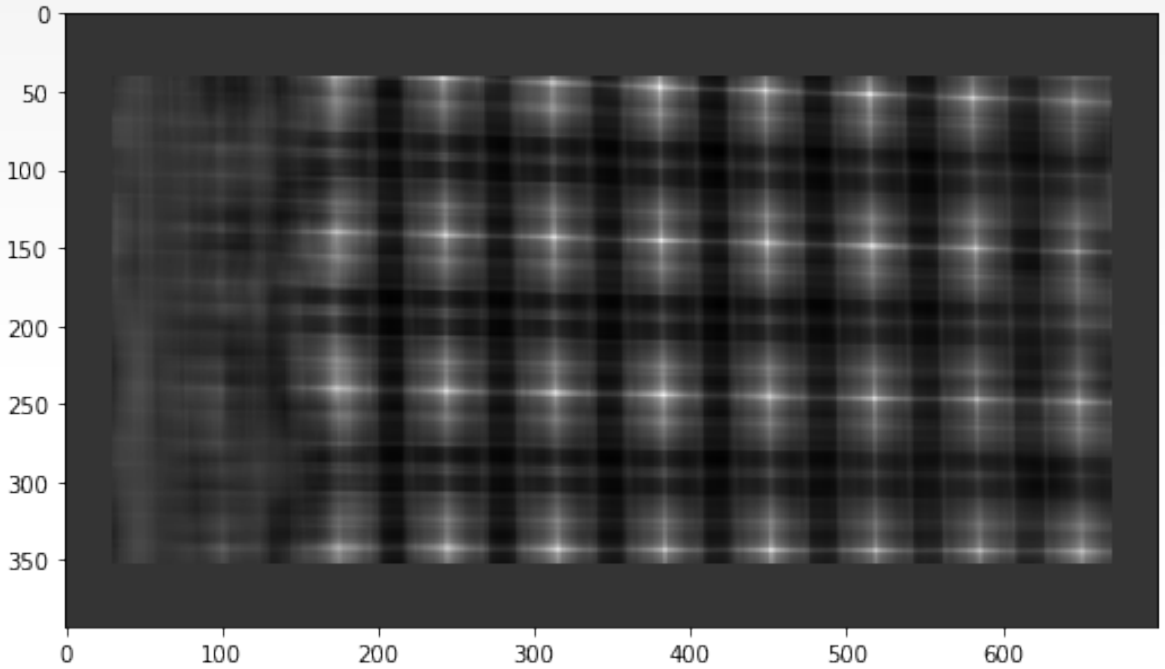
- This is probably what you got from the exercise:



- Where do you thing your templates are?
- What values do these spots have compared to the others?
- How can we select these spots?

Non-maxima Suppression

- This is probably what you got from the exercise:



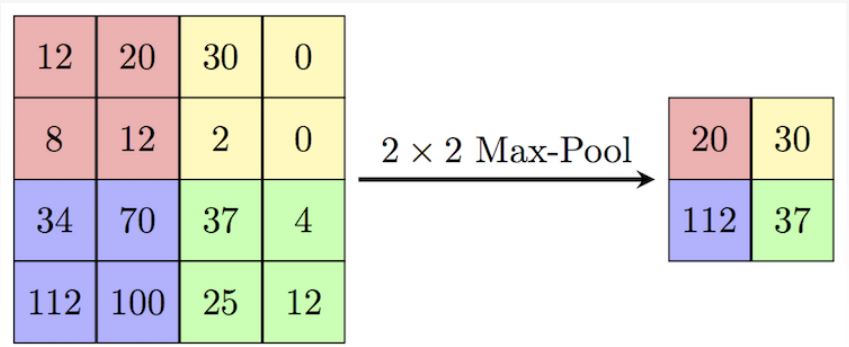
- Where do you thing your templates are?
- What values do these spots have compared to the others?
- How can we select these spots?
- What does “Non-maxima Suppression” mean?

Max pooling

- The last operation before Deep Learning!

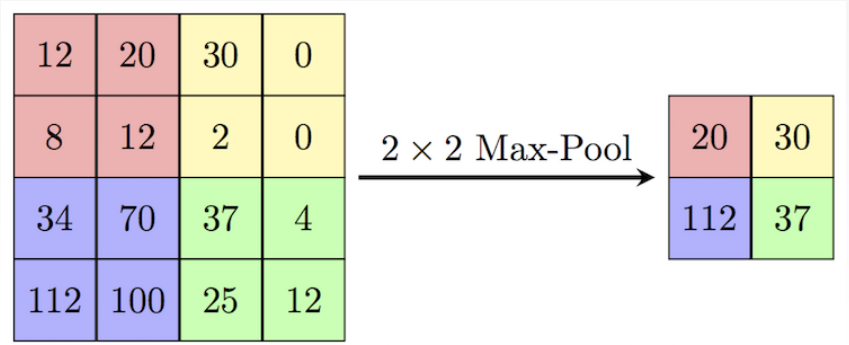
Max pooling

- The last operation before Deep Learning!
- Here's what it does:

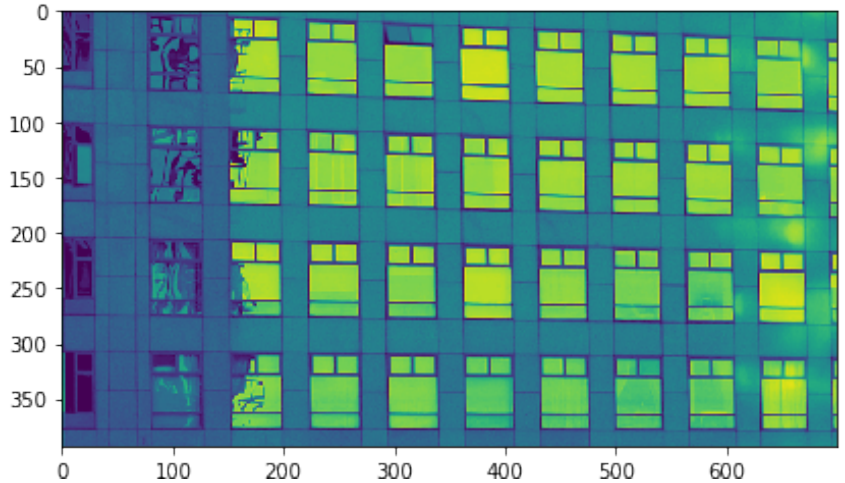


Max pooling

- The last operation before Deep Learning!
- Here's what it does:

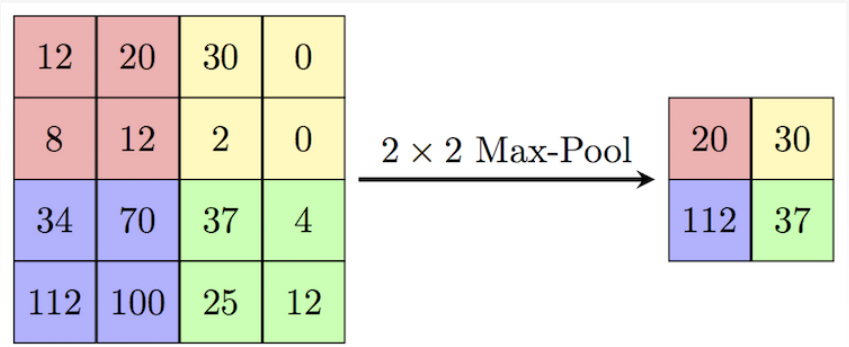


- An example:

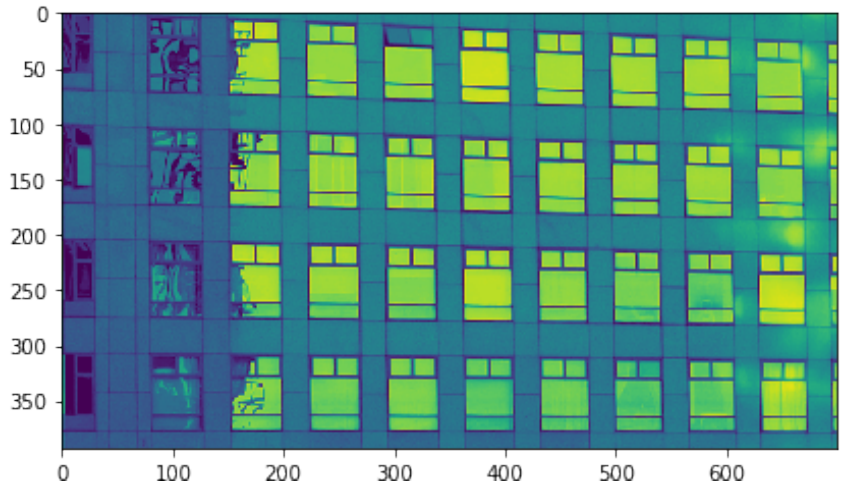


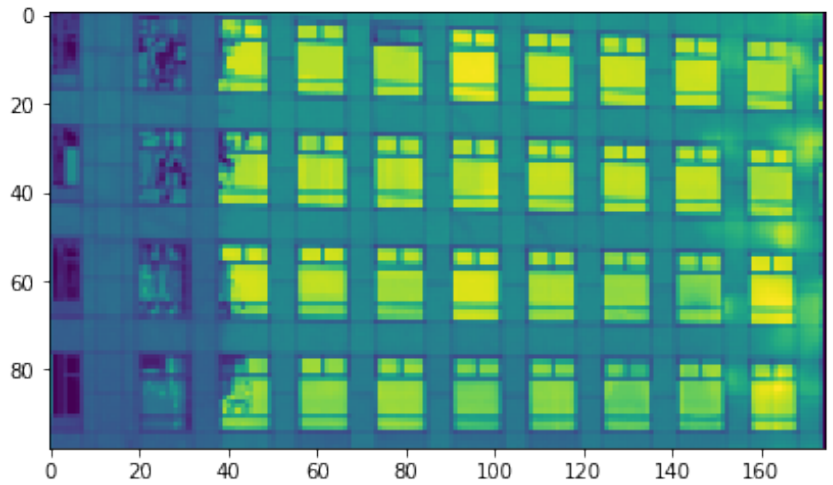
Max pooling

- The last operation before Deep Learning!
- Here's what it does:



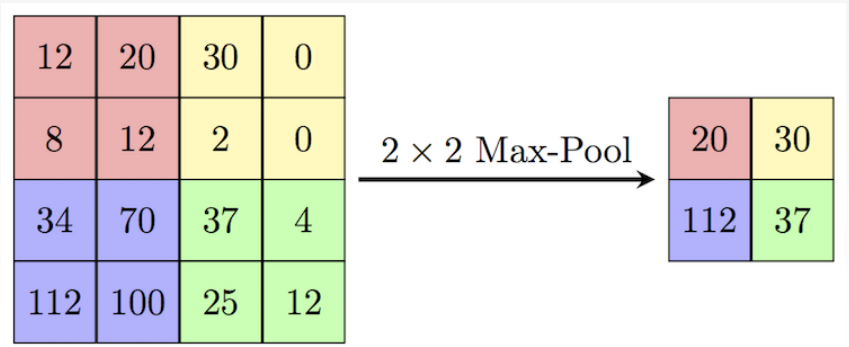
- An example:



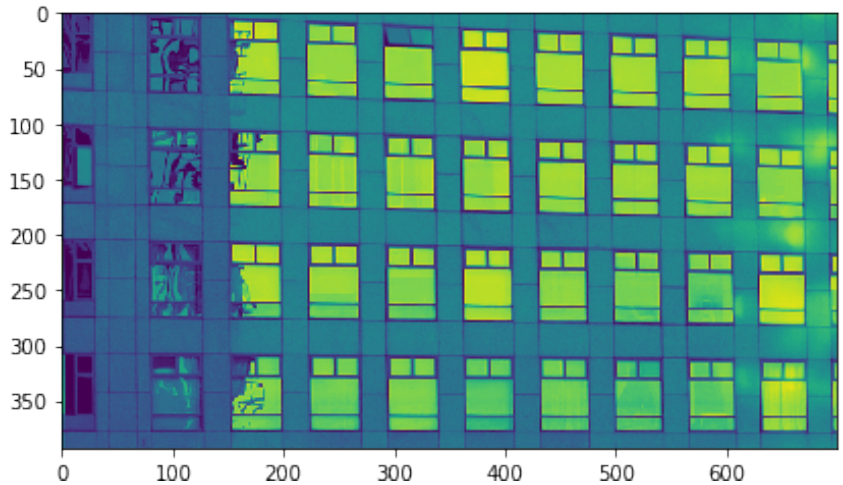


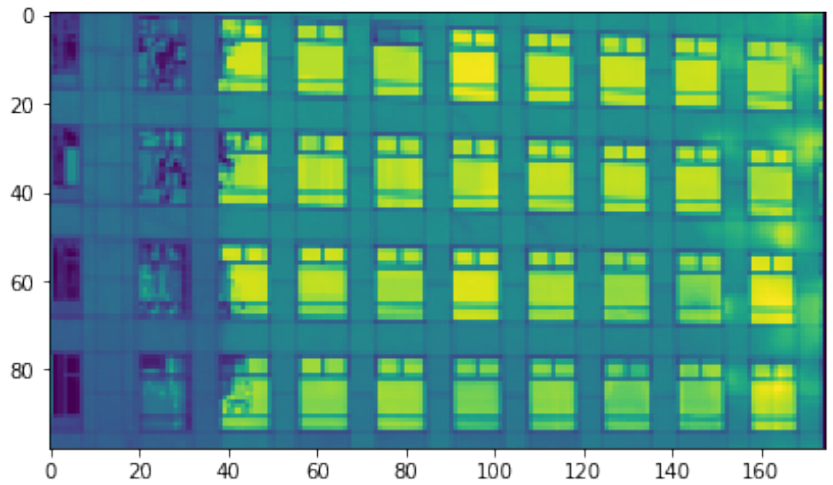
Max pooling

- The last operation before Deep Learning!
- Here's what it does:



- An example:





- What happened to the image?