Unveiling Bird Morphology Evolution through Deep-Learning Based Image Embedding and Gene Association Testing

Final presentation gene2birds group A

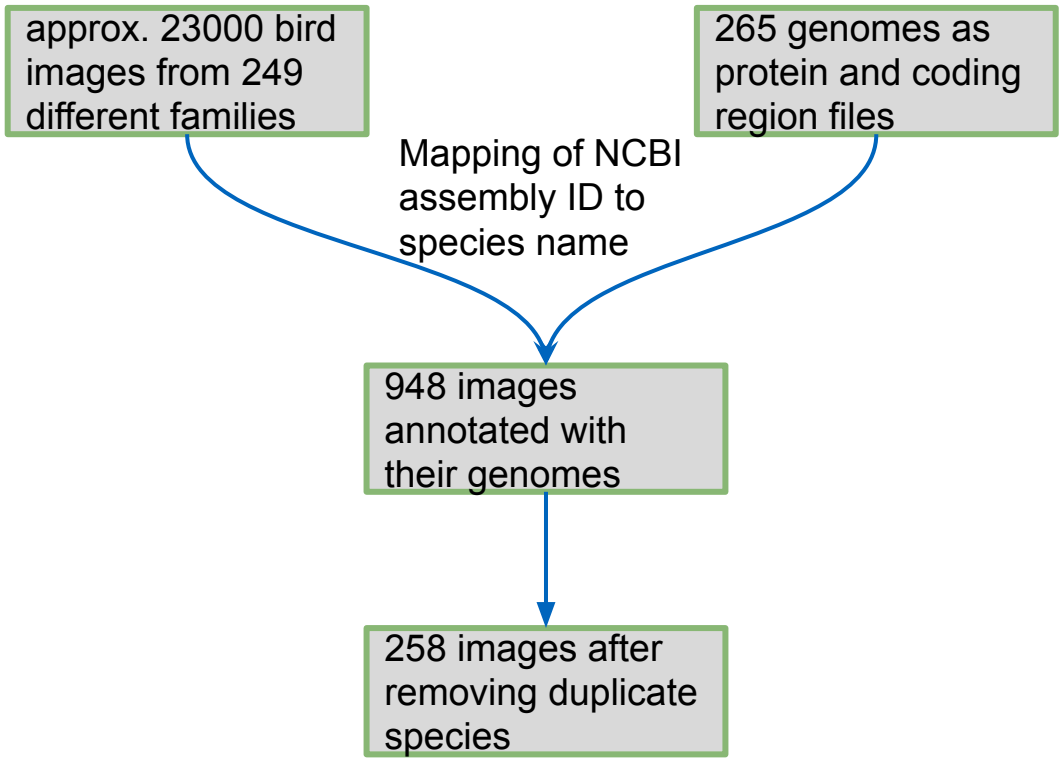
Joshua Günther
Berfin Erdoğan
Ufuk Demir
Mahima Arunkumar



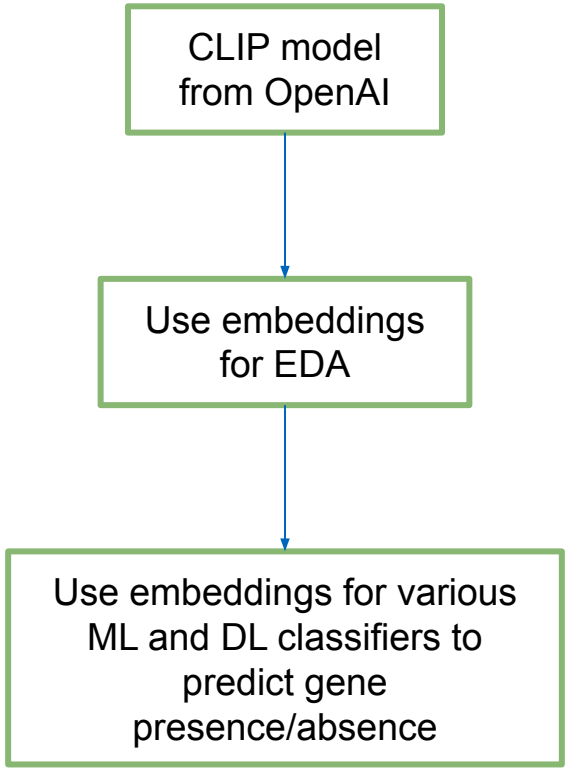





Input data


approx. 23000 bird images from 249 different families
265 genomes as protein and coding region files
Mapping of NCBI assembly ID to species name
948 images annotated with their genomes
258 images after removing duplicate species

Our approach



Absence/presence table by species

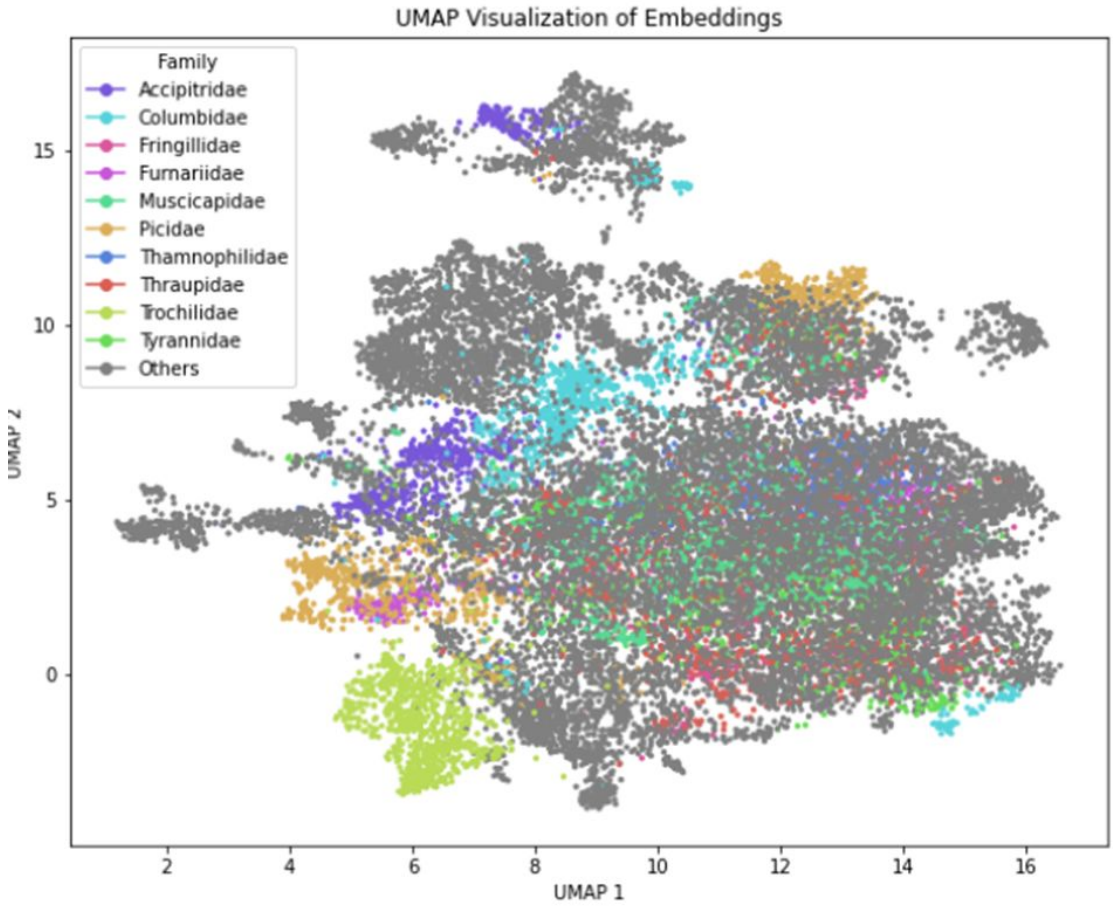
	Nothoprocta ornata	Smithornis capensis	Formicarius rufipectus	Sylvia atricapilla	Lanius ludovicianus	Amazona guildingii	Probosciger aterrimus	Eolophus roseicapilla	Chunga burmeisteri	Herpetotheres cachinnans
104K protein	False	False	False	False	False	False	False	False	False	False
110KD protein	False	False	False	False	False	False	False	False	False	False
1433B protein	True	False	True	True	False	True	True	True	True	True
1433E protein	True	True	True	True	True	True	True	True	False	False
1433F protein	True	True	True	True	True	True	True	True	True	True
...	...	...	...	...	...	...	...	...	...	...
ZY11B protein	True	True	False	True	True	True	True	True	True	True
ZYX protein	False	True	True	False	True	True	True	True	True	False
ZZEF1 protein	True	True	False	True	True	True	True	False	True	True
ZZZ3 protein	True	True	False	True	False	True	False	False	True	True
protein	False	False	False	False	False	False	False	False	False	False

17060 rows × 258 columns

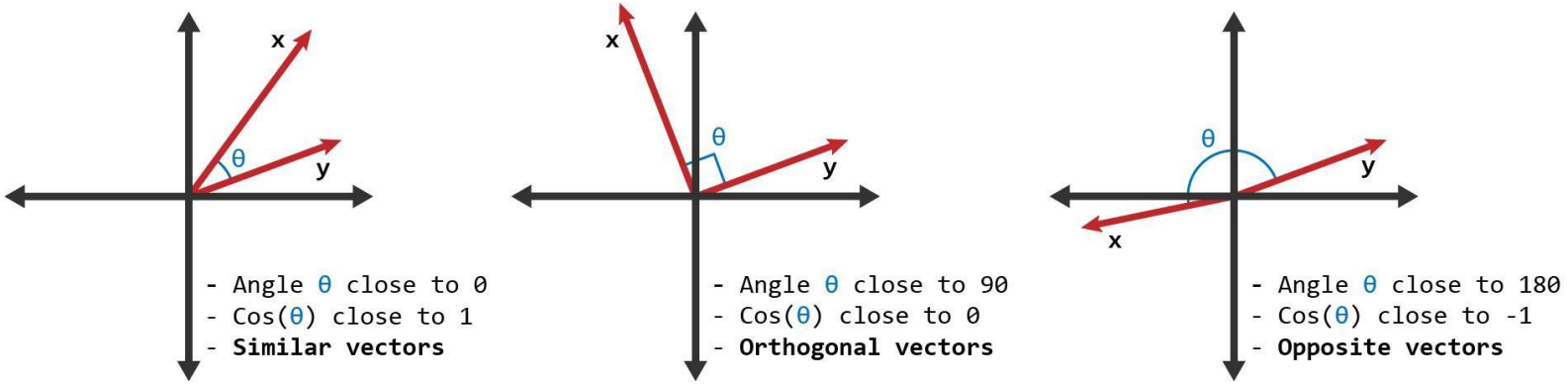
Generated embeddings with CLIP

	All 512 CLIP embedding features				
All 258 species					
Nothoprocta ornata	0.4072875	0.36193112	-0.31036815	...	0.13646944
Smithornis capensis	0.3131959	0.20072602	-0.43035		0.18086748
Formicarius rufipectus	0.37125307	-0.050516717	-0.52256465		0.25555477
Sylvia atricapilla	0.21032035	0.24915919	-0.6387599		0.5333118
Lanius ludovicianus	0.6003466	0.18232825	0.022367803		0.04154638
Amazona guildingii	0.33798376	0.14806776	-0.47427034		-0.19996244
Probosciger aterrimus	0.36870858	-0.056442954	-0.33242458		-0.04306676
Eolophus roseicapilla	0.40817946	0.23949404	-0.13612387		0.18286693
Chunga burmeisteri	0.3471905	0.34163687	-0.21079393		0.046399638
Herpetotheres cachin…	0.49282703	0.13153149	-0.42115825		0.22466694

Data exploration with CLIP embeddings

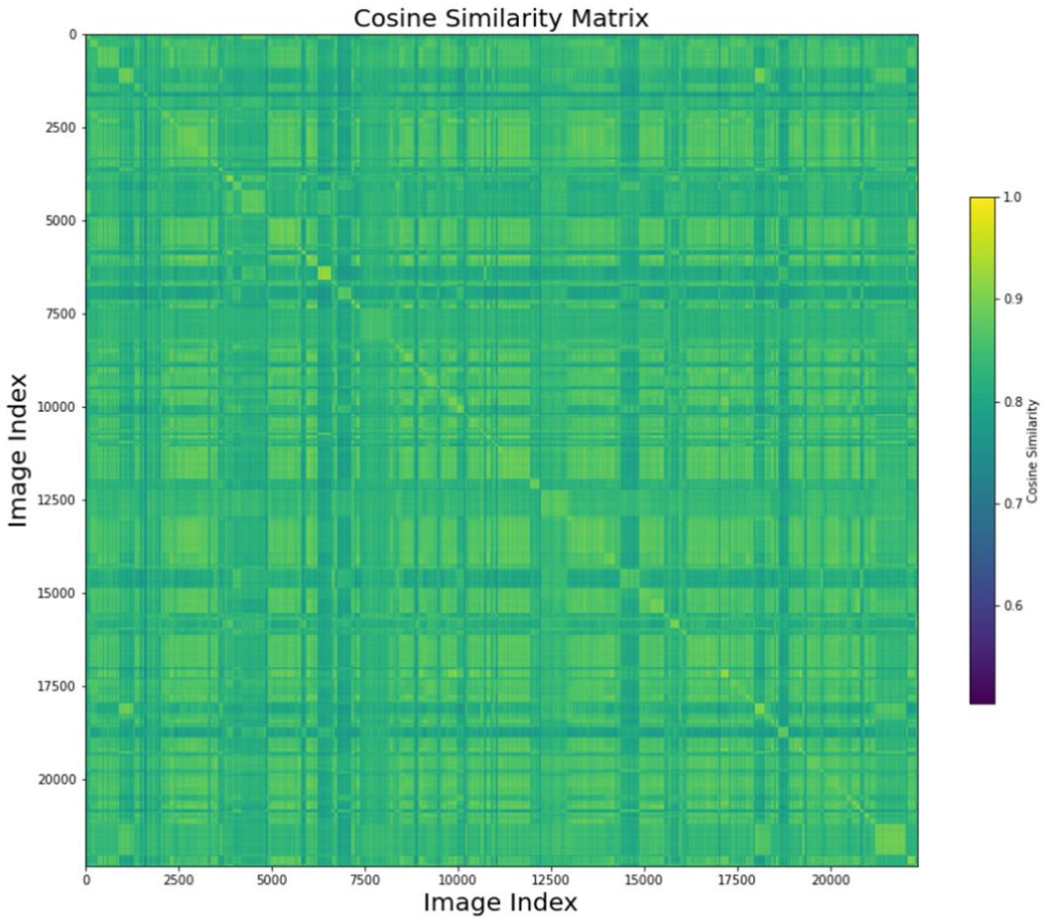


Data exploration with CLIP embeddings

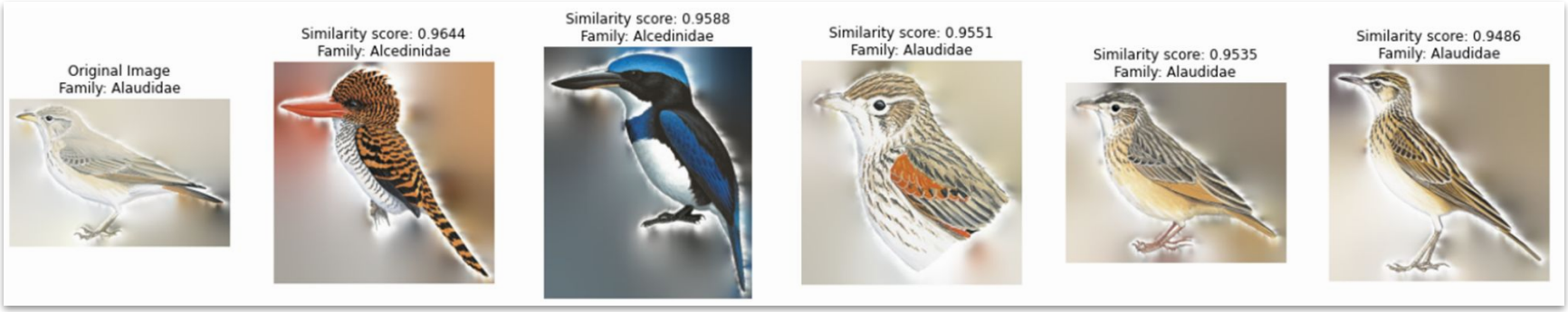


$$similarity(A, B) = cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

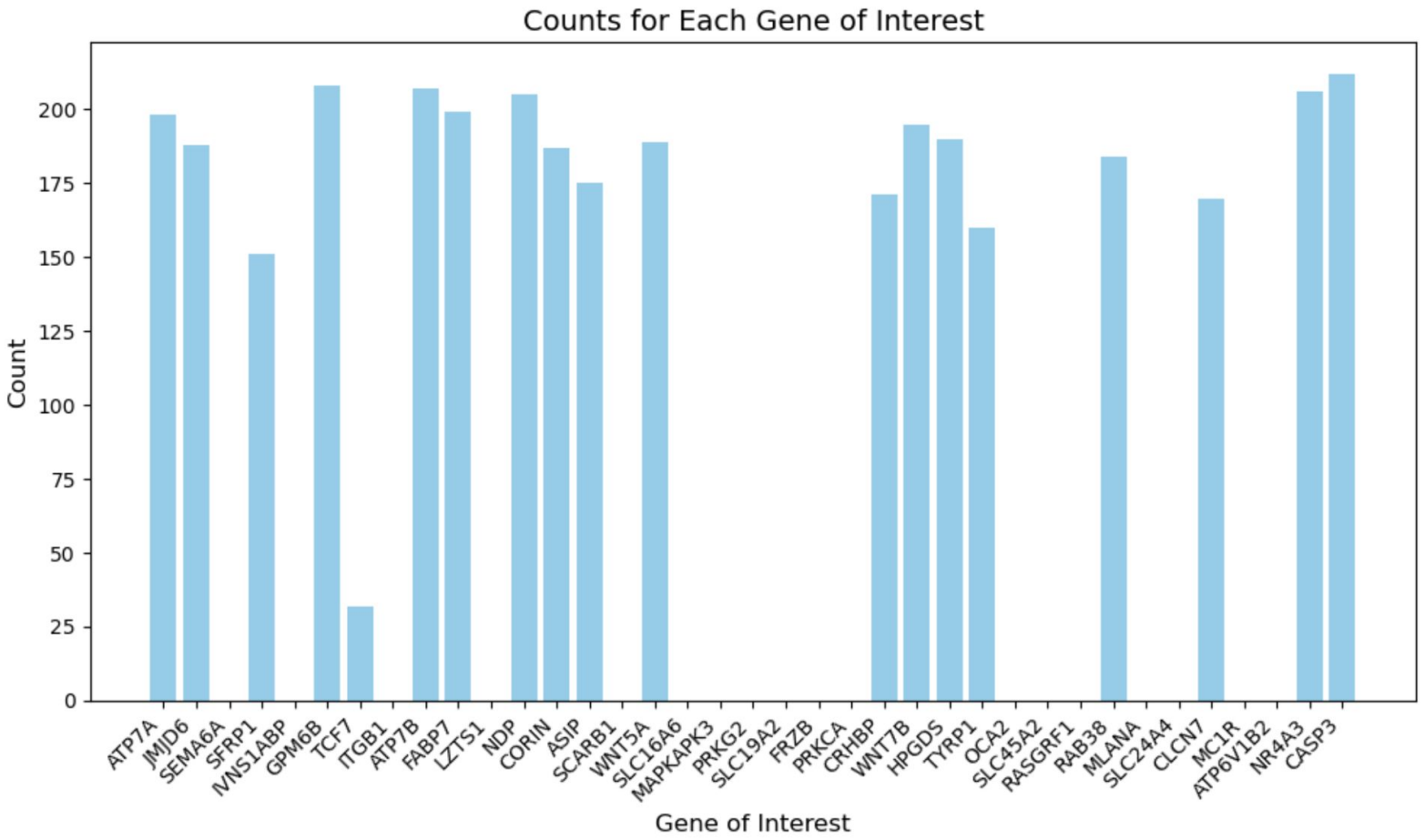
Data exploration with CLIP embeddings



Data exploration with CLIP embeddings



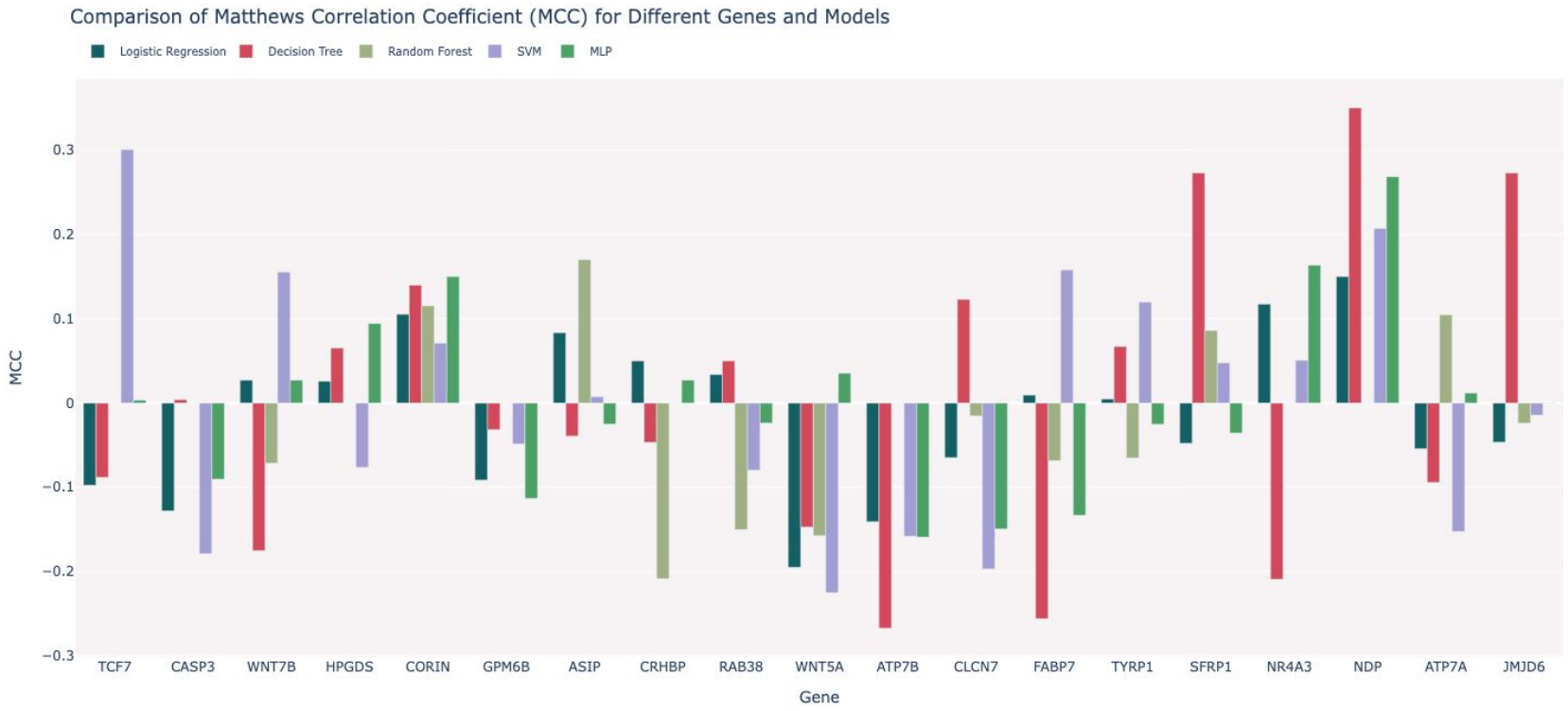
Coloration genes



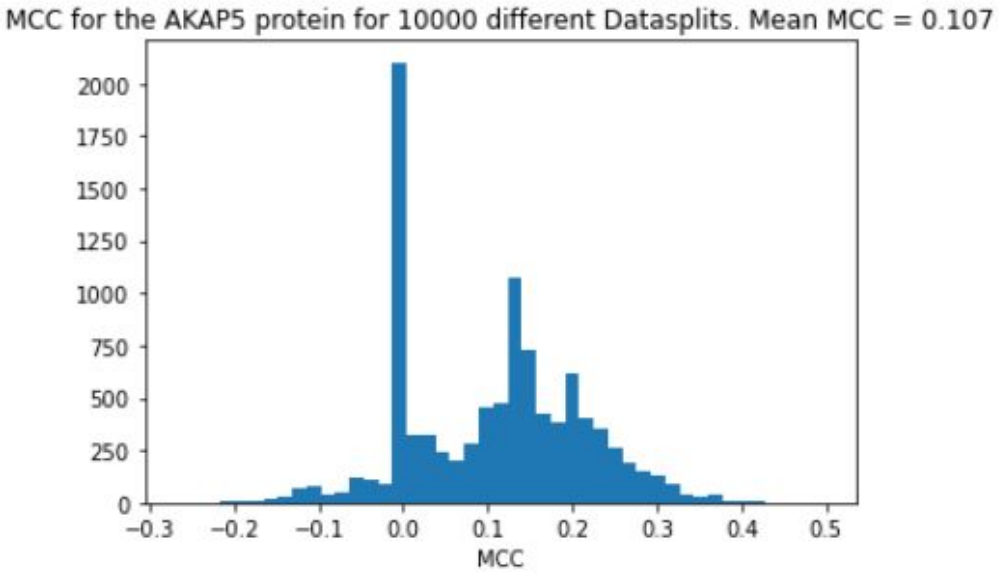
Input for our classifiers

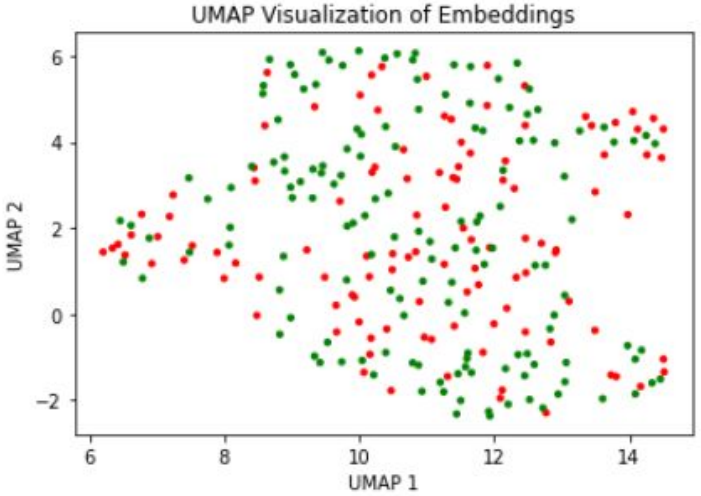
	All 512 CLIP embedding features				Y
All 258 species	Nothoprocta ornata	0.4072875	0.36193112	-0.31036815	...	0.13646944	1
	Smithornis capensis	0.3131959	0.20072602	-0.43035		0.18086748	0
	Formicarius rufipectus	0.37125307	-0.050516717	-0.52256465		0.25555477	1
	Sylvia atricapilla	0.21032035	0.24915919	-0.6387599		0.5333118	0
	Lanius ludovicianus	0.6003466	0.18232825	0.022367803		0.04154638	1
	Amazona guildingii	0.33798376	0.14806776	-0.47427034		-0.19996244	1
	Probosciger aterrimus	0.36870858	-0.056442954	-0.33242458		-0.04306676	0
	Eolophus roseicapilla	0.40817946	0.23949404	-0.13612387		0.18286693	0
	Chunga burmeisteri	0.3471905	0.34163687	-0.21079393		0.046399638	1
	Herpetotheres cachin…	0.49282703	0.13153149	-0.42115825		0.22466694	...

Classical ML models



Bootstrapping





Gene2Bird

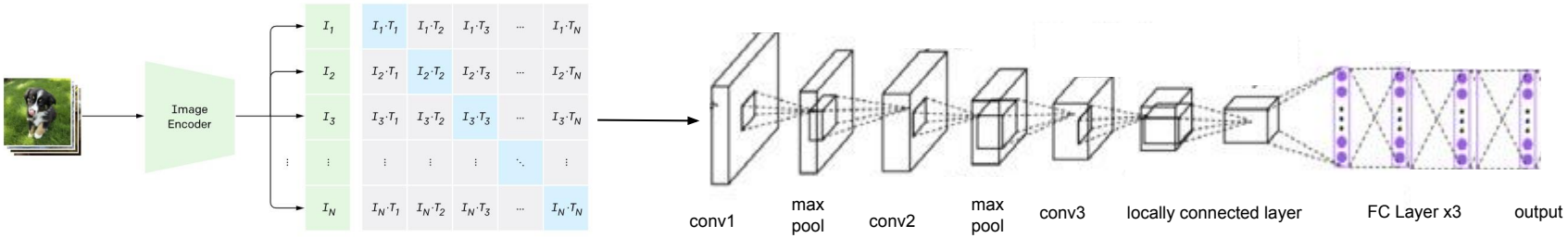


Image Encoder
conv1
max pool
conv2
max pool
conv3
locally connected layer
FC Layer x3
output

ResNet18

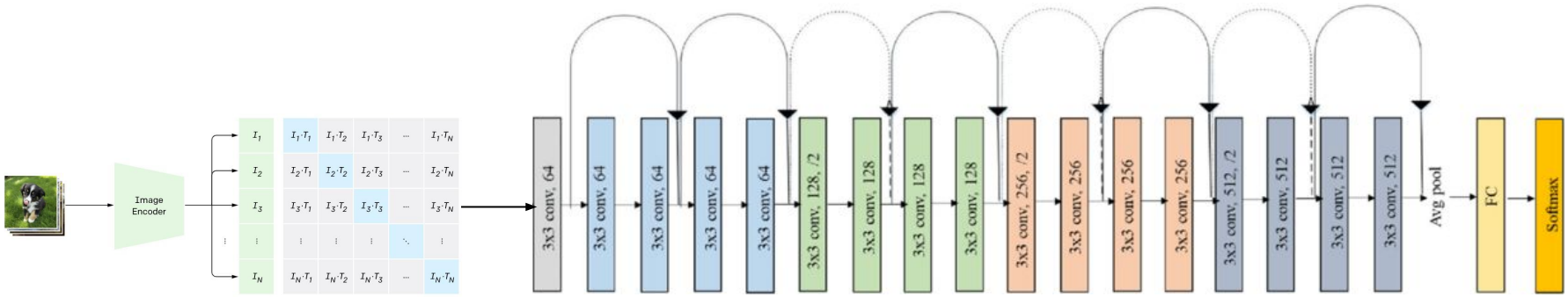


Image Encoder
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 256, /2
3x3 conv, 512, /2
Avg pool
FC
Softmax

ResNet50

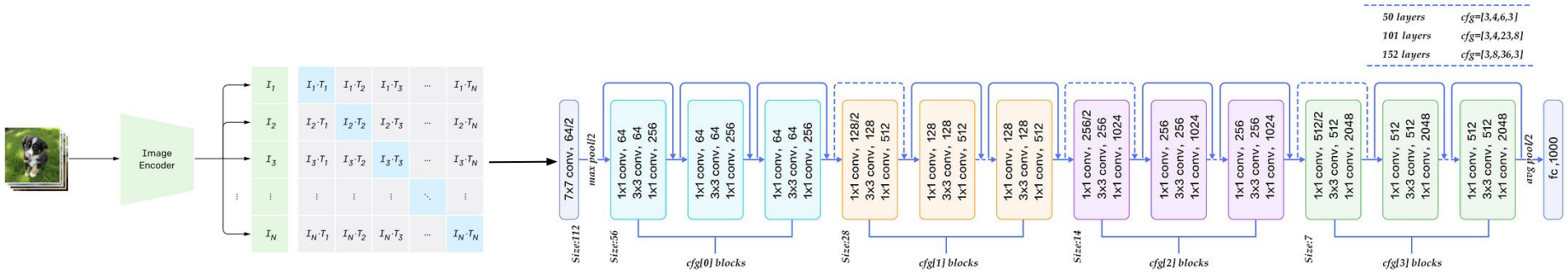


Image Encoder
50 layers cfg=[3,4,6,3]
101 layers cfg=[3,4,23,8]
152 layers cfg=[3,8,36,3]
7x7 conv, 64/2
max pool/2
1x1 conv, 64
3x3 conv, 64
1x1 conv, 256
1x1 conv, 128/2
3x3 conv, 128
1x1 conv, 512
1x1 conv, 128
1x1 conv, 256/2
3x3 conv, 256
1x1 conv, 1024
1x1 conv, 256
1x1 conv, 512/2
3x3 conv, 512
1x1 conv, 2048
1x1 conv, 512
avg pool/2
fc, 1000
Size:112
Size:56
Size:28
Size:14
Size:7
cfg[0] blocks
cfg[1] blocks
cfg[2] blocks
cfg[3] blocks

ResNeSt

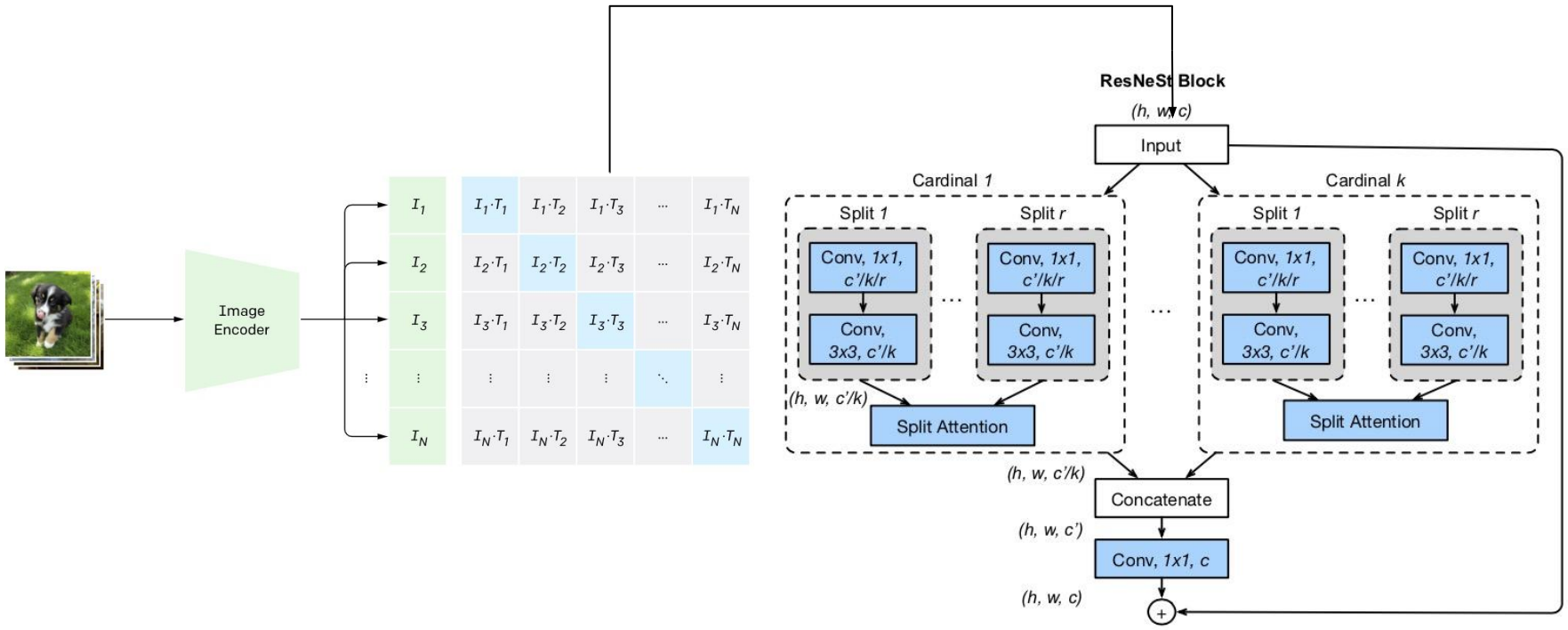
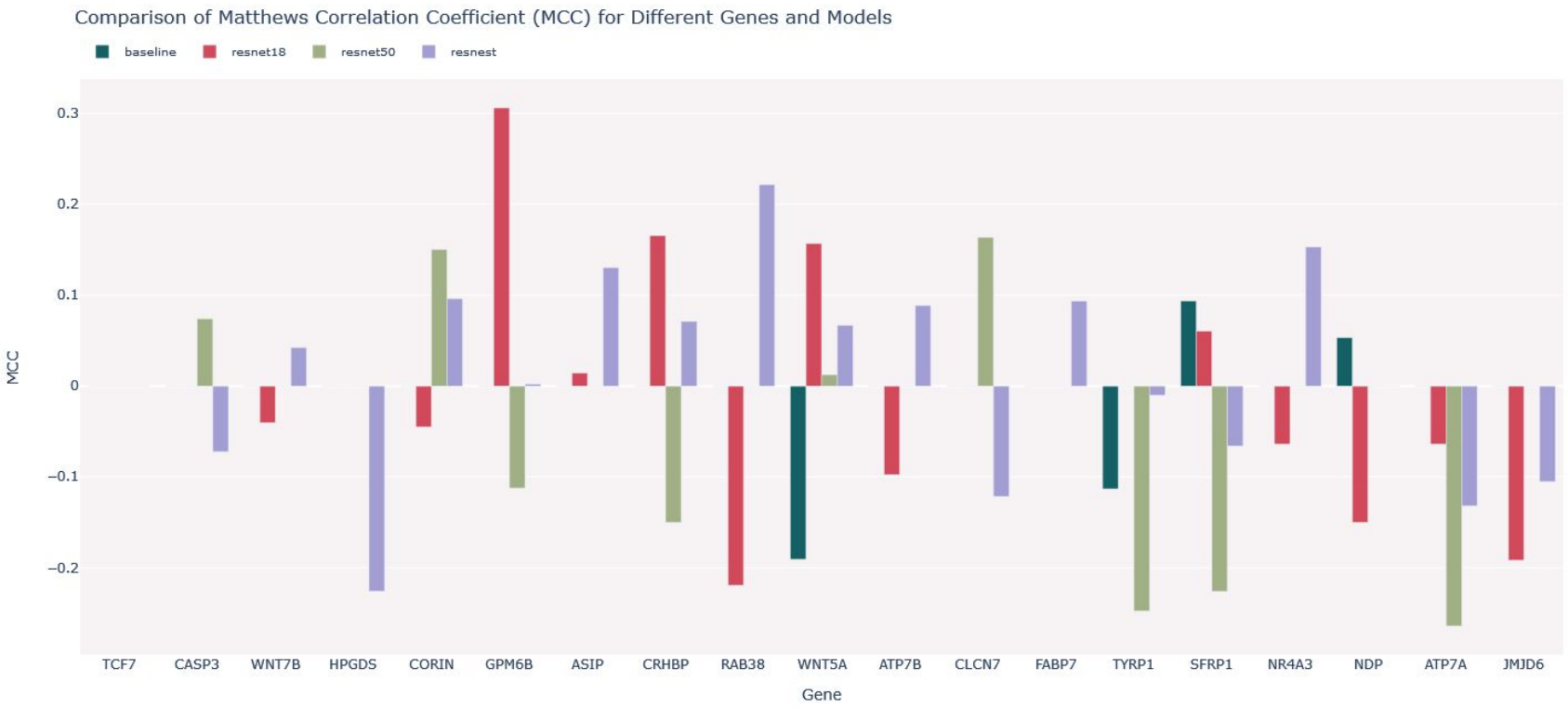


Image Encoder
ResNeSt Block
(h, w, c)
Input
Cardinal 1
Cardinal k
Split 1
Split r
Conv, 1x1, c'/k/r
Conv, 3x3, c'/k
(h, w, c'/k)
Split Attention
(h, w, c')
Concatenate
Conv, 1x1, c
(h, w, c)

Results of Deep Learning Models



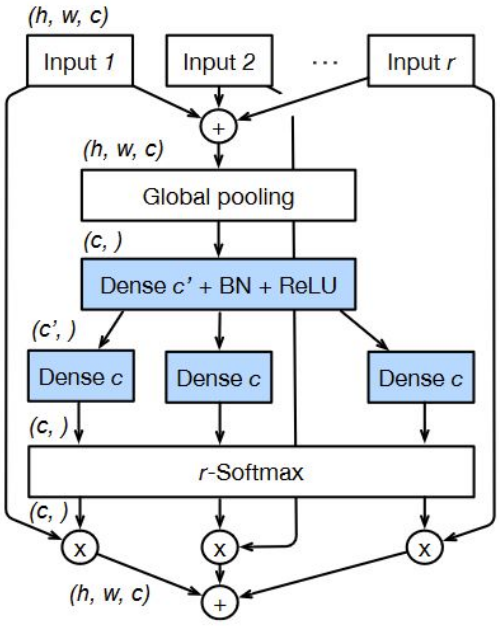
Potential Issues

- CLIP might not have captured all the necessary information
- Poor genbank annotation
- Presence or absence of a gene may not provide sufficient information

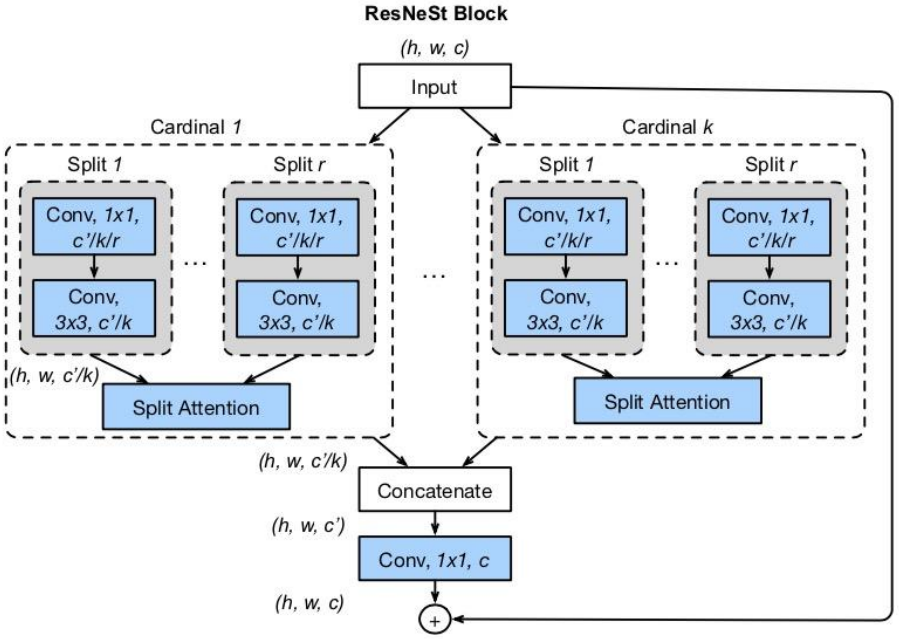
Thank you for your time and attention!

Any questions?

ResNeSt - Details



Split-Attention within a cardinal group



ResNeSt block