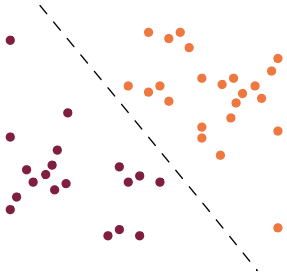
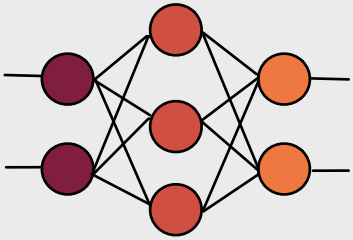
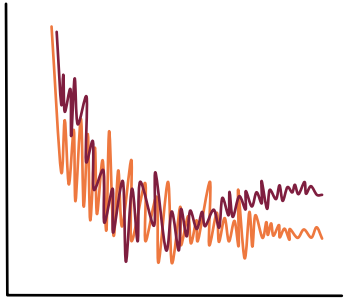
Overview



1. Find some data



2. Design a network



3. Train the network

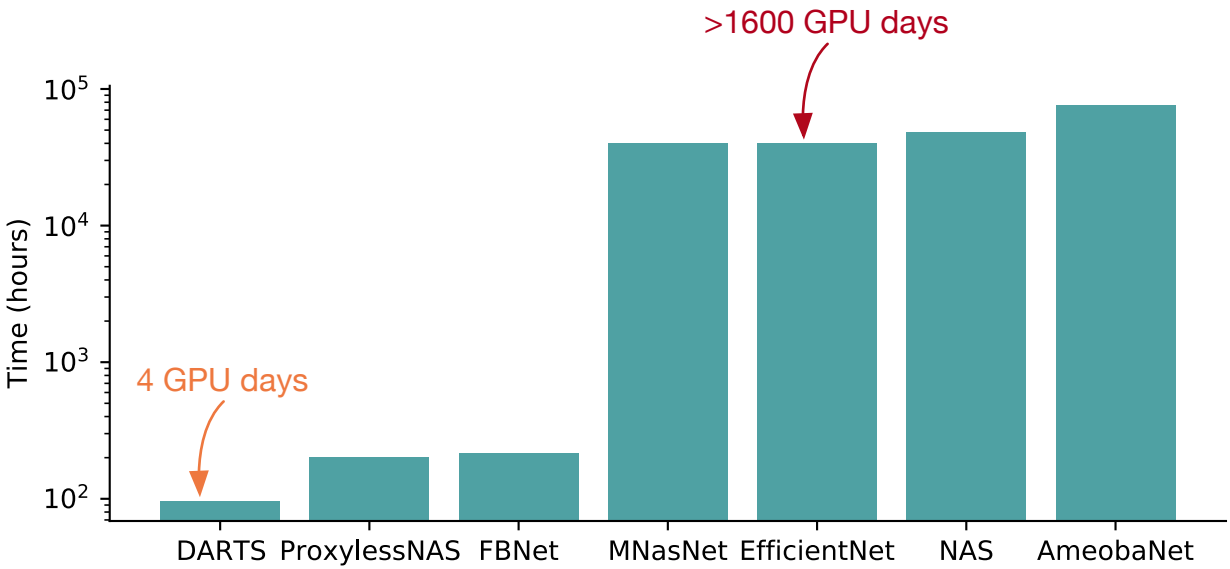
Neural Architecture Search (NAS)

Algorithm 1 Standard NAS

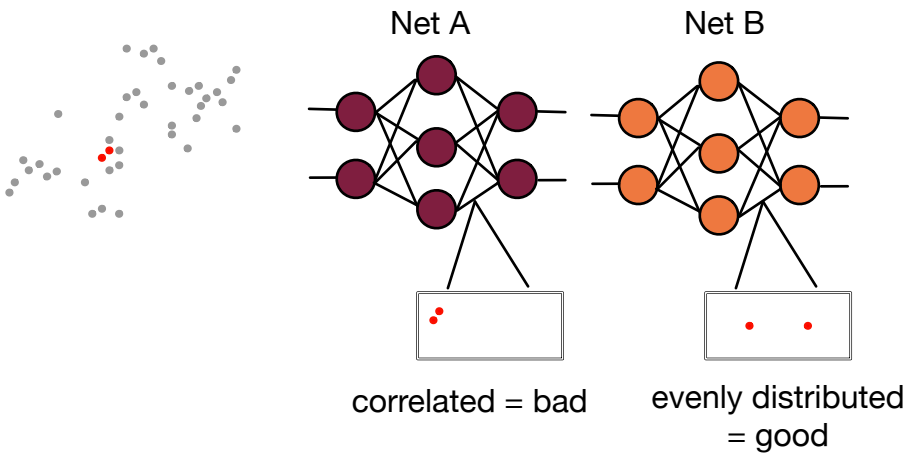
```
generator = RNN()

for i=1:N do
    net = generator.generate()
    trained_net = net.train()
    ▷ Training a net every step is expensive
    generator.update(trained_net)
chosen_net = generator.generate()
```

NAS is slow


>1600 GPU days
4 GPU days
Time (hours)
10^5
10^4
10^3
10^2
DARTS
ProxylessNAS
FBNet
MNasNet
EfficientNet
NAS
AmeobaNet

Is there a property of networks *without training* that we can use?

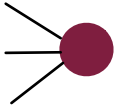


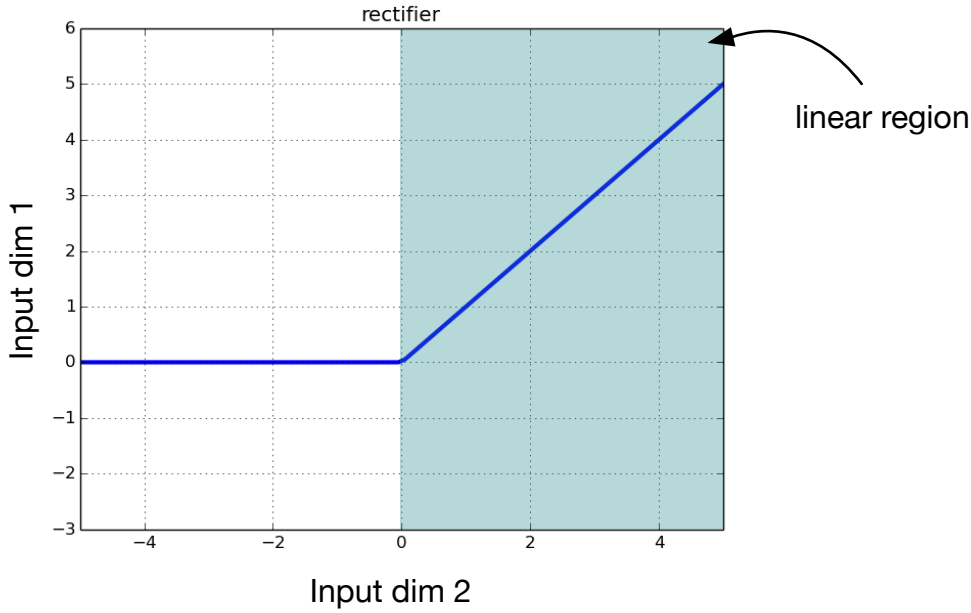
Linear Regions in Neural Networks

Input

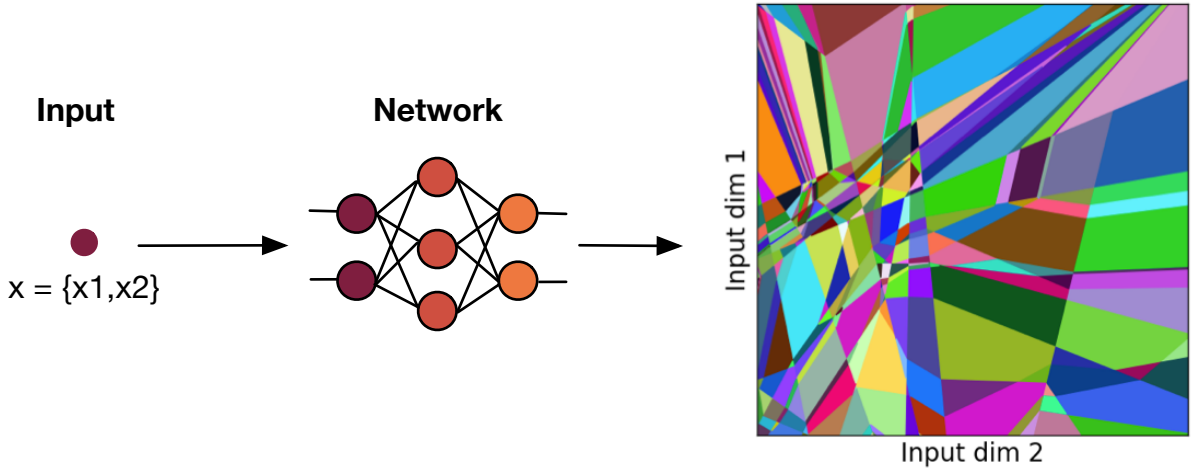
x = {x1,x2}

Neuron



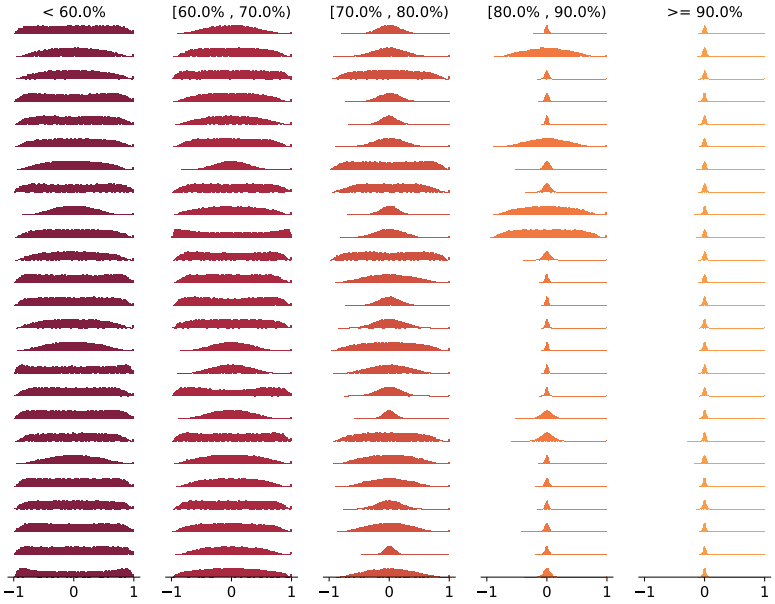


Linear Regions

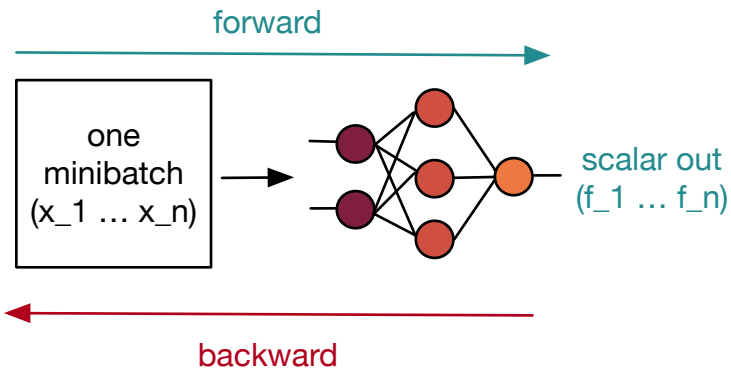


[1] Hanin, B. and Rolnick, D., 2019. Deep relu networks have surprisingly few activation patterns. In *Advances in Neural Information Processing Systems* (pp. 361-370).

Correlating Linear Regions



Correlating Linear Regions



$$\mathbf{J} = \begin{pmatrix} \frac{\partial f_1}{\partial \mathbf{x_1}} & \frac{\partial f_2}{\partial \mathbf{x_2}} & \cdots & \frac{\partial f_N}{\partial \mathbf{x_N}} \end{pmatrix}^\top$$

Scoring Jacobians

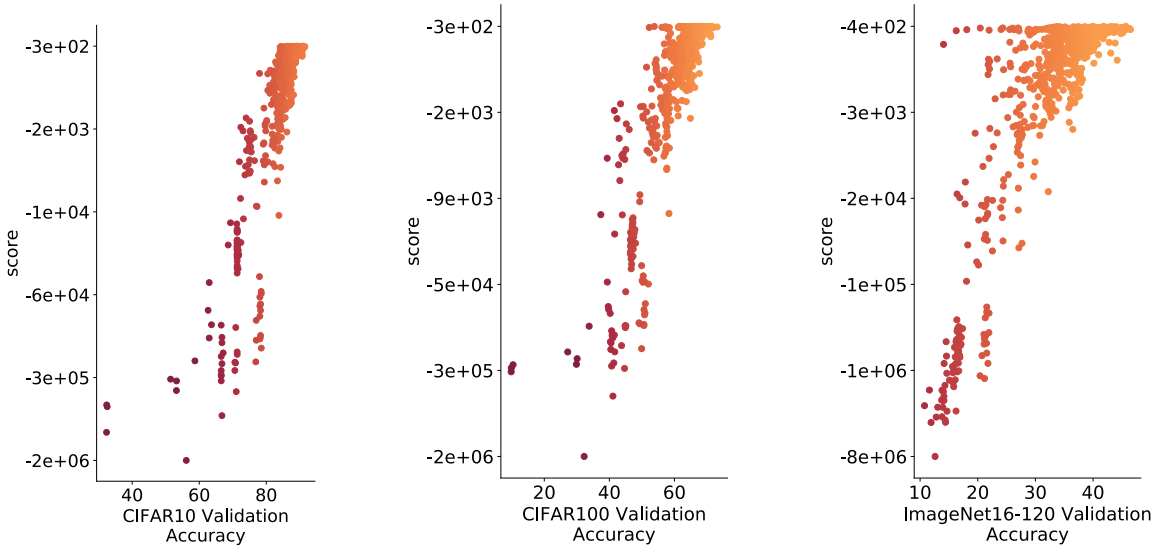
1. Get the input-output Jacobian (J) for one minibatch of data

2. Take eigenvalues of the correlation matrix of J

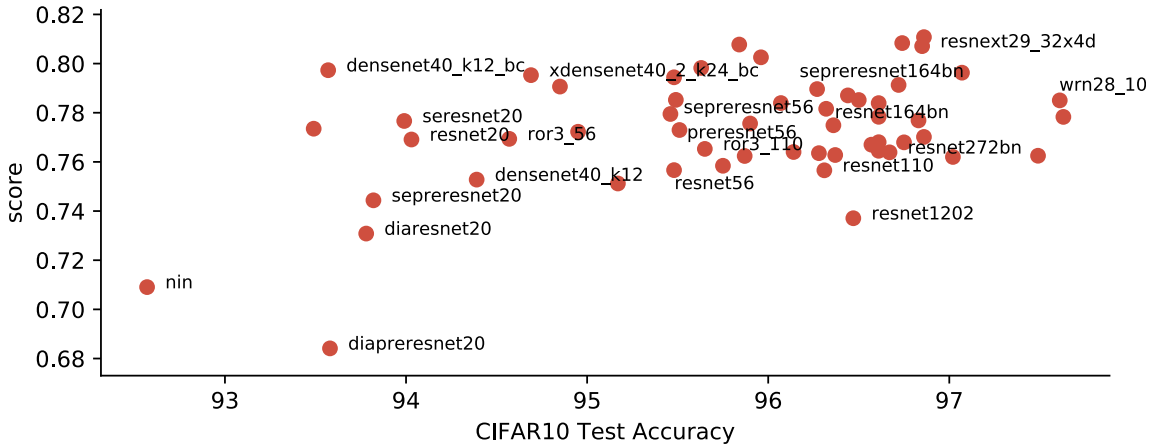
3. Score by KL divergence between two Gaussians:
   - one with 2. as the kernel
   - one with an uncorrelated kernel

```
def eval_score(jacob):
    corrs = np.corrcoef(jacob)
    v, _  = np.linalg.eig(corrs)
    k = 1e-5
    return -np.sum(np.log(v + k) + 1./(v + k))
```

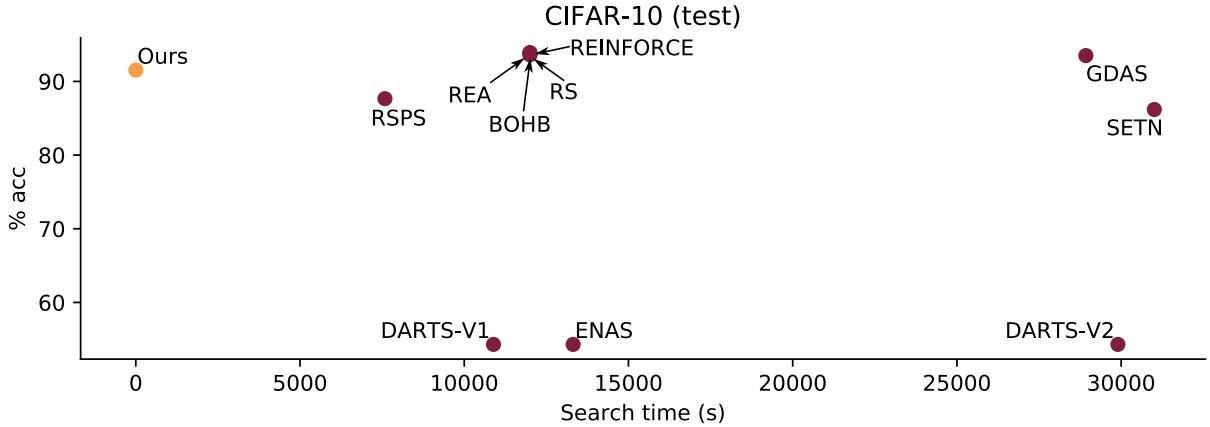
Results


-3e+02
-2e+03
-1e+04
-6e+04
-3e+05
-2e+06
score
40
60
80
CIFAR10 Validation Accuracy
-3e+02
-2e+03
-9e+03
-5e+04
-3e+05
-2e+06
score
20
40
60
CIFAR100 Validation Accuracy
-4e+02
-3e+03
-2e+04
-1e+05
-1e+06
-8e+06
score
10
20
30
40
ImageNet16-120 Validation Accuracy

Results


0.82
0.80
0.78
0.76
0.74
0.72
0.70
0.68
score
93
94
95
96
97
CIFAR10 Test Accuracy
resnext29_32x4d
densenet40_k12_bc
xdensenet40_2_k24_bc
sepreresnet164bn
wrn28_10
seresnet20
sepreresnet56
resnet164bn
resnet20
ror3_56
preresnet56
ror3_110
resnet272bn
densenet40_k12
resnet110
resnet56
sepreresnet20
resnet1202
diaresnet20
nin
diapreresnet20

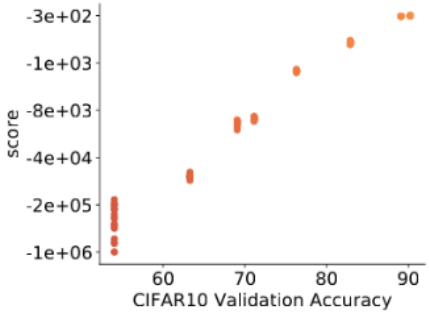
Results


CIFAR-10 (test)
Ours
REINFORCE
GDAS
REA
RS
RSPS
BOHB
SETN
DARTS-V1
ENAS
DARTS-V2
% acc
90
80
70
60
0
5000
10000
15000
20000
25000
30000
Search time (s)

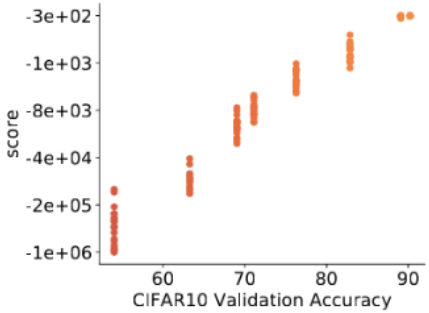
Questions?

Code: https://github.com/BayesWatch/nas-without-training

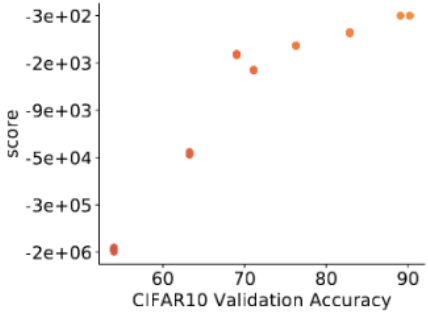
Control studies



(a) vary minibatch



(b) vary initialisation



(c) vary white noise input