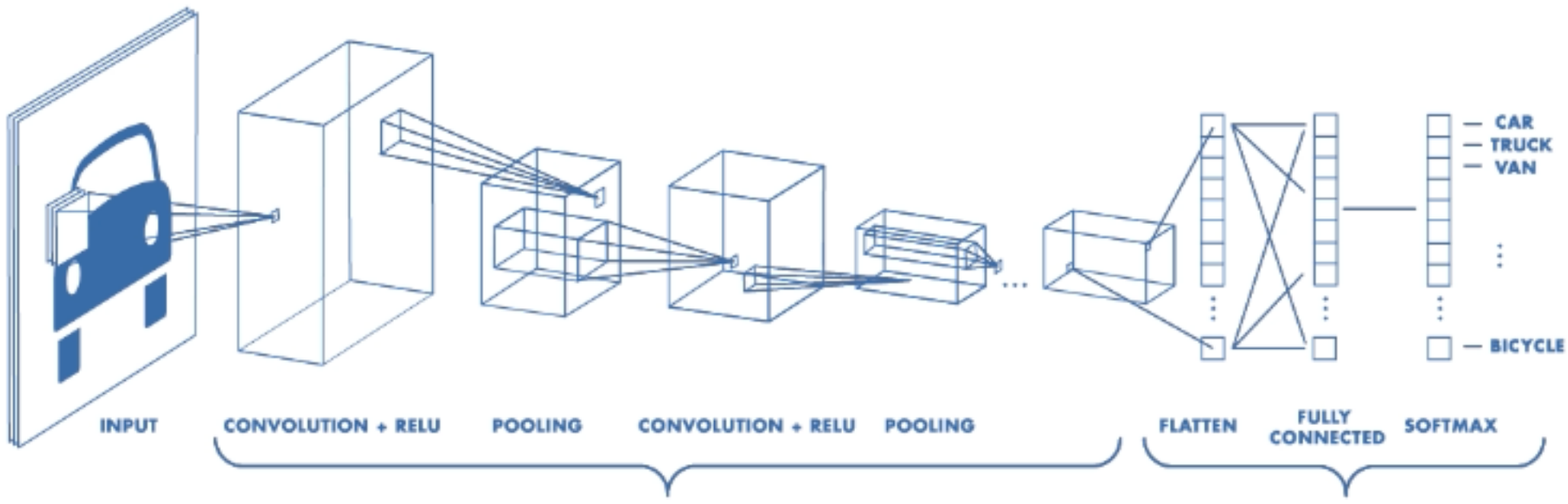
DENSER: Deep Evolutionary Network Structured Representation

Filipe Assunção, Nuno Lourenço, Penousal Machado and Bernardete Ribeiro
University of Coimbra, Coimbra, Portugal
{fga, naml, machado, bribeiro}@dei.uc.pt

AUTOMATED DEEP NEURAL NETWORK DESIGN

- Select the Artificial Neural Network (ANN) type;
- Choose the sequence, type, and number of layers;
- Fine-tune the parameters of each layer;
- Decide the learning algorithm;
- Optimise the parameters of the learning algorithm.

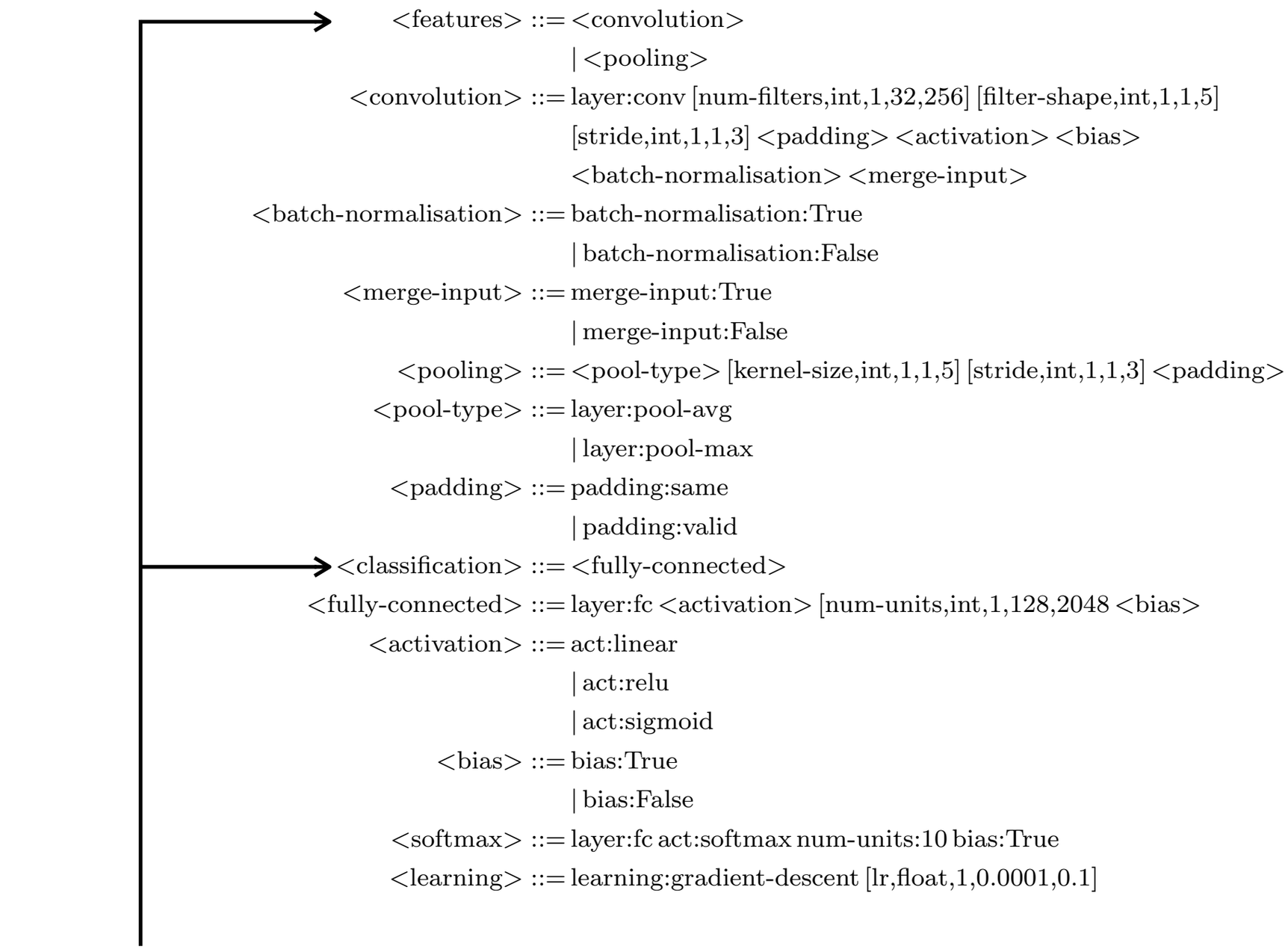
CONVOLUTIONAL NEURAL NETWORK



feature extraction / representation learning

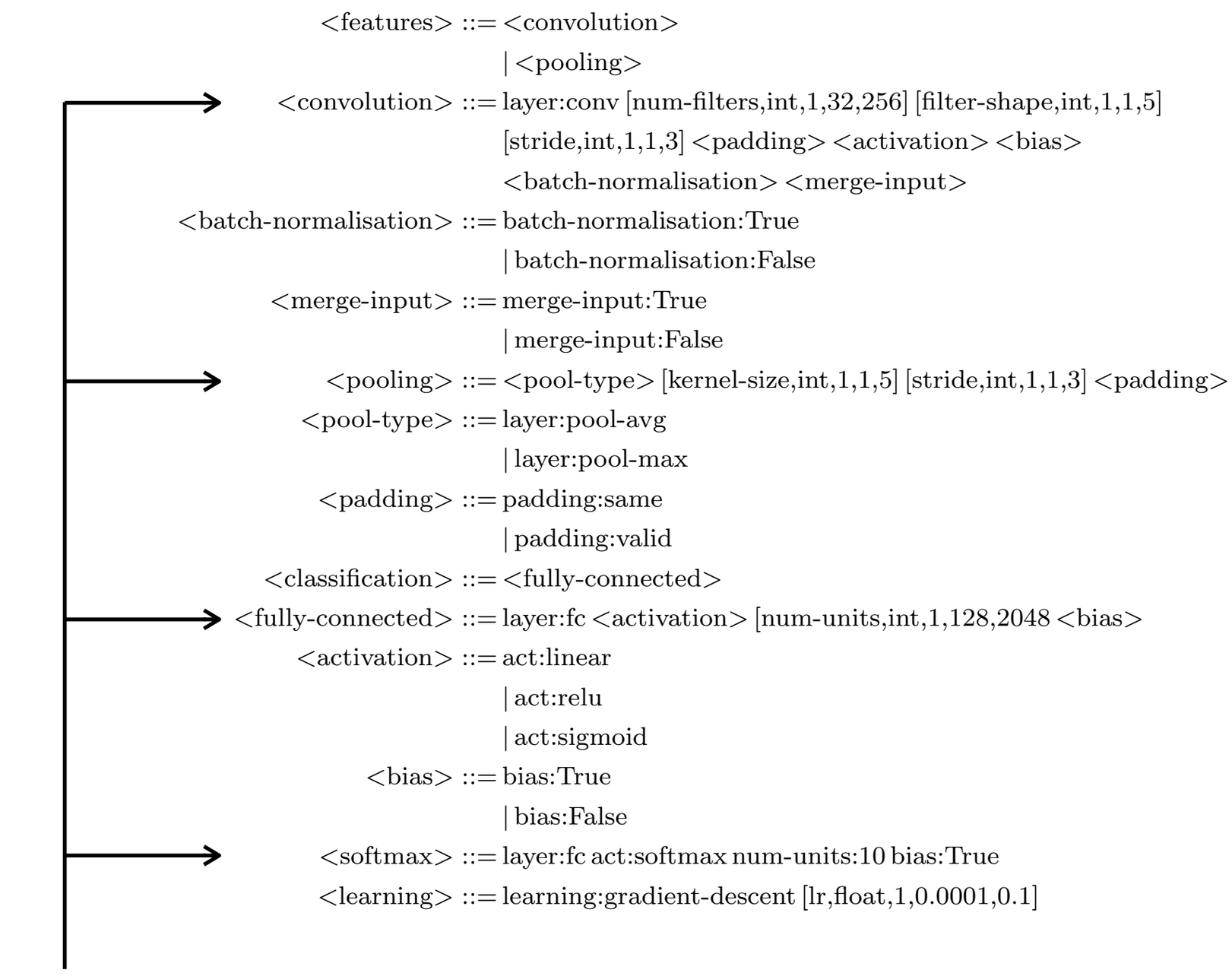
classification

DENSER



<features> ::= <convolution>
| <pooling>
<convolution> ::= layer:conv [num-filters,int,1,32,256] [filter-shape,int,1,1,5]
[stride,int,1,1,3] <padding> <activation> <bias>
<batch-normalisation> <merge-input>
<batch-normalisation> ::= batch-normalisation:True
| batch-normalisation:False
<merge-input> ::= merge-input:True
| merge-input:False
<pooling> ::= <pool-type> [kernel-size,int,1,1,5] [stride,int,1,1,3] <padding>
<pool-type> ::= layer:pool-avg
| layer:pool-max
<padding> ::= padding:same
| padding:valid
<classification> ::= <fully-connected>
<fully-connected> ::= layer:fc <activation> [num-units,int,1,128,2048 <bias>
<activation> ::= act:linear
| act:relu
| act:sigmoid
<bias> ::= bias:True
| bias:False
<softmax> ::= layer:fc act:softmax num-units:10 bias:True
<learning> ::= learning:gradient-descent [lr,float,1,0.0001,0.1]

DENSER



<features> ::= <convolution>
| <pooling>

<convolution> ::= layer:conv [num-filters,int,1,32,256] [filter-shape,int,1,1,5]
[stride,int,1,1,3] <padding> <activation> <bias>
<batch-normalisation> <merge-input>

<batch-normalisation> ::= batch-normalisation:True
| batch-normalisation:False

<merge-input> ::= merge-input:True
| merge-input:False

<pooling> ::= <pool-type> [kernel-size,int,1,1,5] [stride,int,1,1,3] <padding>

<pool-type> ::= layer:pool-avg
| layer:pool-max

<padding> ::= padding:same
| padding:valid

<classification> ::= <fully-connected>

<fully-connected> ::= layer:fc <activation> [num-units,int,1,128,2048 <bias>

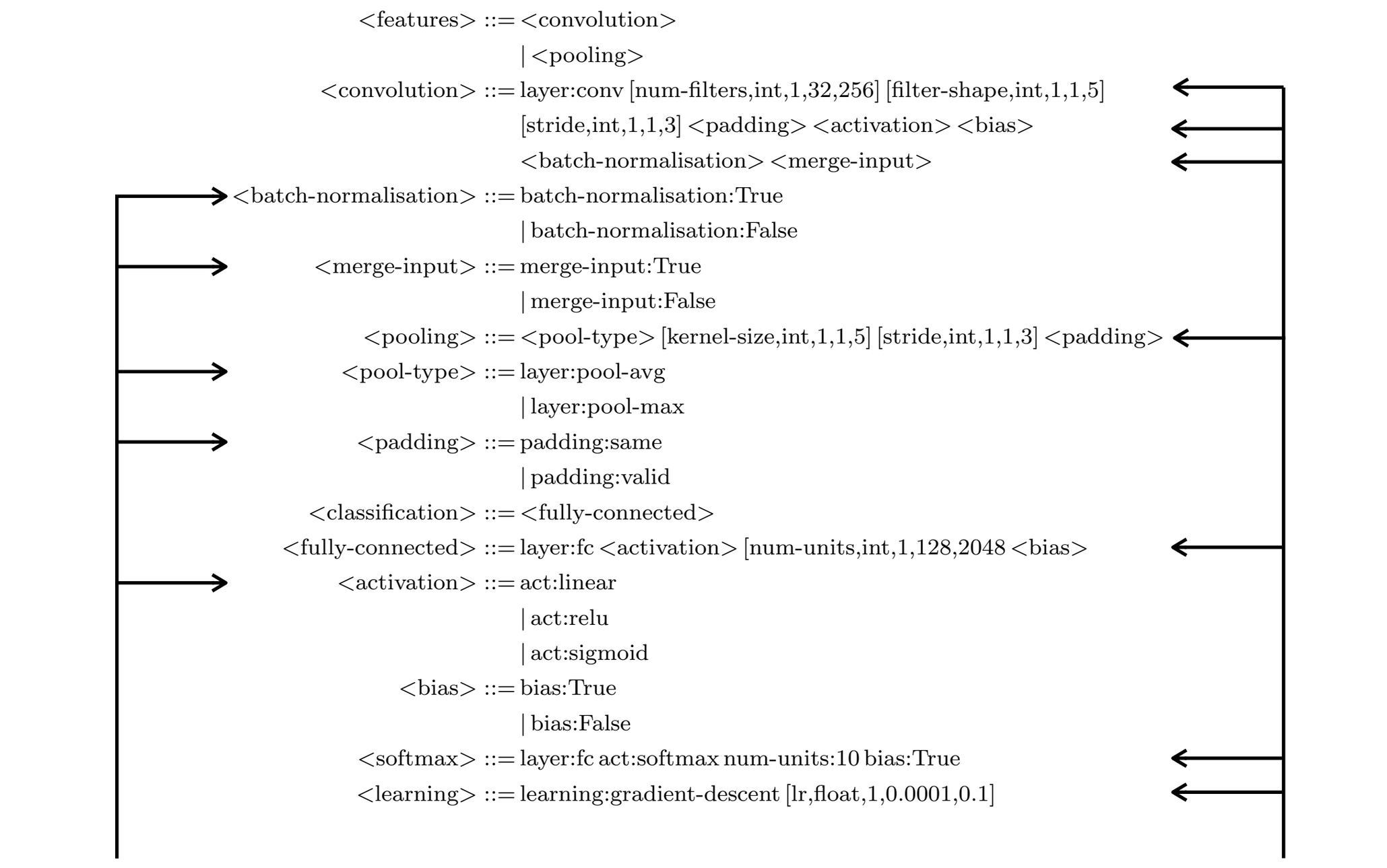
<activation> ::= act:linear
| act:relu
| act:sigmoid

<bias> ::= bias:True
| bias:False

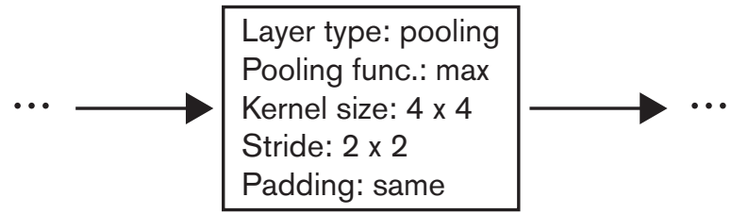
<softmax> ::= layer:fc act:softmax num-units:10 bias:True

<learning> ::= learning:gradient-descent [lr,float,1,0.0001,0.1]

DENSER


<features> ::= <convolution>
| <pooling>
<convolution> ::= layer:conv [num-filters,int,1,32,256] [filter-shape,int,1,1,5]
[stride,int,1,1,3] <padding> <activation> <bias>
<batch-normalisation> <merge-input>
<batch-normalisation> ::= batch-normalisation:True
| batch-normalisation:False
<merge-input> ::= merge-input:True
| merge-input:False
<pooling> ::= <pool-type> [kernel-size,int,1,1,5] [stride,int,1,1,3] <padding>
<pool-type> ::= layer:pool-avg
| layer:pool-max
<padding> ::= padding:same
| padding:valid
<classification> ::= <fully-connected>
<fully-connected> ::= layer:fc <activation> [num-units,int,1,128,2048 <bias>
<activation> ::= act:linear
| act:relu
| act:sigmoid
<bias> ::= bias:True
| bias:False
<softmax> ::= layer:fc act:softmax num-units:10 bias:True
<learning> ::= learning:gradient-descent [lr,float,1,0.0001,0.1]
close-choice parameters
real-valued parameters

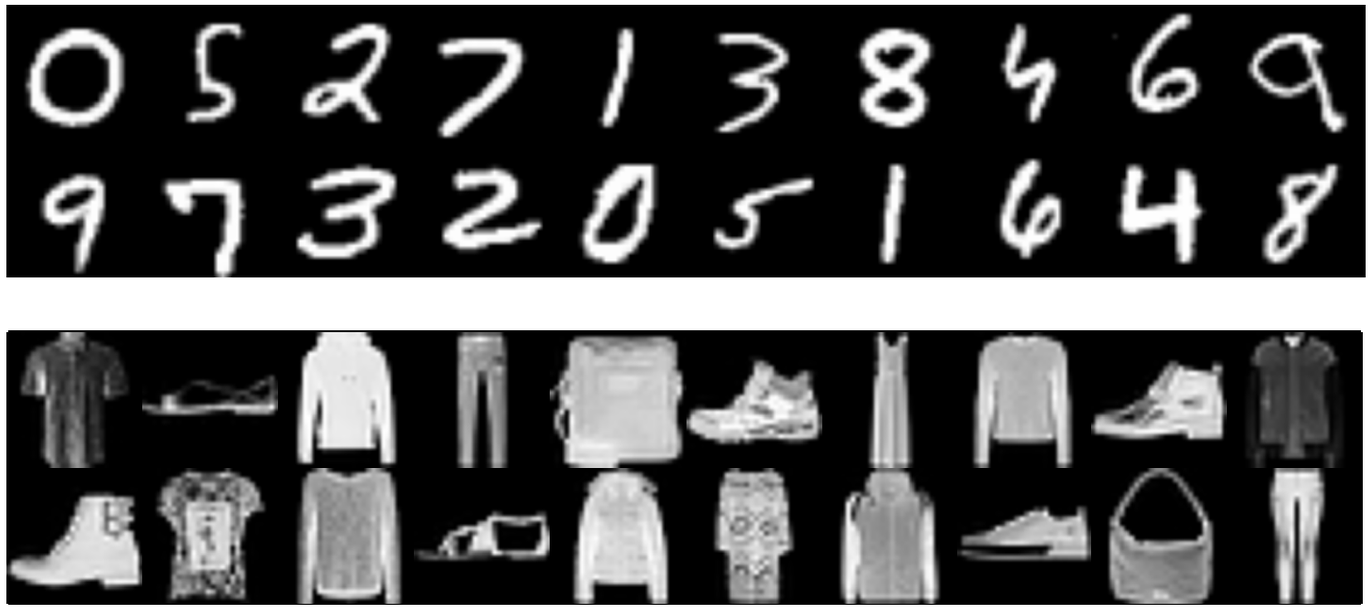
EXAMPLE OF A CANDIDATE SOLUTION

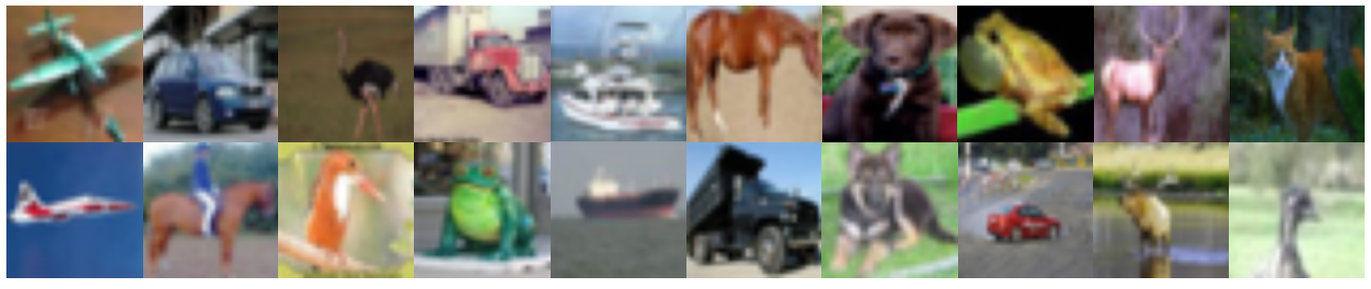


HINTON

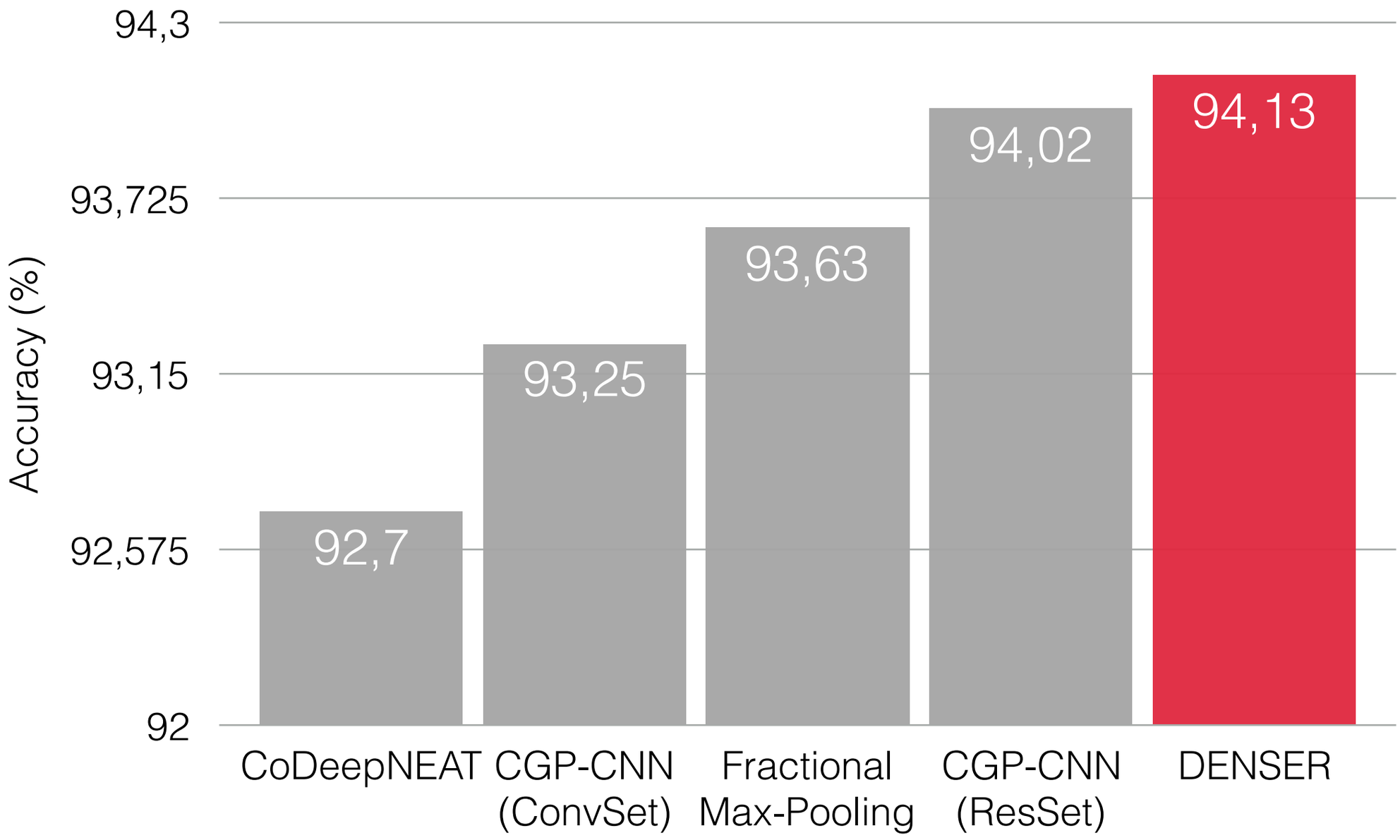
HINTON

DENSER BENCHMARKING

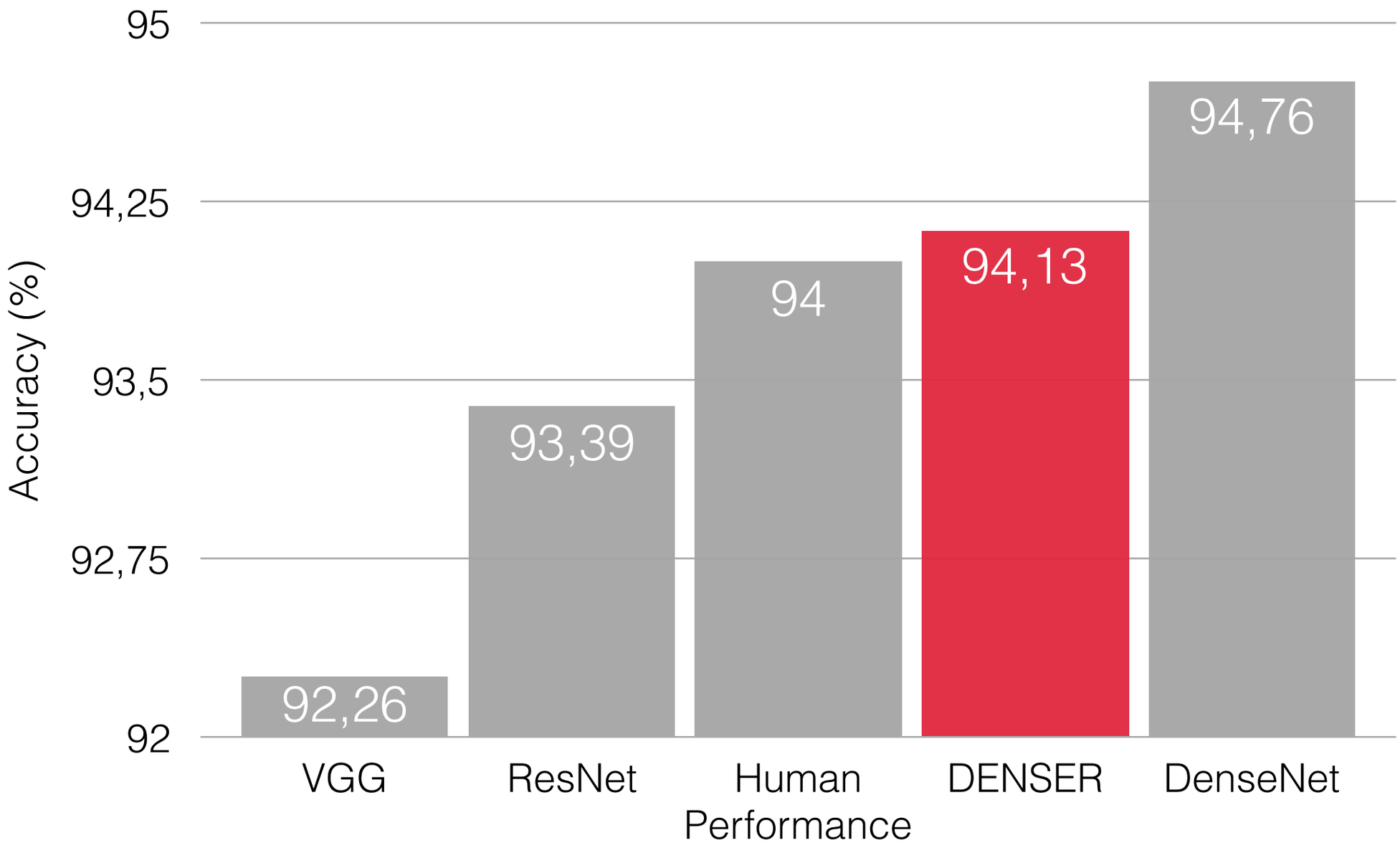




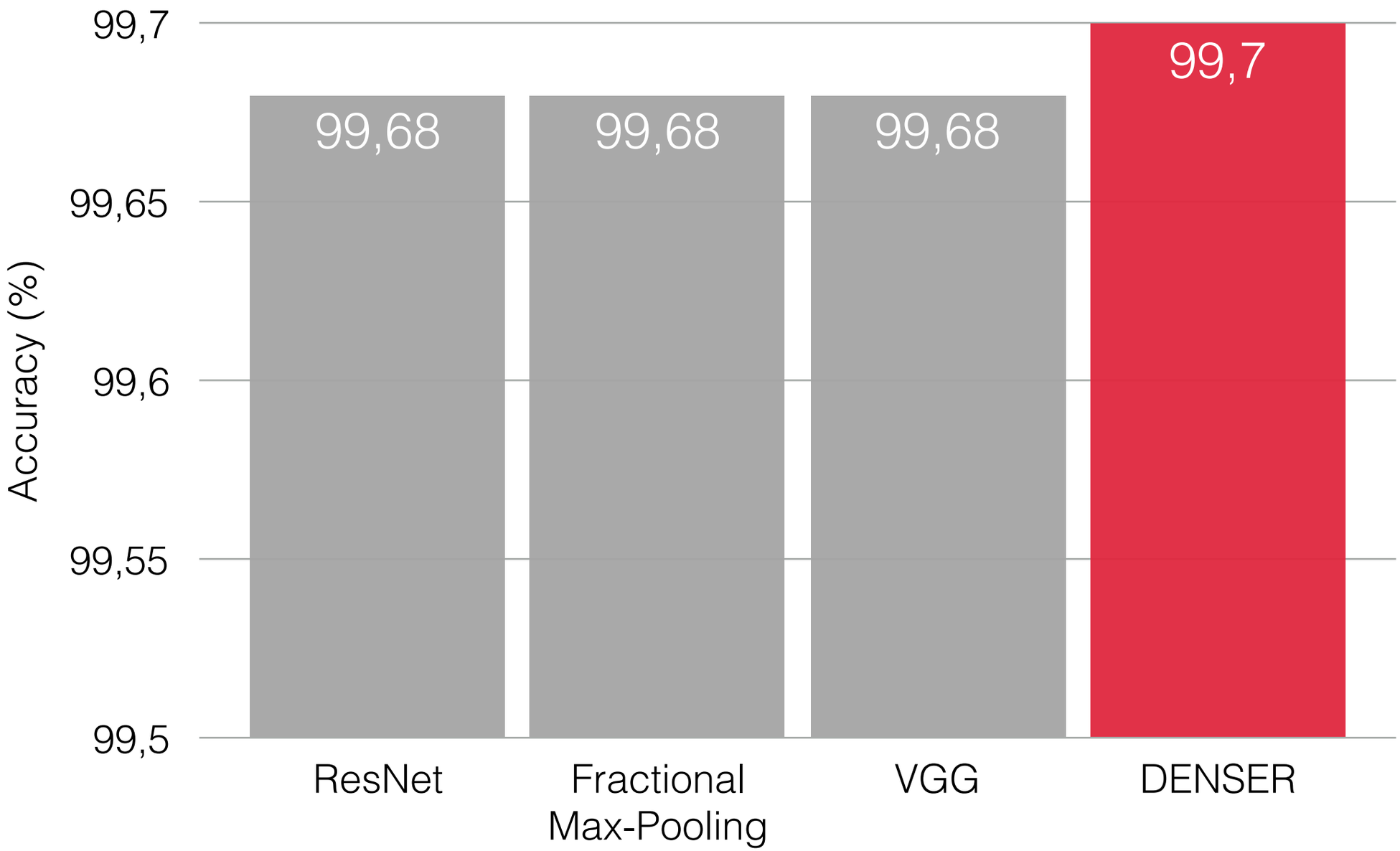
DENSER VS. OTHER AUTOMATIC DESIGN METHODS (CIFAR-10)



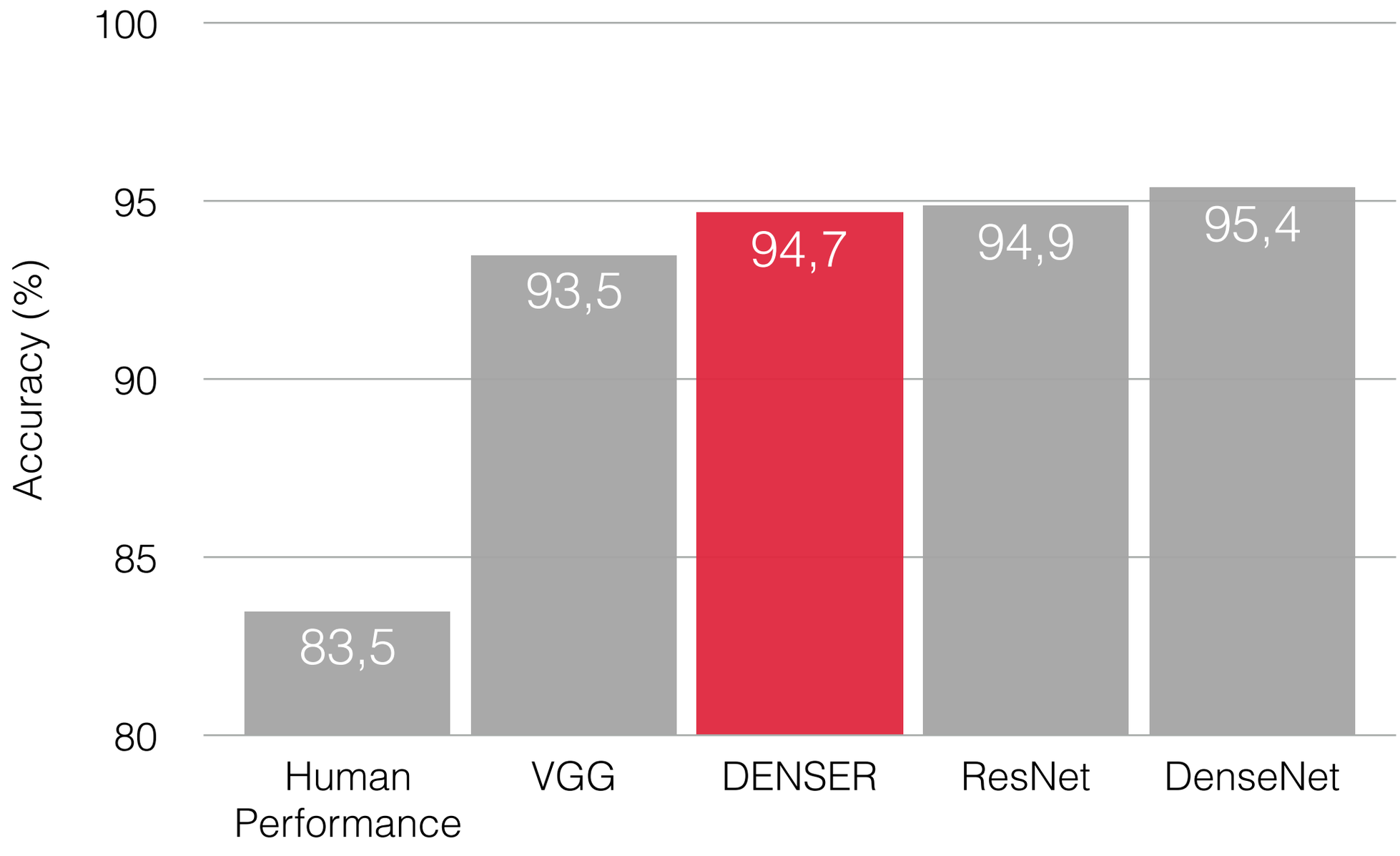
DENSER VS. HUMAN-DESIGNED NETWORKS (CIFAR-10)



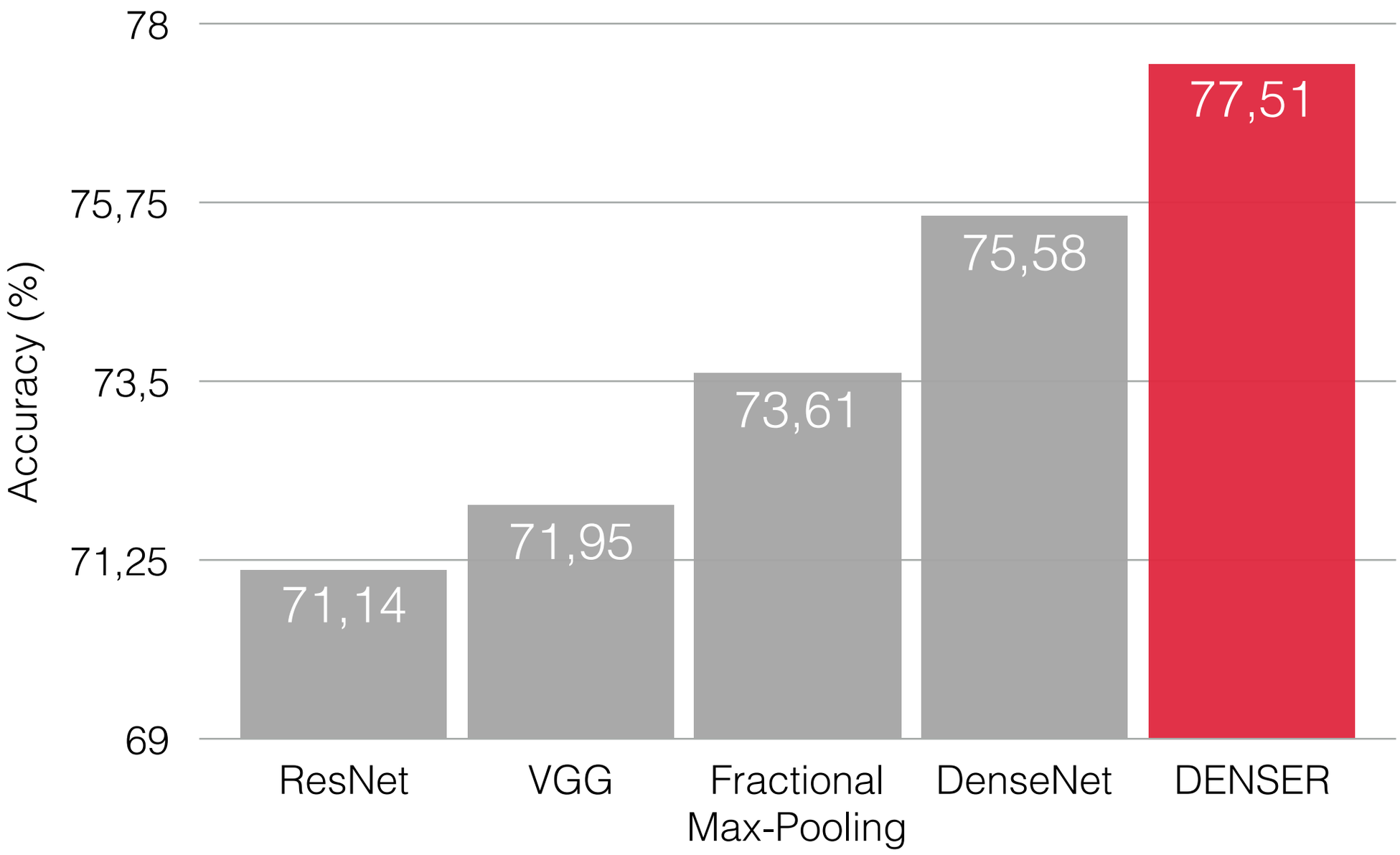
DENSER VS. HUMAN-DESIGNED NETWORKS (MNIST)



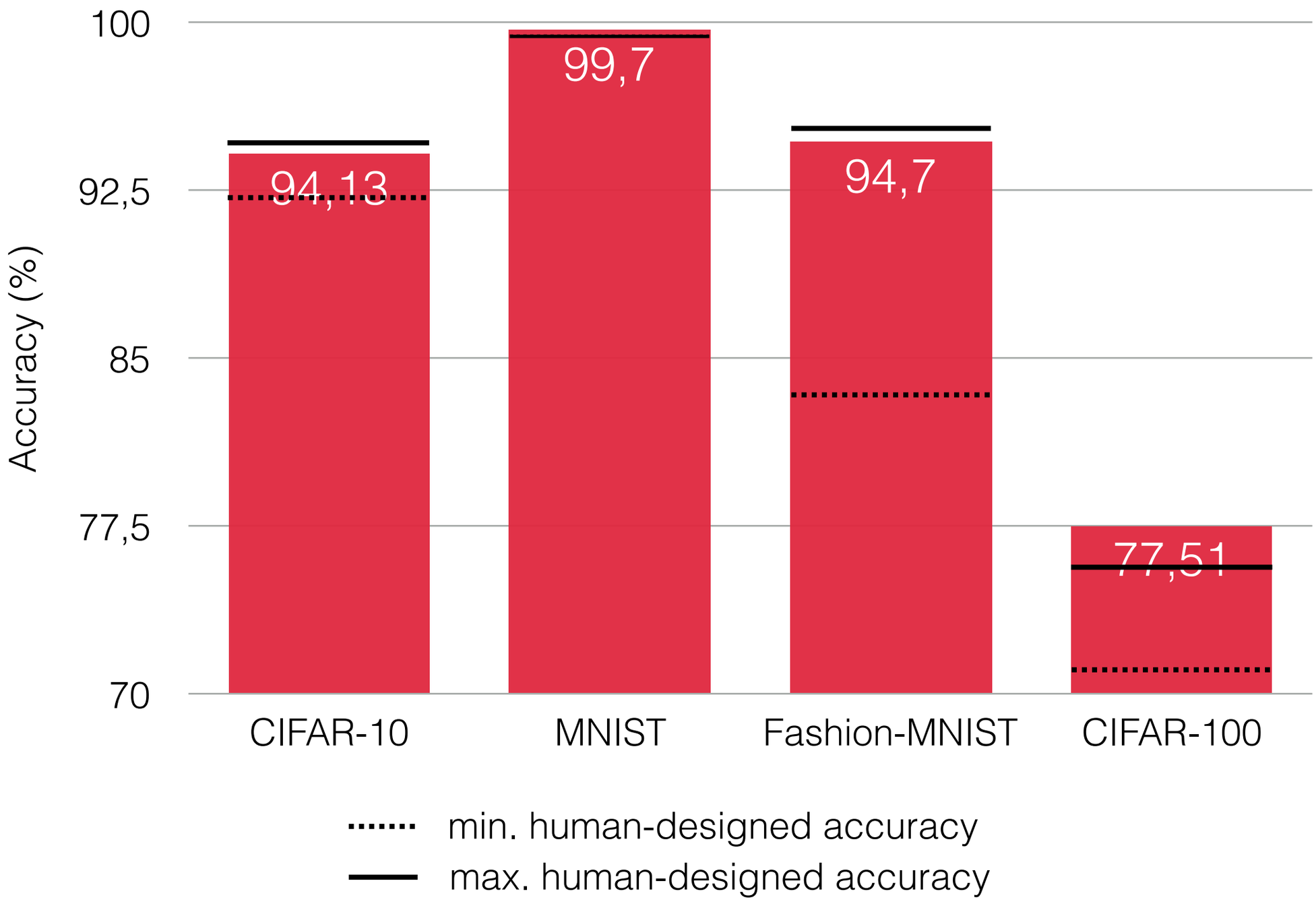
DENSER VS. HUMAN-DESIGNED NETWORKS (FASHION-MNIST)



DENSER VS. HUMAN-DESIGNED NETWORKS (CIFAR-100)



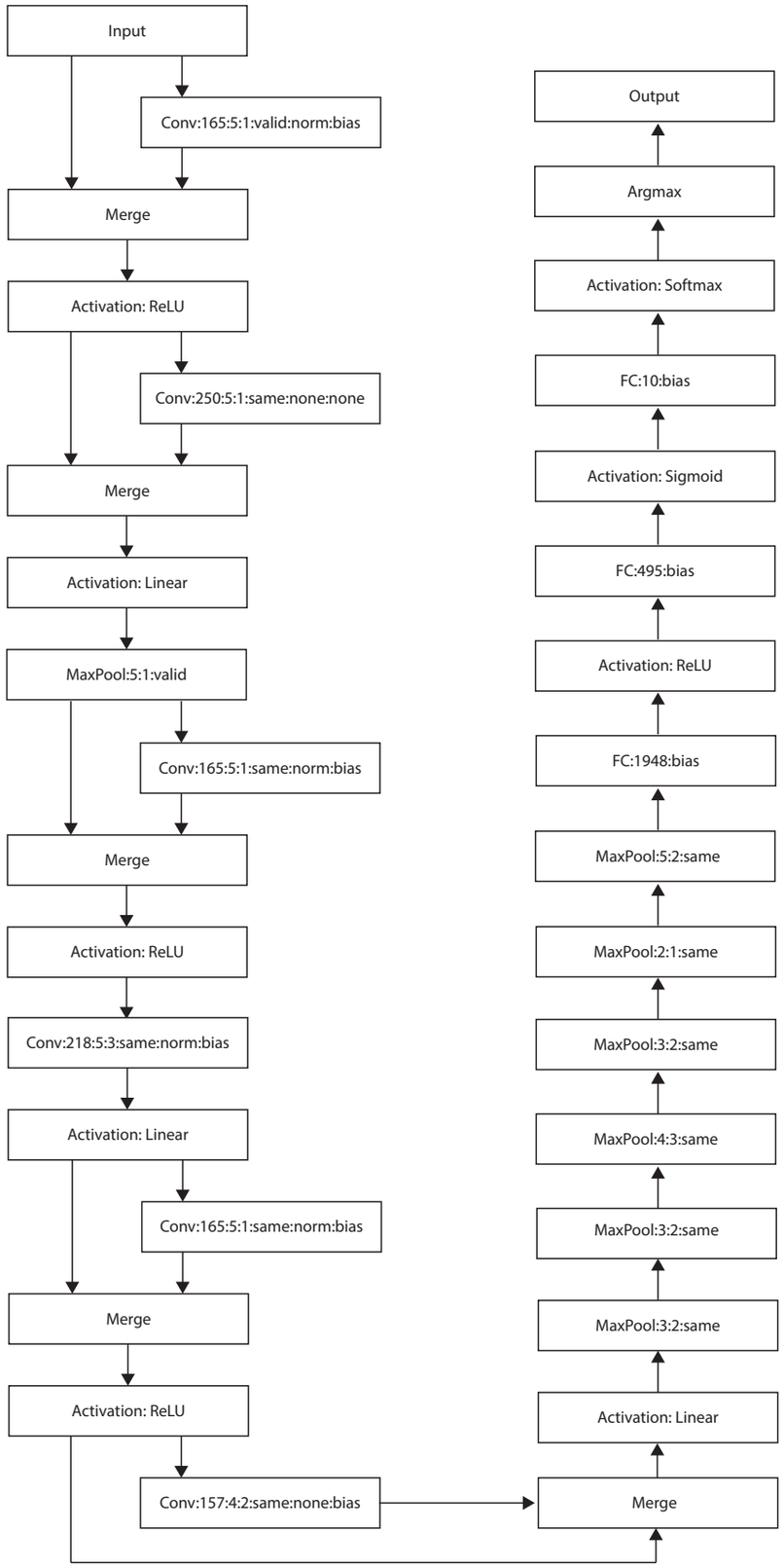
ROBUSTNESS, GENERALISATION, SCALABILITY



WHY THE BEST ENTRY?

- General purpose-framework for the automatisation of the search of Deep Artificial Neural Networks (DANNs);
- Results show that, without any prior-knowledge, DENSER can effectively discover (and even surpass) other automatic and human-designed DANNs;
- The CIFAR-100 result defines a new state-of-the-art;
- The evolved DANNs have proven to be robust, generalisable, and scalable;
- Low cost evolutionary ML approach.

WHY THE BEST ENTRY?



DENSER: Deep Evolutionary Network Structured Representation

cdv.dei.uc.pt/denser

Filipe Assunção, Nuno Lourenço, Penousal Machado and Bernardete Ribeiro
University of Coimbra, Coimbra, Portugal
{fga, naml, machado, bribeiro}@dei.uc.pt