I chose one hidden layer of 16 nodes and an output layer of 10 nodes, in my test results it seems (unlike what you would expect because of over fitting) that more nodes in one hidden layer always means better validation results. However the additional nodes seem to have less of an impact the more you add. That is why I chose to use one layer of 16 nodes because this seems to be a good compromise between speed and accuracy.

output:

Train on 50000 samples, validate on 10000 samples

Epoch 1/5

32/50000 [..............................] - ETA: 2:31 - loss: 0.2462 - acc: 0.1562

1568/50000 [..............................] - ETA: 4s - loss: 0.0976 - acc: 0.2921

3040/50000 [>.............................] - ETA: 3s - loss: 0.0862 - acc: 0.4059

4448/50000 [=>............................] - ETA: 2s - loss: 0.0786 - acc: 0.4975

5952/50000 [==>...........................] - ETA: 2s - loss: 0.0716 - acc: 0.5679

7392/50000 [===>..........................] - ETA: 2s - loss: 0.0657 - acc: 0.6209

8960/50000 [====>.........................] - ETA: 1s - loss: 0.0604 - acc: 0.6590

10464/50000 [=====>........................] - ETA: 1s - loss: 0.0560 - acc: 0.6887

12160/50000 [======>.......................] - ETA: 1s - loss: 0.0521 - acc: 0.7132

13760/50000 [=======>......................] - ETA: 1s - loss: 0.0489 - acc: 0.7317

15136/50000 [========>.....................] - ETA: 1s - loss: 0.0466 - acc: 0.7449

16800/50000 [=========>....................] - ETA: 1s - loss: 0.0441 - acc: 0.7592

18496/50000 [==========>...................] - ETA: 1s - loss: 0.0420 - acc: 0.7704

20224/50000 [===========>..................] - ETA: 1s - loss: 0.0401 - acc: 0.7810

21696/50000 [============>.................] - ETA: 1s - loss: 0.0387 - acc: 0.7887

23328/50000 [============>.................] - ETA: 0s - loss: 0.0373 - acc: 0.7961

24672/50000 [=============>................] - ETA: 0s - loss: 0.0362 - acc: 0.8016

26272/50000 [==============>...............] - ETA: 0s - loss: 0.0352 - acc: 0.8066

27808/50000 [===============>..............] - ETA: 0s - loss: 0.0344 - acc: 0.8107

29216/50000 [================>.............] - ETA: 0s - loss: 0.0336 - acc: 0.8149

30272/50000 [=================>............] - ETA: 0s - loss: 0.0331 - acc: 0.8177

31488/50000 [=================>............] - ETA: 0s - loss: 0.0324 - acc: 0.8213

33024/50000 [==================>...........] - ETA: 0s - loss: 0.0317 - acc: 0.8249

34560/50000 [===================>..........] - ETA: 0s - loss: 0.0310 - acc: 0.8280

35968/50000 [====================>.........] - ETA: 0s - loss: 0.0305 - acc: 0.8308

37632/50000 [=====================>........] - ETA: 0s - loss: 0.0299 - acc: 0.8340

39200/50000 [======================>.......] - ETA: 0s - loss: 0.0293 - acc: 0.8367

40736/50000 [=======================>......] - ETA: 0s - loss: 0.0288 - acc: 0.8393

42336/50000 [========================>.....] - ETA: 0s - loss: 0.0283 - acc: 0.8423

44128/50000 [=========================>....] - ETA: 0s - loss: 0.0277 - acc: 0.8450

45568/50000 [==========================>...] - ETA: 0s - loss: 0.0274 - acc: 0.8464

47264/50000 [===========================>..] - ETA: 0s - loss: 0.0269 - acc: 0.8487

48608/50000 [============================>.] - ETA: 0s - loss: 0.0266 - acc: 0.8503

50000/50000 [==============================] - 2s 40us/sample - loss: 0.0263 - acc: 0.8517 - val\_loss: 0.0133 - val\_acc: 0.9188

Epoch 2/5

32/50000 [..............................] - ETA: 3s - loss: 0.0069 - acc: 0.9688

1536/50000 [..............................] - ETA: 1s - loss: 0.0143 - acc: 0.9115

2912/50000 [>.............................] - ETA: 1s - loss: 0.0150 - acc: 0.9069

4448/50000 [=>............................] - ETA: 1s - loss: 0.0148 - acc: 0.9098

5856/50000 [==>...........................] - ETA: 1s - loss: 0.0150 - acc: 0.9073

7072/50000 [===>..........................] - ETA: 1s - loss: 0.0149 - acc: 0.9089

8640/50000 [====>.........................] - ETA: 1s - loss: 0.0147 - acc: 0.9104

10368/50000 [=====>........................] - ETA: 1s - loss: 0.0145 - acc: 0.9119

11904/50000 [======>.......................] - ETA: 1s - loss: 0.0145 - acc: 0.9115

13472/50000 [=======>......................] - ETA: 1s - loss: 0.0145 - acc: 0.9117

15040/50000 [========>.....................] - ETA: 1s - loss: 0.0144 - acc: 0.9117

16352/50000 [========>.....................] - ETA: 1s - loss: 0.0145 - acc: 0.9107

17920/50000 [=========>....................] - ETA: 1s - loss: 0.0145 - acc: 0.9109

19488/50000 [==========>...................] - ETA: 1s - loss: 0.0144 - acc: 0.9116

21056/50000 [===========>..................] - ETA: 0s - loss: 0.0144 - acc: 0.9115

22592/50000 [============>.................] - ETA: 0s - loss: 0.0143 - acc: 0.9120

23936/50000 [=============>................] - ETA: 0s - loss: 0.0143 - acc: 0.9123

25536/50000 [==============>...............] - ETA: 0s - loss: 0.0143 - acc: 0.9124

26976/50000 [===============>..............] - ETA: 0s - loss: 0.0143 - acc: 0.9120

28544/50000 [================>.............] - ETA: 0s - loss: 0.0142 - acc: 0.9127

30112/50000 [=================>............] - ETA: 0s - loss: 0.0142 - acc: 0.9126

31680/50000 [==================>...........] - ETA: 0s - loss: 0.0143 - acc: 0.9125

33312/50000 [==================>...........] - ETA: 0s - loss: 0.0142 - acc: 0.9126

34944/50000 [===================>..........] - ETA: 0s - loss: 0.0142 - acc: 0.9131

36576/50000 [====================>.........] - ETA: 0s - loss: 0.0141 - acc: 0.9133

38016/50000 [=====================>........] - ETA: 0s - loss: 0.0141 - acc: 0.9139

39360/50000 [======================>.......] - ETA: 0s - loss: 0.0141 - acc: 0.9140

40704/50000 [=======================>......] - ETA: 0s - loss: 0.0140 - acc: 0.9148

42304/50000 [========================>.....] - ETA: 0s - loss: 0.0139 - acc: 0.9149

43936/50000 [=========================>....] - ETA: 0s - loss: 0.0140 - acc: 0.9150

45408/50000 [==========================>...] - ETA: 0s - loss: 0.0140 - acc: 0.9149

47136/50000 [===========================>..] - ETA: 0s - loss: 0.0139 - acc: 0.9152

48480/50000 [============================>.] - ETA: 0s - loss: 0.0139 - acc: 0.9155

50000/50000 [==============================] - 2s 38us/sample - loss: 0.0138 - acc: 0.9157 - val\_loss: 0.0120 - val\_acc: 0.9258

Epoch 3/5

32/50000 [..............................] - ETA: 4s - loss: 0.0041 - acc: 1.0000

1664/50000 [..............................] - ETA: 1s - loss: 0.0123 - acc: 0.9285

3232/50000 [>.............................] - ETA: 1s - loss: 0.0118 - acc: 0.9319

4768/50000 [=>............................] - ETA: 1s - loss: 0.0122 - acc: 0.9276

6304/50000 [==>...........................] - ETA: 1s - loss: 0.0120 - acc: 0.9281

7584/50000 [===>..........................] - ETA: 1s - loss: 0.0123 - acc: 0.9252

8800/50000 [====>.........................] - ETA: 1s - loss: 0.0124 - acc: 0.9233

10432/50000 [=====>........................] - ETA: 1s - loss: 0.0123 - acc: 0.9245

12000/50000 [======>.......................] - ETA: 1s - loss: 0.0125 - acc: 0.9236

13632/50000 [=======>......................] - ETA: 1s - loss: 0.0124 - acc: 0.9244

15040/50000 [========>.....................] - ETA: 1s - loss: 0.0125 - acc: 0.9239

16256/50000 [========>.....................] - ETA: 1s - loss: 0.0125 - acc: 0.9237

17888/50000 [=========>....................] - ETA: 1s - loss: 0.0126 - acc: 0.9232

19264/50000 [==========>...................] - ETA: 1s - loss: 0.0126 - acc: 0.9237

20800/50000 [===========>..................] - ETA: 1s - loss: 0.0125 - acc: 0.9244

22336/50000 [============>.................] - ETA: 0s - loss: 0.0124 - acc: 0.9250

23776/50000 [=============>................] - ETA: 0s - loss: 0.0125 - acc: 0.9246

25408/50000 [==============>...............] - ETA: 0s - loss: 0.0125 - acc: 0.9244

27072/50000 [===============>..............] - ETA: 0s - loss: 0.0124 - acc: 0.9243

28608/50000 [================>.............] - ETA: 0s - loss: 0.0125 - acc: 0.9237

30208/50000 [=================>............] - ETA: 0s - loss: 0.0124 - acc: 0.9239

31712/50000 [==================>...........] - ETA: 0s - loss: 0.0125 - acc: 0.9234

33152/50000 [==================>...........] - ETA: 0s - loss: 0.0125 - acc: 0.9235

34720/50000 [===================>..........] - ETA: 0s - loss: 0.0125 - acc: 0.9236

36128/50000 [====================>.........] - ETA: 0s - loss: 0.0124 - acc: 0.9237

37824/50000 [=====================>........] - ETA: 0s - loss: 0.0124 - acc: 0.9234

39392/50000 [======================>.......] - ETA: 0s - loss: 0.0125 - acc: 0.9227

40960/50000 [=======================>......] - ETA: 0s - loss: 0.0125 - acc: 0.9225

42592/50000 [========================>.....] - ETA: 0s - loss: 0.0125 - acc: 0.9226

44128/50000 [=========================>....] - ETA: 0s - loss: 0.0125 - acc: 0.9229

45664/50000 [==========================>...] - ETA: 0s - loss: 0.0124 - acc: 0.9232

47008/50000 [===========================>..] - ETA: 0s - loss: 0.0124 - acc: 0.9231

47904/50000 [===========================>..] - ETA: 0s - loss: 0.0124 - acc: 0.9232

49184/50000 [============================>.] - ETA: 0s - loss: 0.0124 - acc: 0.9232

50000/50000 [==============================] - 2s 38us/sample - loss: 0.0124 - acc: 0.9230 - val\_loss: 0.0115 - val\_acc: 0.9266

Epoch 4/5

32/50000 [..............................] - ETA: 3s - loss: 0.0134 - acc: 0.9062

1568/50000 [..............................] - ETA: 1s - loss: 0.0116 - acc: 0.9267

2848/50000 [>.............................] - ETA: 1s - loss: 0.0111 - acc: 0.9312

4224/50000 [=>............................] - ETA: 1s - loss: 0.0114 - acc: 0.9306

5568/50000 [==>...........................] - ETA: 1s - loss: 0.0115 - acc: 0.9289

7040/50000 [===>..........................] - ETA: 1s - loss: 0.0112 - acc: 0.9308

8352/50000 [====>.........................] - ETA: 1s - loss: 0.0110 - acc: 0.9322

9184/50000 [====>.........................] - ETA: 1s - loss: 0.0111 - acc: 0.9309

9952/50000 [====>.........................] - ETA: 1s - loss: 0.0113 - acc: 0.9301

11040/50000 [=====>........................] - ETA: 1s - loss: 0.0113 - acc: 0.9303

12480/50000 [======>.......................] - ETA: 1s - loss: 0.0114 - acc: 0.9307

14176/50000 [=======>......................] - ETA: 1s - loss: 0.0114 - acc: 0.9302

15840/50000 [========>.....................] - ETA: 1s - loss: 0.0115 - acc: 0.9295

17472/50000 [=========>....................] - ETA: 1s - loss: 0.0114 - acc: 0.9298

19104/50000 [==========>...................] - ETA: 1s - loss: 0.0115 - acc: 0.9294

19936/50000 [==========>...................] - ETA: 1s - loss: 0.0114 - acc: 0.9294

21152/50000 [===========>..................] - ETA: 1s - loss: 0.0114 - acc: 0.9299

22688/50000 [============>.................] - ETA: 1s - loss: 0.0114 - acc: 0.9294

24128/50000 [=============>................] - ETA: 0s - loss: 0.0114 - acc: 0.9295

25696/50000 [==============>...............] - ETA: 0s - loss: 0.0115 - acc: 0.9293

27232/50000 [===============>..............] - ETA: 0s - loss: 0.0114 - acc: 0.9297

28736/50000 [================>.............] - ETA: 0s - loss: 0.0114 - acc: 0.9299

30176/50000 [=================>............] - ETA: 0s - loss: 0.0114 - acc: 0.9296

31520/50000 [=================>............] - ETA: 0s - loss: 0.0115 - acc: 0.9290

33120/50000 [==================>...........] - ETA: 0s - loss: 0.0116 - acc: 0.9284

34656/50000 [===================>..........] - ETA: 0s - loss: 0.0115 - acc: 0.9289

36224/50000 [====================>.........] - ETA: 0s - loss: 0.0115 - acc: 0.9286

37760/50000 [=====================>........] - ETA: 0s - loss: 0.0116 - acc: 0.9278

39328/50000 [======================>.......] - ETA: 0s - loss: 0.0117 - acc: 0.9273

40800/50000 [=======================>......] - ETA: 0s - loss: 0.0117 - acc: 0.9271

42496/50000 [========================>.....] - ETA: 0s - loss: 0.0117 - acc: 0.9273

44032/50000 [=========================>....] - ETA: 0s - loss: 0.0117 - acc: 0.9271

45728/50000 [==========================>...] - ETA: 0s - loss: 0.0118 - acc: 0.9268

47360/50000 [===========================>..] - ETA: 0s - loss: 0.0118 - acc: 0.9267

48704/50000 [============================>.] - ETA: 0s - loss: 0.0117 - acc: 0.9268

50000/50000 [==============================] - 2s 40us/sample - loss: 0.0118 - acc: 0.9267 - val\_loss: 0.0112 - val\_acc: 0.9285

Epoch 5/5

32/50000 [..............................] - ETA: 3s - loss: 0.0205 - acc: 0.8750

1600/50000 [..............................] - ETA: 1s - loss: 0.0118 - acc: 0.9281

3072/50000 [>.............................] - ETA: 1s - loss: 0.0116 - acc: 0.9303

4672/50000 [=>............................] - ETA: 1s - loss: 0.0115 - acc: 0.9311

6272/50000 [==>...........................] - ETA: 1s - loss: 0.0116 - acc: 0.9297

8000/50000 [===>..........................] - ETA: 1s - loss: 0.0114 - acc: 0.9304

9376/50000 [====>.........................] - ETA: 1s - loss: 0.0114 - acc: 0.9311

11008/50000 [=====>........................] - ETA: 1s - loss: 0.0111 - acc: 0.9331

12448/50000 [======>.......................] - ETA: 1s - loss: 0.0110 - acc: 0.9340

14016/50000 [=======>......................] - ETA: 1s - loss: 0.0112 - acc: 0.9326

15648/50000 [========>.....................] - ETA: 1s - loss: 0.0112 - acc: 0.9329

16992/50000 [=========>....................] - ETA: 1s - loss: 0.0112 - acc: 0.9328

18528/50000 [==========>...................] - ETA: 1s - loss: 0.0112 - acc: 0.9327

20064/50000 [===========>..................] - ETA: 0s - loss: 0.0113 - acc: 0.9321

21440/50000 [===========>..................] - ETA: 0s - loss: 0.0113 - acc: 0.9322

23072/50000 [============>.................] - ETA: 0s - loss: 0.0112 - acc: 0.9328

24480/50000 [=============>................] - ETA: 0s - loss: 0.0112 - acc: 0.9328

26016/50000 [==============>...............] - ETA: 0s - loss: 0.0111 - acc: 0.9330

27232/50000 [===============>..............] - ETA: 0s - loss: 0.0111 - acc: 0.9331

28352/50000 [================>.............] - ETA: 0s - loss: 0.0111 - acc: 0.9327

30048/50000 [=================>............] - ETA: 0s - loss: 0.0111 - acc: 0.9329

31680/50000 [==================>...........] - ETA: 0s - loss: 0.0111 - acc: 0.9332

33120/50000 [==================>...........] - ETA: 0s - loss: 0.0111 - acc: 0.9329

34720/50000 [===================>..........] - ETA: 0s - loss: 0.0112 - acc: 0.9324

36224/50000 [====================>.........] - ETA: 0s - loss: 0.0112 - acc: 0.9318

37696/50000 [=====================>........] - ETA: 0s - loss: 0.0113 - acc: 0.9316

39264/50000 [======================>.......] - ETA: 0s - loss: 0.0113 - acc: 0.9314

40704/50000 [=======================>......] - ETA: 0s - loss: 0.0113 - acc: 0.9314

42272/50000 [========================>.....] - ETA: 0s - loss: 0.0113 - acc: 0.9314

43840/50000 [=========================>....] - ETA: 0s - loss: 0.0113 - acc: 0.9313

45312/50000 [==========================>...] - ETA: 0s - loss: 0.0113 - acc: 0.9311

46944/50000 [===========================>..] - ETA: 0s - loss: 0.0113 - acc: 0.9308

48416/50000 [============================>.] - ETA: 0s - loss: 0.0113 - acc: 0.9308

50000/50000 [==============================] - 2s 38us/sample - loss: 0.0113 - acc: 0.9306 - val\_loss: 0.0111 - val\_acc: 0.9318

Model: "sequential"

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

dense (Dense) multiple 12560

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) multiple 170

=================================================================

Total params: 12,730

Trainable params: 12,730

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_