

CNN of KMNIST

By Geoffrey Hughes and Nic Fergie

欽察從迹如
嘉嬰擒廟私
鷄玉瑞真傳
坡替羽位致
盤妻之從德
外政仍聘雌
務業夫業菲
得身答拒獲
杖書粹快果
招恣幣外婚
喧訖熱羹飽
勿牆久淚淋
枕喉咽斗指
嘴般冰凍竊
矣哭為害病
將車尔奔炭
情款急標粗
乞賴極弑綠
蹄單鯁凝麓
勃夾却進多
急文權威鵬
罰咭隣名獨
據諸稜積甄
晉輝服血通
仇沒襄夷浣
敢牛耐悔羞
允少涿意蒼
苦屈札了黃
治資側布以

Dataset

Handwritten kanji characters



Can we classify
handwritten
Japanese
characters with
a high degree of
accuracy?

True:4 Pred:2



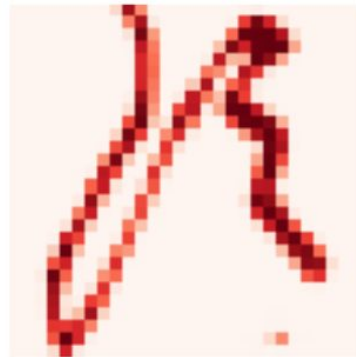
True:2 Pred:4



True:3 Pred:2

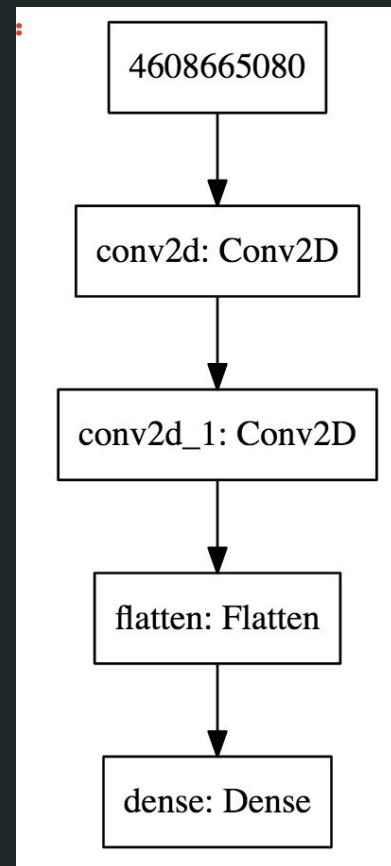


True:2 Pred:8



The Model

We've used a CNN to try and classify the data. It is a two layer network with 64 and then 32 nodes.



Early Results

91% across precision and recall
with the test data

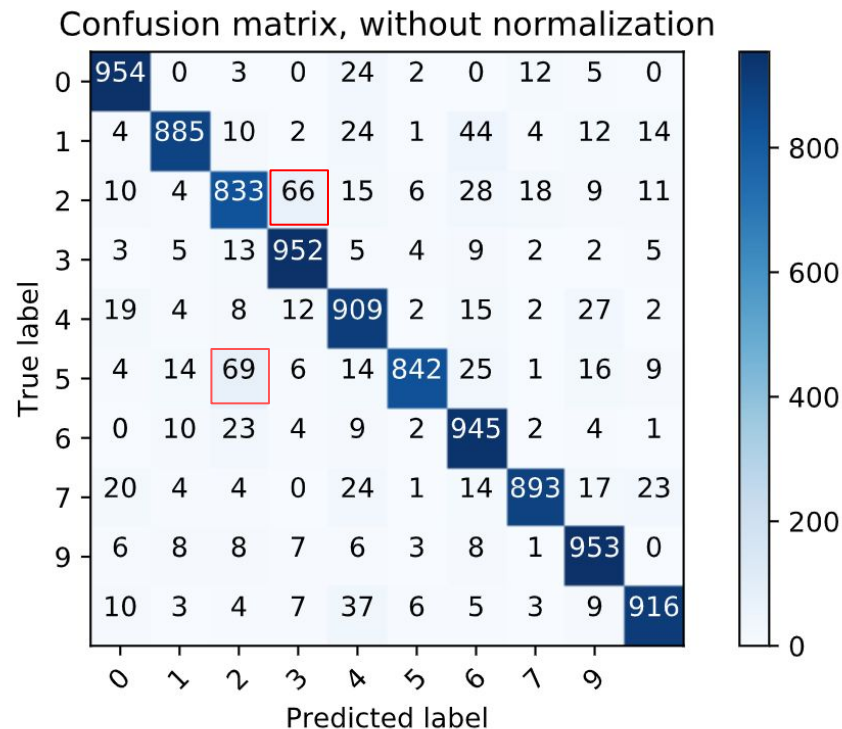
Correct predicted classes: 9141

Incorrect predicted classes: 859

	precision	recall	f1-score	support
Class 0 (お):	0.95	0.93	0.94	1000
Class 1 (き):	0.89	0.93	0.91	1000
Class 2 (す):	0.86	0.85	0.86	1000
Class 3 (つ):	0.94	0.92	0.93	1000
Class 4 (な):	0.90	0.90	0.90	1000
Class 5 (は):	0.94	0.89	0.92	1000
Class 6 (ま):	0.90	0.94	0.92	1000
Class 7 (や):	0.91	0.91	0.91	1000
Class 8 (れ):	0.91	0.95	0.93	1000
Class 9 (を):	0.94	0.93	0.94	1000
avg / total	0.91	0.91	0.91	10000

Our Model is Confused

The highest incidences of confusion are occurring with the label 2 being incorrectly assigned to 5 and the label 3 being incorrectly assigned to 2.



Are we overfitting?

