

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ

ΣΥΣΤΗΜΑΤΑ ΕΜΠΙΝΕΥΣΜΕΝΑ ΑΠΟ ΤΗ ΒΙΟΛΟΓΙΑ

MARIA BPANA · mvrana.96@outlook.com
ΠΡΟΗΓΜΕΝΑ ΣΥΣΤΗΜΑΤΑ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

Περιεχόμενα

1	Εισαγωγή: περιγραφή συνόλου δεδομένων και στόχου	1
2	Το μοντέλο και παραλλαγές	1
2.1	Τα δεδομένα	1
2.2	Το μοντέλο	2
2.2.1	Παραλλαγή 1: Softmax χωρίς ομαλοποίηση	2
2.2.2	Παραλλαγή 2: Softmax με ομαλοποίηση	3
2.2.3	Παραλλαγή 3: Sigmoid χωρίς ομαλοποίηση	4
2.2.4	Παραλλαγή 4: Sigmoid με ομαλοποίηση	4
3	Αποτελέσματα	5
3.1	Training - validation scores	5
3.2	Testing scores - Classification reports	6
4	Σύνοψη	9

1 Εισαγωγή: περιγραφή συνόλου δεδομένων και στόχου

Στόχος της εργασίας είναι η δημιουργία ενός νευρωνικού δικτύου για ταξινόμηση εικόνων (πινακίδες οδοσήμανσης). Το σύνολο δεδομένων αποτελείται από εικόνες 20 κλάσεων: limit30, limit50, limit60, limit80, noleft, noright, nouturn, gostraight, goleft, goright, keepright, roundabout, danger, pedestrcross, bikecross, childcross, traincross, construction, nostopping, noentry. Παρουσιάζεται αναλυτικά η προεπεξεργασία των δεδομένων, η δημιουργία κάθε στρώματος των νευρωνικών δικτύων και τα αποτελέσματα σε classification reports. Επιπλέον, παρέχονται σχόλια για τις δυσκολίες που αντιμετώπισαν τα μοντέλα στην κατάταξη των δειγμάτων και τα τελικά συμπεράσματα από την σύγκριση των διαφορετικών μοντέλων.

2 Το μοντέλο και παραλλαγές

2.1 Τα δεδομένα

Για την αποτελεσματικότερη εκπαίδευση του αλγορίθμου, έγιναν τροποποιήσεις στα δεδομένα με χρήση του `tf.keras.preprocessing.image.ImageDataGenerator`. Οι παράμετροι που διαφοροποιήθηκαν από τις default ρυθμίσεις είναι οι παρακάτω:

- `shear_range=0.2`,

- `zoom_range=0.2`,
- `brightness_range=[0.4,0.9]`

και επιπλέον σε αυτό το βήμα διαχωρίστηκε το training set σε training (70%) και validation set (30%). Στη συνέχεια, έγινε real-time augmentation κατά την εκπαίδευση του νευρωνικού.

Κατά την προεπεξεργασία του συνόλου δεδομένων χρειάστηκε να γίνει one hot encoding ώστε να δουλεύουν σωστά οι activation functions (η softmax είχε καλά αποτελέσματα και χωρίς one hot encoding όμως η sigmoid δεν μπορούσε να λειτουργήσει σωστά). Αυτό πραγματοποιήθηκε με χρήση της [tf.keras.utils.to_categorical](#).

2.2 Το μοντέλο

Η ομάδα μοντέλων που επιλέχθηκε αποτελείται από 4 convolutional layers με kernel 5×5 και 2 dense layers με activation function: 'relu' για τα ενδιάμεσα layers (για τα convolutional και τα dense). Ο λόγος που επιλέχθηκε αυτή η ομάδα είναι ότι το μεγαλύτερο kernel καταλήγει σε καλύτερα αποτελέσματα από την πιο δημοφιλή επιλογή kernel 3×3 . Μία ενδεικτική σύγκριση για μοντέλα με ίδια χαρακτηριστικά και μόνη διαφορά το kernel size σε όλα τα convolutional layers είναι η παρακάτω:

- kernel size: 3×3 , activation function: sigmoid, accuracy: 78.08%, loss= 1.34
- kernel size: 5×5 , activation function: sigmoid, accuracy: 84.15%, loss= 0.82
- kernel size: 3×3 , activation function: softmax, accuracy: 81.51%, loss= 1.17
- kernel size: 5×5 , activation function: softmax, accuracy: 89.43%, loss= 0.78

Έγιναν 4 παραλλαγές με kernel size 5×5 και αναλύονται στις παρακάτω υποενότητες.

2.2.1 Παραλλαγή 1: Softmax χωρίς ομαλοποίηση

Το μοντέλο αποτελείται από 3 convolutional layers με filters=32, kernel= 5×5 , activation='relu' ένα ακόμη convolutional layer με filters=64, kernel= 5×5 , activation='relu'. Ενδιάμεσα, εφαρμόζονται layers max pooling με poolsize=(2, 2) και στο τέλος ένα layer Dropout(0.25). Στη συνέχεια, για την εισαγωγή των δεδομένων στο dense layer πρέπει να γίνει flatten. Προστίθεται ένα dense layer με units=128, activation='relu', ένα ακόμη layer Dropout(0.5) για την αποφυγή του over-fitting και τελευταίο το dense layer με units=20, activation='softmax'. Ως μετρική έχει επιλεγθεί το accuracy ενώ ως κόστος η categorical cross-entropy. Το model summary είναι:

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 124, 124, 32)	2432
max_pooling2d (MaxPooling2D)	(None, 62, 62, 32)	0
conv2d_1 (Conv2D)	(None, 58, 58, 32)	25632
max_pooling2d_1 (MaxPooling2D)	(None, 29, 29, 32)	0
conv2d_2 (Conv2D)	(None, 25, 25, 32)	25632
max_pooling2d_2 (MaxPooling2D)	(None, 12, 12, 32)	0
conv2d_3 (Conv2D)	(None, 8, 8, 64)	51264
max_pooling2d_3 (MaxPooling2D)	(None, 4, 4, 64)	0
dropout (Dropout)	(None, 4, 4, 64)	0
flatten (Flatten)	(None, 1024)	0
dense (Dense)	(None, 128)	131200
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 20)	2580

=====
Total params: 238,740
Trainable params: 238,740
Non-trainable params: 0
=====

Σχήμα 1: Model Summary μοντέλου 1

2.2.2 Παραλλαγή 2: Softmax με ομαλοποίηση

Στην παραλλαγή αυτή, χρησιμοποιήθηκε ένα επιπλέον layer (συνολικά 4) ανάμεσα στα convolutional layers, το BatchNormalization ενώ το πρώτο dropout layer είναι Dropout(0.35). Το [tf.keras.layers.BatchNormalization](#) εφαρμόζει έναν μετασχηματισμό που διατηρεί το μέσο output κοντά στο 0 ενώ τη μέση απόκλιση κοντά στο 1. Το model summary είναι:

Model: "sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 124, 124, 32)	2432
batch_normalization (Batch Normalization)	(None, 124, 124, 32)	128
max_pooling2d (MaxPooling2D)	(None, 62, 62, 32)	0
conv2d_1 (Conv2D)	(None, 58, 58, 32)	25632
batch_normalization_1 (Batch Normalization)	(None, 58, 58, 32)	128
max_pooling2d_1 (MaxPooling2D)	(None, 29, 29, 32)	0
conv2d_2 (Conv2D)	(None, 25, 25, 32)	25632
batch_normalization_2 (Batch Normalization)	(None, 25, 25, 32)	128
max_pooling2d_2 (MaxPooling2D)	(None, 12, 12, 32)	0
conv2d_3 (Conv2D)	(None, 8, 8, 64)	51264
batch_normalization_3 (Batch Normalization)	(None, 8, 8, 64)	256
max_pooling2d_3 (MaxPooling2D)	(None, 4, 4, 64)	0
dropout (Dropout)	(None, 4, 4, 64)	0
flatten (Flatten)	(None, 1024)	0
dense (Dense)	(None, 128)	131200
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 20)	2580
Total params: 239,380		
Trainable params: 239,060		
Non-trainable params: 320		

Σχήμα 2: Model Summary μοντέλου 2

2.2.3 Παραλλαγή 3: Sigmoid χωρίς ομαλοποίηση

Η παραλλαγή αυτή, διαφέρει από την παραλλαγή 1 μόνο στην activation function του τελευταίου dense layer που σε αυτή την περίπτωση είναι sigmoid. Το model summary είναι ακριβώς ίδιο με του μοντέλου 1 [1].

2.2.4 Παραλλαγή 4: Sigmoid με ομαλοποίηση

Η παραλλαγή αυτή, διαφέρει από την παραλλαγή 2 μόνο στην activation function του τελευταίου dense layer που σε αυτή την περίπτωση είναι sigmoid. Το model summary είναι ακριβώς ίδιο με του μοντέλου 2 [2].

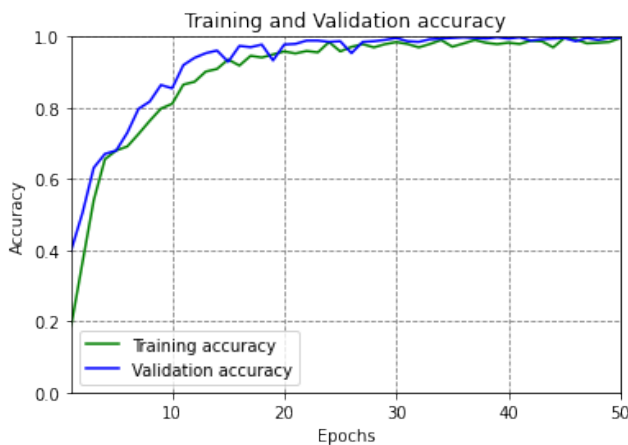
3 Αποτελέσματα

Σε αυτή την ενότητα θα συγκριθούν τα αποτελέσματα των 4 μοντέλων. Αρχικά έχει γίνει οπτικοποίηση της μετρικής απόδοσης και του κόστους για το training και validation set. Έπειτα, μετρήθηκε το accuracy και categorical cross-entropy στο test set. Τέλος, σχεδιάστηκε το classification report για κάθε μοντέλο. Στο [sklearn.metrics.classification_report](#) σχεδιάζεται ένας πίνακας ο οποίος περιέχει τα precision, recall, f1-score για κάθε κλάση και τέλος τα accuracy, macro avg, weighted avg.

Τα precision, recall και f1-score παίρνουν τιμές στο διάστημα $[0,1]$. Οι μετρικές αυτές θα είναι ίσες με 0 όταν δεν υπάρχει κανένα TRUE POSITIVE, δηλαδή για τη θετική κλάση έχει προβλεφθεί ότι δεν ανήκει κανένα δείγμα που όντως ανήκει στην θετική κλάση. Το precision θα είναι 1 αν δεν υπάρχει κανένα FALSE POSITIVE, δηλαδή κάποιο δείγμα που να έχει προβλεφθεί ως θετικό ενώ στην πραγματικότητα δεν είναι. Το recall θα είναι 1 όταν δεν υπάρχει κανένα FALSE NEGATIVE, δηλαδή κάποιο δείγμα που να έχει προβλεφθεί αρνητικό ενώ είναι θετικό. Το f1-score αποτελεί το συνδυασμό των precision και recall και υπολογίζεται από τη σχέση $F1 = 2 * (precision * recall) / (precision + recall)$.

3.1 Training - validation scores

Στα σχήματα [3],[4],[5],[6] φαίνονται τα accuracy και categorical cross-entropy για κάθε μοντέλο στο training και validation set ανά κύκλο.

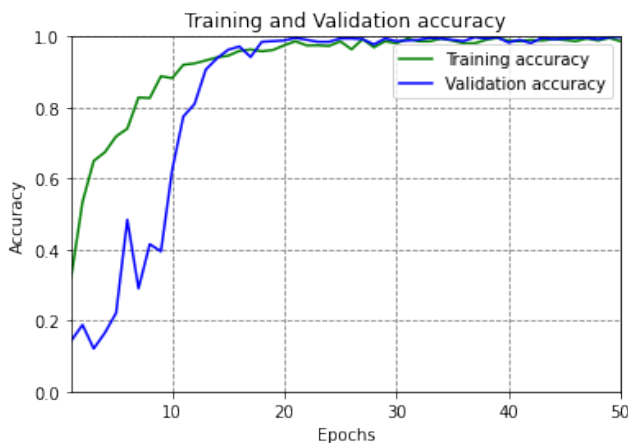


(α') Accuracy στο training και validation set.

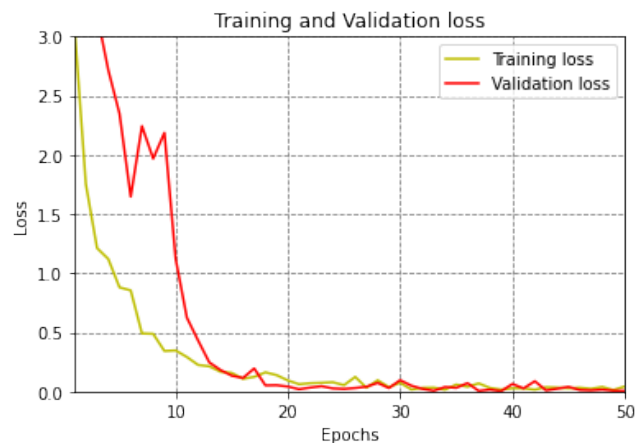


(β') Loss στο training και validation set.

Σχήμα 3: Accuracy και loss στο training και validation set του μοντέλου 1.

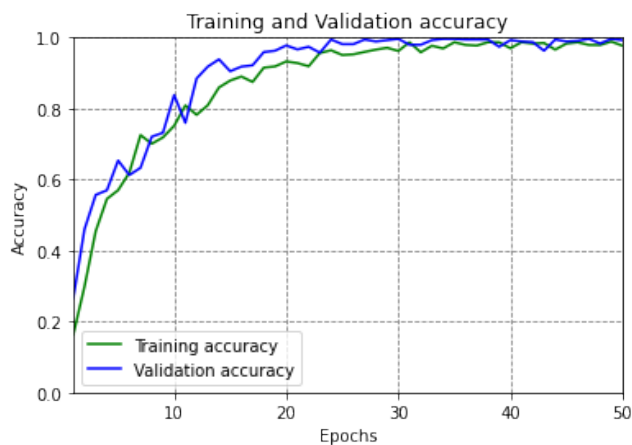


(α') Accuracy στο training και validation set.

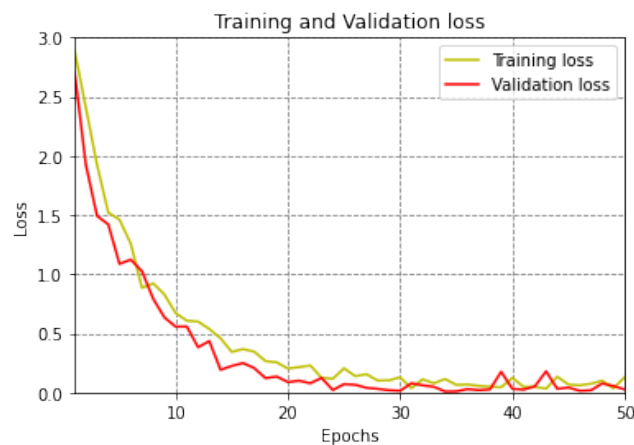


(β') Loss στο training και validation set.

Σχήμα 4: Accuracy και loss στο training και validation set του μοντέλου 2.



(α') Accuracy στο training και validation set.

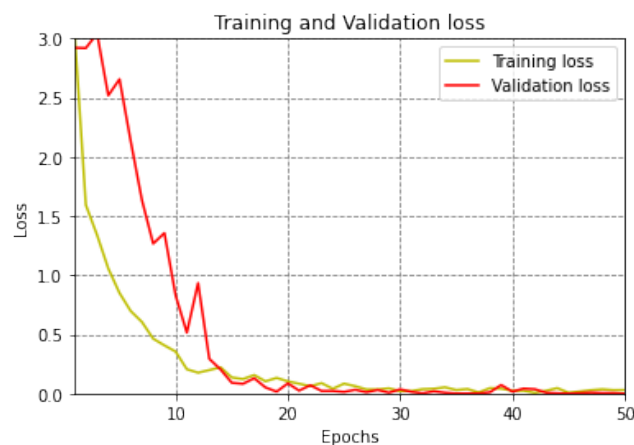


(β') Loss στο training και validation set.

Σχήμα 5: Accuracy και loss στο training και validation set του μοντέλου 3.



(α') Accuracy στο training και validation set.



(β') Loss στο training και validation set.

Σχήμα 6: Accuracy και loss στο training και validation set του μοντέλου 4.

Η χρήση του BatchNormalization φαίνεται να δημιουργεί αρχικά μία πιο ακανόνιστη συμπεριφορά του accuracy και του cross-entropy στο validation set κατά τους αρχικούς κύκλους εκπαίδευσης, όμως τελικά καταλήγει σε ομαλότερη συμπεριφορά στο training και το validation set (κατά τους τελευταίους κύκλους) από τα αντίστοιχα μοντέλα χωρίς την ομαλοποίηση.

3.2 Testing scores - Classification reports

Όπως φαίνεται από τις γραφικές παραστάσεις τα μοντέλα καταλήγουν να έχουν πολύ μεγάλο accuracy και πολύ μικρό loss στο training και το validation set. Η τελική αξιολόγηση όμως δε μπορεί να γίνει σε δεδομένα που το μοντέλο έχει συναντήσει κατά την εκπαίδευση του. Αναλυτικότερα, εφαρμόζοντας κάθε το μοντέλο στο test set προκύπτουν τα παρακάτω αποτελέσματα:

- μοντέλο 1 (softmax) \Rightarrow loss: 0.8742 - accuracy: 0.8806
- μοντέλο 2 (softmax με BatchNormalization) \Rightarrow loss: 0.5409 - accuracy: 0.9119
- μοντέλο 3 (sigmoid) \Rightarrow loss: 0.8225 - accuracy: 0.8415
- μοντέλο 4 (sigmoid με BatchNormalization) \Rightarrow loss: 0.6226 - accuracy: 0.8933

Εκ πρώτης όψεως τα αποτελέσματα είναι ικανοποιητικά, με αποτελεσματικότερο στην ταξινόμηση το μοντέλο 2. Είναι όμως σημαντικό να γίνει μελέτη σε μεγαλύτερο βάθος ώστε να βρεθεί ποιες κλάσεις

μπορούν τα μοντέλα να προβλέψουν αποτελεσματικά. Για αυτό το λόγο σχεδιάστηκαν τα classification reports. Στα σχήματα [7],[8],[9] και [10] φαίνονται τα classification reports για κάθε μοντέλο.

	precision	recall	f1-score	support
limit30	0.63	0.63	0.63	60
limit50	0.78	0.79	0.79	58
limit60	0.61	0.92	0.74	50
limit80	0.51	0.60	0.55	50
noleft	0.98	0.98	0.98	130
noright	0.99	0.87	0.92	92
nouturn	1.00	0.28	0.43	36
gostraight	1.00	1.00	1.00	12
goleft	0.89	1.00	0.94	8
goright	1.00	0.92	0.96	26
keepright	1.00	0.95	0.97	134
roundabout	0.92	0.92	0.92	24
danger	0.78	0.88	0.82	8
pedestrcross	1.00	0.91	0.95	46
bikecross	0.86	1.00	0.92	12
childcross	0.81	0.85	0.83	26
traincross	0.82	0.90	0.86	10
construction	0.43	1.00	0.60	6
nostopping	0.98	0.98	0.98	176
noentry	0.98	1.00	0.99	58
accuracy			0.88	1022
macro avg	0.85	0.87	0.84	1022
weighted avg	0.90	0.88	0.88	1022

Σχήμα 7: Classification report μοντέλου 1

	precision	recall	f1-score	support
limit30	0.79	0.87	0.83	60
limit50	0.93	0.88	0.90	58
limit60	0.60	0.94	0.73	50
limit80	0.65	0.72	0.69	50
noleft	0.96	0.98	0.97	130
noright	0.98	0.91	0.94	92
nouturn	0.60	0.08	0.15	36
gostraight	0.75	1.00	0.86	12
goleft	0.88	0.88	0.88	8
goright	1.00	0.81	0.89	26
keepright	1.00	0.90	0.94	134
roundabout	0.70	0.96	0.81	24
danger	1.00	0.75	0.86	8
pedestrcross	0.94	0.98	0.96	46
bikecross	0.75	1.00	0.86	12
childcross	1.00	0.88	0.94	26
traincross	1.00	0.90	0.95	10
construction	1.00	1.00	1.00	6
nostopping	0.97	0.98	0.97	176
noentry	0.98	1.00	0.99	58
accuracy			0.90	1022
macro avg	0.87	0.87	0.86	1022
weighted avg	0.90	0.90	0.89	1022

Σχήμα 8: Classification report μοντέλου 2

	precision	recall	f1-score	support
limit30	0.69	0.55	0.61	60
limit50	0.92	0.79	0.85	58
limit60	0.71	0.80	0.75	50
limit80	0.44	0.82	0.57	50
noleft	0.84	0.98	0.91	130
noright	0.93	0.55	0.69	92
nouturn	0.88	0.19	0.32	36
gostraight	1.00	1.00	1.00	12
goleft	1.00	1.00	1.00	8
goright	1.00	1.00	1.00	26
keepright	0.95	0.94	0.94	134
roundabout	0.81	0.92	0.86	24
danger	0.30	0.75	0.43	8
pedestrcross	0.90	0.96	0.93	46
bikecross	0.63	1.00	0.77	12
childcross	0.72	0.81	0.76	26
traincross	0.75	0.60	0.67	10
construction	0.55	1.00	0.71	6
nostopping	1.00	0.95	0.98	176
noentry	1.00	0.86	0.93	58
accuracy			0.83	1022
macro avg	0.80	0.82	0.78	1022
weighted avg	0.87	0.83	0.83	1022

Σχήμα 9: Classification report μοντέλου 3

	precision	recall	f1-score	support
limit30	0.84	0.72	0.77	60
limit50	0.91	0.90	0.90	58
limit60	0.76	0.88	0.81	50
limit80	0.64	0.78	0.70	50
noleft	0.98	0.97	0.98	130
noright	0.99	0.88	0.93	92
nouturn	1.00	0.36	0.53	36
gostraight	0.75	1.00	0.86	12
goleft	1.00	1.00	1.00	8
goright	0.91	0.77	0.83	26
keepright	0.99	0.95	0.97	134
roundabout	0.83	0.83	0.83	24
danger	0.54	0.88	0.67	8
pedestrcross	0.77	1.00	0.87	46
bikecross	0.73	0.67	0.70	12
childcross	0.56	0.88	0.69	26
traincross	1.00	0.90	0.95	10
construction	0.75	1.00	0.86	6
nostopping	0.99	0.98	0.99	176
noentry	1.00	1.00	1.00	58
accuracy			0.90	1022
macro avg	0.85	0.87	0.84	1022
weighted avg	0.91	0.90	0.89	1022

Σχήμα 10: Classification report μοντέλου 4

Αξίζει σε αυτό το σημείο να αναφερθεί ότι οι κλάσεις είναι unbalanced και αυτό δημιουργεί προβλήματα κατά την εκπαίδευση και την αξιολόγηση των μοντέλων. Ξεκινώντας από το μοντέλο 1 [7], ενώ οι κλάσεις

limit30, limit50, limit60, limit80 έχουν ικανοποιητικό αριθμό δειγμάτων στα υποσύνολα δεδομένων το precision είναι σχετικά χαμηλό, το ίδιο και το recall με εξαίρεση το limit60. Αυτό συμβαίνει καθώς οι πινακίδες αυτές έχουν πολλά ίδια χαρακτηριστικά οπότε ήταν αναμενόμενο. Το nouturn έχει πολύ χαμηλό recall 0.28 ενώ έχει precision 1, αυτό σημαίνει ότι πολλά δείγματα της κλάσης nouturn κατατάσσονται από το μοντέλο σε άλλες κλάσεις. Κάτι τέτοιο είναι επίσης λογικό, καθώς παρόμοιες πινακίδες διαφορετικής κλάσης έχουν πολλά περισσότερα δείγματα (μόνο 22/1792). Το construction με precision 0.48 είναι η κλάση με τα λιγότερα δείγματα στο σύνολο εκπαίδευσης (10/1792).

Στο μοντέλο 2 [8] οι πρώτες 4 κλάσεις έχουν σχετικά χαμηλά precision και recall όμως είναι καλύτερα από του μοντέλου 1. Το μεγάλο πρόβλημα σε αυτό το μοντέλο συναντάται στην κλάση nouturn με precision 0.6 και recall 0.08, δηλαδή πολλά δείγματα της κλάσης nouturn κατατάσσονται από το μοντέλο σε άλλες κλάσεις. Το construction με όλες τις μετρικές ίσες με 1 είναι ιδανικό. Γενικά το μοντέλο 2 είναι βελτιωμένο σε σχέση με το μοντέλο 1.

Στο μοντέλο 3 [9] οι πρώτες 4 κλάσεις έχουν σχετικά χαμηλά precision και recall με εξαίρεση το precision του limit50. Όπως και στα προηγούμενα μοντέλα η κλάση nouturn έχει εξαιρετικά χαμηλό recall 0.19. Το precision των danger, bikecross, childcross, traincross, construction είναι αρκετά χαμηλά.

Στο μοντέλο 4 [10], όμοια με τα προηγούμενα, υπάρχει πρόβλημα στις πρώτες 4 κλάσεις όμως τα αποτελέσματα είναι εμφανώς βελτιωμένα σε σχέση με το μοντέλο 3. Το nouturn έχει ακόμη πολύ χαμηλό recall παρόλα αυτά είναι το υψηλότερο που έχει παρουσιαστεί στα μοντέλα. Το precision των danger, bikecross, childcross, construction είναι αρκετά χαμηλά αλλά σχετικά βελτιωμένα σε σύγκριση με το μοντέλο 3.

4 Σύνοψη

Στα πλαίσια της εργασίας αναπτύχθηκαν μοντέλα ταξινόμησης με στόχο την αποτελεσματική ταξινόμηση των δειγμάτων στις κλάσεις που ανήκουν. Τα 4 μοντέλα που αναπτύχθηκαν αποτελούνται από 4 convolutional layers με kernel size 5×5 , 4 maxpooling layers 2×2 , 2 dropout layers (το πρώτο μετά το τελευταίο maxpooling με 0.25 ή 0.35 για τις περιπτώσεις με batch normalization και το τελευταίο μετά το πρώτο dense layer με 0.5), flatten και τέλος 2 dense layers. Οι παράμετροι επιλέχθηκαν καθώς προσέφεραν τα καλύτερα αποτελέσματα κατά τις δοκιμές.

Η εκπαίδευση των μοντέλων έγινε πάνω σε τροποποιημένα δεδομένα και σε διάστημα 50 κύκλων. Στη συνέχεια, έγινε οπτικοποίηση της μετρικής απόδοσης και του κόστους στα training και validation set. Ένα μοντέλο φαίνεται επιτυχημένο, γενικά, όταν έχει υψηλό accuracy και χαμηλό κόστος στο training και το validation. Η τελικά αξιολόγηση βέβαια γίνεται κατά το testing. Είναι σημαντικό στο testing να προκύπτει υψηλό accuracy όμως αυτό δεν είναι αρκετό για να βρεθεί ακριβώς πως τα πάει το μοντέλο. Για μία καλύτερη εικόνα χρησιμοποιήθηκε το classification report. Με αυτό, φαίνεται ακριβώς που ο αλγόριθμος συναντά προβλήματα και πως με κάθε παραλλαγή χειρίζεται τις προβληματικές κλάσεις.

Και τα 4 μοντέλα έχουν πρόβλημα με τις κλάσεις limit30, limit50, limit60, limit80 παρόλο που για τις συγκεκριμένες υπάρχει επαρκής αριθμός δειγμάτων στο σύνολο εκπαίδευσης. Αυτό συμβαίνει καθώς οι πινακίδες που ανήκουν σε αυτές τις κλάσεις έχουν πολύ παρόμοια χαρακτηριστικά. Μία άλλη προβληματική κλάση που δυσκολεύει και τα 4 μοντέλα είναι η nouturn με εξαιρετικά χαμηλό recall. Εδώ αυτό είναι λογικό καθώς προσφέρονται περισσότερα δείγματα άλλων κλάσεων με παρόμοια χαρακτηριστικά στο σύνολο εκπαίδευσης και ο αλγόριθμος κατατάσσει τα δείγματα του test set στις κλάσεις αυτές. Τέλος, η κλάση construction είναι η κλάση με τα λιγότερα δείγματα στο σύνολο εκπαίδευσης και φαίνεται να δημιουργεί πρόβλημα στους αλγορίθμους χωρίς την ομαλοποίηση. Στα μοντέλα που έγινε batch normalization παρατηρήθηκε υψηλότερο accuracy και χαμηλότερο cross-entropy από τα αντίστοιχα τους χωρίς την ομαλοποίηση.