

Ασκηση 1 Εργαστήριο Βαθιάς Μάθησης Ακ. έτος 2022-2023

Image Classification on Cifar-10

Περιγραφή

Θέμα της 1ης εργασίας του μαθήματος είναι η Ταξινόμηση Εικόνων (Image Classification). Συγκεκριμένα, θα φτιάξουμε ένα ταξινομητή εικόνων χρησιμοποιώντας παραλλαγές του συνελικτικού νευρωνικού δικτύου ["Wide residual networks"](#).

Οι παράμετροι ενός νευρωνικού δικτύου εκπαιδεύονται με βάση τα δεδομένα εκπαίδευσης (training data). Αν έχουμε αρκετά δεδομένα και αν το νευρωνικό δίκτυο είναι όσο deep πρέπει για να μπορεί να δημιουργήσει μια ικανοποιητική αναπαράσταση των δεδομένων τότε θα υπάρξει ένα καλό αποτέλεσμα ταξινόμησης. Δυστυχώς, πολύ σπάνια αυτό συμβαίνει και γι αυτό θα πρέπει να δοκιμάσουμε και άλλους τρόπους για να βοηθήσουμε το μοντέλο να δημιουργήσει μια καλύτερη αναπαράσταση και κατ' επέκταση καλύτερο αποτέλεσμα.

1) Αρχικά θα δοκιμάσουμε να αλλάξουμε την αρχιτεκτονική του δικτύου και να δούμε πως αυτές οι αλλαγές επηρεάζουν την ακρίβεια των προβλέψεων.

2) Αφού έχουμε βρει την αρχιτεκτονική του μοντέλου με το καλύτερο αποτέλεσμα, θα υλοποιήσουμε μια προηγμένη τεχνική επαύξησης δεδομένων (data augmentation) ["MixUp"](#). Αυτή η τεχνική μας επιτρέπει να δημιουργήσουμε τεχνητές εικόνες οι οποίες θα εμπλουτίσουν το αρχικό training dataset και θα βοηθήσουν στο να καλυφθεί καλύτερα η κατανομή του χώρου των κλάσεων που θέλουμε να ταξινομήσουμε.

Θα σας δοθεί ένα βοηθητικό notebook ως σημείο εκκίνησης των πειραμάτων

Σύνολο Δεδομένων

Τα δεδομένα που θα χρησιμοποιήσουμε προέρχονται από το dataset CIFAR-10. Αυτό το dataset περιέχει 60,000 εικόνες, 32x32 διαστάσεων με τρία κανάλια (RGB). Υπάρχουν 10 διαφορετικές κλάσεις που αναπαριστούν αεροπλάνα, αυτοκίνητα, πτηνά, γάτες, ελάφια, σκύλους, βατράχους, άλογα, πλοία και φορτηγά. Σε κάθε κλάση αντιστοιχούν 6,000 εικόνες.

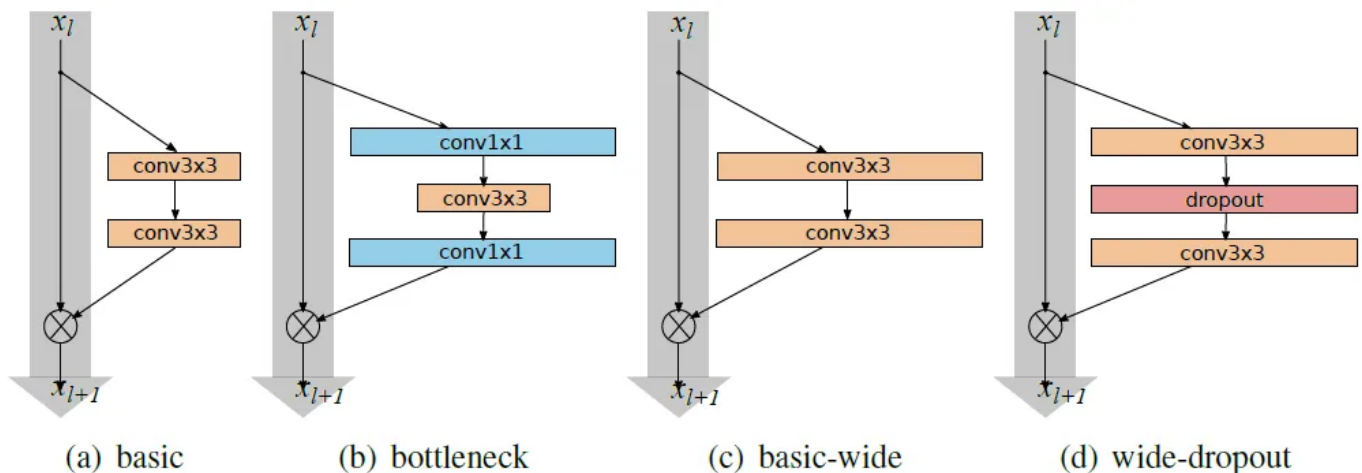


Το μοντέλο

Τα πλατιά (wide) νευρωνικά δίκτυα έχουν λιγότερα κρυφά επίπεδα αλλά περισσότερους νευρώνες ανά επίπεδο. Σε αυτήν την εργασία θα χρησιμοποιήσουμε τα ["Wide residual networks"](#) για να δούμε πως αλλάζοντας την αρχιτεκτονική ενός δικτύου (πλήθος νευρώνων και πλήθος των επιπέδων wide vs deep) αλλάζει και την αναπαράσταση των εικόνων και κατ' επέκταση την απόδοση του μοντέλου.

Πλάτος vs Βάθος στα Residual Networks. Το δίλημμα αυτό έχει απασχολήσει την ερευνητική κοινότητα της μηχανικής μάθησης. Έχει παρατηρηθεί ότι τα λιγότερα βαθιά δίκτυα χρειάζονται εκθετικά περισσότερες παραμέτρους για να έχουν το αποτέλεσμα των βαθιών δικτύων. Γι' αυτό τα ResNets προσπάθησαν να τα κάνουν όσο πιο λεπτά και βαθιά γινόταν.

Στο paper ["Wide residual networks"](#) αποδεικνύουν ότι κάνοντας πιο πλατιά (με τον κατάλληλο τρόπο) τα ResNets, αυξάνεται η απόδοση αλλά και η ταχύτητα εκπαίδευσης.



Αξιολόγηση της ποιότητας της ταξινόμησης

Για το CIFAR-10 dataset, η έξοδος του μοντέλου θα πρέπει να είναι οι προβλέψεις των κλάσεων για κάθε εικόνα. Η αξιολόγηση θα γίνει σύμφωνα με το ποσοστό των προβλέψεων που ανήκουν στις σωστές κλάσεις (accuracy metric)

Βελτιώσεις (και παραδοτέα)

Εφόσον αποκτήσετε μια εικόνα της επίδοσης του έτοιμου δικτύου accuracy και χρόνος εκπαίδευσης στο train και test sets) θα δοκιμάσουμε κάποιες ιδέες για βελτιώσεις στο δίκτυο.

1. Δοκιμή τουλάχιστον των 3 καλύτερων συνδυασμών των παραμέτρων width και depth του Δικτύου wideResNet σύμφωνα με αυτά που προτείνονται και στο paper (results with moderate data augmentation) (3.5 μονάδες)
2. Regularization. Το δίκτυο έχει μηχανισμό ομαλοποίησης (Dropout). Δοκιμάστε να προσθέσετε το Dropout (παράμετρος dropout) όπως αναφέρει το paper και καταγράψτε τα αποτελέσματα για τα παραπάνω μοντέλα. (3.5 μονάδες)
3. α). Κρατώντας τις παραμέτρους του δικτύου με τα καλύτερα αποτελέσματα (load best model από το ζητούμενο 1 και 2) θα κάνετε εκ νέου training με το custom Data Generator (CIFAR_Dataset) στο CIFAR και evaluation στο CIFAR-C χωρίς το mixup και β) Εφαρμόστε τη μέθοδο mixup (συμπληρώστε τον κώδικα στο κελί με το CifarDataset) κάντε εκ νέου την εκπαίδευση και καταγράψτε τα αποτελέσματα. (3 μονάδες)
4. (Bonus) Κάντε επίσης ένα plot των confidence scores (softmax predictions) και συγκρίνετε το με το αντίστοιχο plot όταν το μοντέλο δεν είχε mixup. Τι παρατηρείτε? (2 μονάδες)

Παραδοτέο

1. Από το HELIOS κατεβάστε το zip file που περιέχει την εκφώνηση (DL23_Project1.pdf) , ένα template notebook (ex1_template.ipynb), τον φάκελο με το dataset (data) και άλλα χρήσιμα αρχεία. Θα πρέπει έπειτα να ανεβάζετε και να αποσυμπιέσετε αυτό το αρχείο (zip file) στο home directory σας στο Drive ([/content/drive/My Drive/](#)).
2. Θα πρέπει να χρησιμοποιήσετε το **ex1_template.ipynb** και να συμπληρώσετε τα ζητούμενα. Χρησιμοποιήστε markdown για να εξηγήσετε τις επιλογές σας και τα αποτελέσματά σας.
3. Το παραδοτέο θα πρέπει να υποβληθεί στο HELIOS από το ένα μέλος της ομάδας - δεν χρειάζεται να υποβληθεί και από τα δύο μέλη.