



ΑΛΕΞΑΝΔΡΕΙΟ Τ.Ε.Ι. ΘΕΣΣΑΛΟΝΙΚΗΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Ανάλυση των ResNet και DenseNet και η σύγκρισή τους

Του φοιτητή

Χατζή Χαλήλ Νιαζή

Αρ. Μητρώου: 123859

Επιβλέπων καθηγητής

Κώστας Διαμαντάρας

Θεσσαλονίκη 2020

ΠΡΟΛΟΓΟΣ

Τα Νευρωνικά Δίκτυα (NN , Neural Networks), έχουν χρησιμοποιηθεί εκτεταμένα σε διάφορα προβλήματα της μηχανικής όρασης. Μία πολύ δημοφιλής κατηγορία νευρωνικών δικτύων τα οποία έχουν άριστες επιδόσεις σε προβλήματα ταξινόμησης εικόνας είναι τα Συνελκτικά Νευρωνικά Δίκτυα (CNN, Convolutional NN), τα οποία αποτελούν υποκατηγορία των νευρωνικών δικτύων βαθιάς μάθησης. Η λειτουργία τους βασίζεται στην μοντελοποίηση αναπαραστάσεων υψηλού επιπέδου στα δείγματα εκπαίδευσης μέσω αρχιτεκτονικών πολλών επιπέδων και μη γραμμικών μετασχηματισμών.

Κατά τον παραδοσιακό τρόπο αναγνώρισης αντικειμένων σε εικόνα ή βίντεο, μεμονωμένες εικόνες ή αλληλουχία εικόνων (βίντεο) δίνονται σαν είσοδο σε κάποιο σύστημα προκειμένου να εκτελεστεί ο αλγόριθμος αναγνώρισης. Συνήθως δε, τα χαρακτηριστικά που χρησιμοποιούνται για την αναγνώριση είναι κατά βάση χαρακτηριστικά επιλεγμένα από τον ανθρώπινο παράγοντα. Δηλαδή δεν μεσολαβεί κάποια διαδικασία εκμάθησης μέσω της οποίας εκπαιδεύεται ένας ταξινομητής για την εκπλήρωση του σκοπού της ταξινόμησης. Έτσι, από νωρίς έγινε γνωστό ότι το κλειδί στην αναγνώριση και ταξινόμηση, είναι η επιλογή σωστών χαρακτηριστικών και η καλή περιγραφή τους με τις πιο κλασικές πλέον μεθόδους όπως π.χ. SIFT, HOG κλπ. Το ερώτημα που προέκυψε λοιπόν, συνοψίζεται στην ανάγκη για εύρεση μιας διαδικασίας εκμάθησης αυτών των χαρακτηριστικών, χωρίς την ανάγκη παρέμβασης του ανθρώπινου παράγοντα. Μια καλή προσέγγιση και λύση αυτού του προβλήματος, δίνουν τα CNN μέσω της εξαγωγής χαρακτηριστικών από ενδιάμεσα κρυφά επίπεδα.

Η σχεδίαση των νευρωνικών δικτύων CNN εμπνεύστηκε από τη βιολογία και συγκεκριμένα από τον οπτικό φλοιό τη γάτας ο οποίος διαθέτει μία περίπλοκη διάταξη 18 κυττάρων. Η διάταξη αυτή επιβάλλει την σύνδεση κάποιων νευρωνικών κυττάρων που είναι ευαίσθητα σε συγκεκριμένες υποπεριοχές του ορατού πεδίου, τα οποία και λειτουργούν ως τοπικά 'φίλτρα' στις υποπεριοχές αυτές. Με αυτό τον τρόπο, η γάτα εκμεταλλεύεται τη μεγάλη χωρική συσχέτιση των διάφορων υποπεριοχών, με τελικό αποτέλεσμα την καλύτερη

αντίληψη του χώρου. Τα CNN υπήρξαν μερικές από τις σημαντικότερες καινοτομίες στον τομέα της υπολογιστικής όρασης. Το 2012 ήταν η πρώτη χρονιά που τα νευρωνικά δίκτυα μεγάλωσαν σε εξέχουσα θέση, καθώς ο Alex Krizhevsky τα χρησιμοποίησε για να κερδίσει τον διαγωνισμό ImageNet του ίδιου έτους (οι ετήσιοι Ολυμπιακοί Αγώνες υπολογιστικής όρασης), μειώνοντας το ποσοστό ταξινόμησης σφαλμάτων από 26% σε 15%. Από τότε, μια σειρά εταιρειών έχουν χρησιμοποιήσει τη βαθιά εκπαίδευση στον πυρήνα των υπηρεσιών τους. Το Facebook χρησιμοποιεί νευρωνικά δίκτυα για τους αυτόματους αλγόριθμους ετικετών, το Google για την αναζήτηση φωτογραφιών τους, το Amazon για τις συστάσεις των προϊόντων τους, το Pinterest για την προσαρμογή της αρχικής σελίδας του χρήστη και το Instagram για την υποδομή αναζήτησης.

ΠΕΡΙΛΗΨΗ

Σκοπός αυτής της πτυχιακής εργασίας είναι η παρουσίαση και η ανάλυση της αρχιτεκτονικής των Τεχνητών Νευρωνικών δικτύων. Πιο συγκεκριμένα η εργασία απαρτίζεται από τέσσερα κεφάλαια. Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στα Νευρωνικά Δίκτυα, και παρουσιάζεται μια ιστορική αναδρομή στον συγκεκριμένο κλάδο. Στη συνέχεια γίνεται εισαγωγή στα τεχνητά νευρωνικά δίκτυα καθώς και στις χρήσεις και στα πλεονεκτήματά τους. Στο δεύτερο κεφάλαιο της πτυχιακής γίνεται εισαγωγή στα Residual Networks (ResNet). Αρχικά δίνονται οι κατάλληλοι ορισμοί, και αναλύεται ένα υπολειμματικό δίκτυο (ResNet). Έπειτα γίνεται ανάλυση και αναφορά στα δομικά στοιχεία του ResNet και στην αρχιτεκτονική του δικτύου. Το θέμα του τρίτου κεφαλαίου της εργασίας είναι το DenseNet. Αρχικά γίνεται περιγραφή του δικτύου DenseNet , και στη συνέχεια του κεφαλαίου γίνεται ανάλυση του DenseNet και της αρχιτεκτονικής του. Τέλος στο τέταρτο και τελευταίο κεφάλαιο γίνεται μια σύγκριση των δύο αυτών τεχνητών νευρωνικών δικτύων και εξάγονται τα καταλληλα συμπεράσματα.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Μηχανική Μάθηση, Βαθιά Μάθηση, Τεχνητά Νευρωνικά Δίκτυα, Πυκνά συνδεδεμένα συμβατικά δίκτυα, Υπολειμματικά Δίκτυα

ABSTRACT

The purpose of this thesis is to present and analyze the architecture of Artificial Neural Networks. More specifically, the paper consists of four chapters. The first chapter introduces the Neural Networks, and presents a historical overview of this branch. Then we introduce the artificial neural networks as well as their uses and advantages. The second chapter of the dissertation introduces Residual Networks (ResNet). The appropriate definitions are given first, and a residual network (ResNet) is analyzed. Then the structure and structure of ResNet and the architecture of the network are analyzed and reported. The topic of the third chapter of the thesis is DenseNet. The DenseNet network is first described, and then the chapter is analyzed by DenseNet and its architecture. Finally, in the fourth and last chapter, a comparison of these two artificial neural networks is made and the appropriate conclusions are drawn.

SUBJECT AREA: *Deep Learning, Artificial Neural Networks, ResNet, DenseNet*

ΕΥΧΑΡΙΣΤΙΕΣ

Με το παρόν σύγγραμμα θα ήθελα να ευχαρίστησω όλους όσους με στήριξαν εμπράκτως, είτε άμεσα είτε έμμεσα, στην προσπάθεια περάτωσης αυτής της πτυχιακής εργασίας, αλλά και γενικότερα για την ολοκλήρωση των σπουδών μου και την απόκτηση του πτυχίου μου. Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου κύριο Κώστα Διαμαντάρα, που συμφώνησε να εργασθούμε πάνω στο συγκεκριμένο κλάδο , για την άμεση ανταπόκρισή του σε κάθε είδους ζήτημα που ανέκυπτε καθ' όλη τη διάρκεια της εκπόνησης της εργασίας, αλλά και για όλες τις χρήσιμες συμβουλές του προκειμένου να επιτευχθεί ο τελικός στόχος.

ΠΡΟΛΟΓΟΣ.....	2
ΠΕΡΙΛΗΨΗ.....	4
ABSTRACT	5
ΕΥΧΑΡΙΣΤΙΕΣ.....	6
Ευρετήριο σχημάτων	8
Ευρετήριο πινάκων	9
ΚΕΦΑΛΑΙΟ 1 - ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	10
1.1 Εισαγωγή στα Νευρωνικά δίκτυα.....	10
1.2 Ιστορική Αναδρομή.....	11
1.3 Βιολογικά Νευρωνικά Δίκτυα	14
1.4 Τεχνητό Νευρωνικό Δίκτυο	17
1.5 Πλεονεκτήματα και μειονεκτήματα τεχνητών νευρωνικών δικτύων.....	20
1.6 Convolutional Neural Networks	23
1.7 Πλεονεκτήματα Convolutional Neural Networks.....	24
ΚΕΦΑΛΑΙΟ 2 - Residual Networks (ResNet)	27
2.1 Residual Network (resNet).....	28
2.2 Residual μπλοκ - Δομικά στοιχεία του ResNet.....	30
2.3 Το ResNet και οι παραλλαγές του.....	36
2.3.1 ResNeXt.....	39
2.3.2 Deep Network with Stochastic Depth.....	41
ΚΕΦΑΛΑΙΟ 3 - Densely Connected Convolutional Networks (DenseNet)	46
3.1 Densely Connected Convolutional Networks (DenseNet).....	46
3.2 Δομή του Densenet	48
3.3 Dense and Transition Blocks	51
ΚΕΦΑΛΑΙΟ 4 – ΣΥΓΚΡΙΣΗ RESNET - DENSENET.....	57
4.1 Πλεονεκτήματα ResNet - DenseNet.....	57
4.2 Σύγκριση αρχιτεκτονικής ResNet και DenseNet.....	58
4.3 Η αποτελεσματικότητα του Densenet	62
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	67

Ευρετήριο σχημάτων

Σχήμα 1: Βιολογικός Νευρώνας	15
Σχήμα 2: Μοντέλο Τεχνητού Νευρώνα	17
Σχήμα 3: Παράδειγμα απλού τεχνητού νευρωνικού δικτύου	18
Σχήμα 4: Βιολογικό και τεχνητό νευρωνικό δίκτυο	18
Σχήμα 5: Training Error and Test Error on CIFAR-10 with 20-layer and 56-layer "plain" networks	28
Σχήμα 6: Single Residual Block	30
Σχήμα 7: Identity block.....	31
Σχήμα 8: Convolutional block.....	32
Σχήμα 9: Types of Residual Block	34
Σχήμα 10: Διαφορετικές ερμηνείες του Residual Block.....	35
Σχήμα 11: Increasing network depth leads to worse performance	37
Σχήμα 12: Residual Architecture	38
Σχήμα 13: Residual Block and ResNeXt Block	40
Σχήμα 14: Probability of being disabled.....	41
Σχήμα 15: Error Increases smoothly as the number of deleted layers increases	41
Σχήμα 16: Χαρακτηριστικά των διαδρομών - στο ResNet	44
Σχήμα 17: DenseNet with 5 layers with expansion of 4.....	48
Σχήμα 18: Another look at DenseNet - 121.....	52
Σχήμα 19: One level deeper look at DenseNet - 121	52
Σχήμα 20: Full Schematic representation of DenseNet - 121	53
Σχήμα 21: Transition Layer/Block 1	54
Σχήμα 22: Transition Layer/Block 2.....	55
Σχήμα 23: ResNet Architecture and Layer Details	59
Σχήμα 24: Plain Layers and Residual Block.....	60
Σχήμα 25: Comparison of the DenseNet and ResNet	63
Σχήμα 26: ResNet vs DenseNet on the ImageNet dataset.....	65

Ευρετήριο πινάκων

Πίνακας 1: Sizes of outputs and convolutional kernels for different DenseNets [1] architectures on ImageNet.....	50
Πίνακας 2: Results of DenseNets on CIFAR datasets.....	51
Πίνακας 3: DenseNet Architecture	57

ΚΕΦΑΛΑΙΟ 1 - ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

1.1 Εισαγωγή στα Νευρωνικά δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα είναι απλοποιημένα μοντέλα του κεντρικού νευρικού συστήματος του ανθρώπου. Συνεπώς μπορούν και μιμούνται τη λειτουργία των βασική λειτουργία των βιολογικών νευρώνων του εγκεφάλου του ανθρώπου και τη δομή των βιολογικών Νευρωνικών Δικτύων. Αποτελούνται από διασυνδεδεμένα υπολογιστικά στοιχεία που έχουν την ικανότητα να ανταποκρίνονται σε ερεθίσματα που δέχονται στην είσοδο τους και να μαθαίνουν να προσαρμόζονται στο περιβάλλον τους.

Η μελέτη των Τεχνητών Νευρωνικών Δικτύων είναι εμπνευσμένη από τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου. Οι νευρώνες είναι βασικό δομικό στοιχείο του εγκεφάλου, δηλαδή τα νευρικά κύτταρα τα οποία δημιουργούν ένα πυκνό δίκτυο επικοινωνίας μεταξύ τους. Η ελπίδα για ανακάλυψη ενός νέου υπολογιστικού μοντέλου με δικτυακή δομή παρόμοια με αυτή του ανθρώπινου εγκεφάλου δίνει το κίνητρο για τη μελέτη του νευρώνα και των Νευρωνικών Δικτύων. Αυτή η καινούρια υπολογιστική πλατφόρμα ονομάζεται Connectionist Model, και θα είναι η πλέον κατάλληλη για την ανάπτυξη διαδικασιών και ευφών αλγορίθμων που έχουν να κάνουν με τη νοημοσύνη, όπως η μνήμη, η μάθηση η γενίκευση κ.λ.π.

Τα συνηθισμένα τεχνητά νευρωνικά δίκτυα χρησιμοποιούν απλοποιημένα μοντέλα νευρώνων έτσι ώστε να χρησιμοποιούν τα πολύ απλά χαρακτηριστικά των λεπτομερών μοντέλων που χρησιμοποιούνται στη νευρολογία. Ακόμη και αυτά τα απλά μοντέλα μπορούν να δημιουργήσουν ενδιαφέροντα δίκτυα αρκεί να πληρούν δύο βασικά χαρακτηριστικά :

1. Οι νευρώνες να έχουν ρυθμιζόμενες παραμέτρους ώστε να διευκολύνεται η διαδικασία της μάθησης. Αυτή η ιδιότητα είναι γνωστή ως πλαστικότητα των νευρώνων.

2. Το δίκτυο να αποτελείται από πολλούς νευρώνες ώστε να επιτυγχάνεται κατανομή της πληροφορίας και παραλληλισμός της επεξεργασίας.

Το κύριο αντικείμενο των τεχνητών νευρωνικών δικτύων είναι η ανάπτυξη κατάλληλων αλγορίθμων εκπαίδευσης και ανάκλησης της πληροφορίας που αυτά περιέχουν έτσι ώστε να προσομοιάζονται ευφυείς διαδικασίες. Για να γίνει αυτό πραγματικότητα πρέπει να οριστεί το κατάλληλο περιβάλλον εκπαίδευσης πχ. Αν το δίκτυο θα αφήνεται μόνο του να αυτοοργανωθεί με κάποιο συγκεκριμένο κριτήριο και στόχο , ή αν το δίκτυο θα εκπαιδεύεται με επίβλεψη , δηλαδή με την χρήση κάποιων δεδομένων οδηγών-δασκάλων. Η θεμελιώδης μονάδα επεξεργασίας της πληροφορίας στα τεχνητά νευρωνικά δίκτυα είναι ο νευρώνας ή κόμβος.

1.2 Ιστορική Αναδρομή

Η μελέτη του ανθρώπινου εγκεφάλου είναι χιλιάδων ετών. Με την έλευση της σύγχρονης ηλεκτρονικής, ήταν φυσικό να προσπαθήσουμε να εκμεταλλευτούμε αυτήν τη διαδικασία σκέψης. Το πρώτο βήμα προς τα τεχνητά νευρικά δίκτυα ήρθε το 1943 όταν ο Warren McCulloch, ένας νευροφυσιολόγος και ένας νεαρός μαθηματικός, Walter Pitts, έγραψαν μια εργασία για το πώς μπορούν να λειτουργούν οι νευρώνες. Μοντελοποίησαν ένα απλό νευρωνικό δίκτυο με ηλεκτρικά κυκλώματα.

Η ενίσχυση αυτής της έννοιας των νευρώνων και του τρόπου λειτουργίας τους ήταν ένα βιβλίο του Donald Hebb. Ο Οργανισμός Συμπεριφοράς γράφτηκε το 1949. Επισήμανε ότι τα νευρικά μονοπάτια ενισχύονται κάθε φορά που χρησιμοποιούνται.

Καθώς ξεκίνησε η ανάπτυξη των υπολογιστών στα πρώτα χρόνια της δεκαετίας του 1950, κατέστη δυνατό να αρχίσει η μοντελοποίηση των βασικών στοιχείων αυτών των θεωριών σχετικά με την ανθρώπινη σκέψη. Ο Nathaniel Rochester από τα ερευνητικά εργαστήρια της IBM οδήγησε την πρώτη προσπάθεια προσομοίωσης ενός νευρικού δικτύου. Η πρώτη προσπάθεια απέτυχε. Αλλά αργότερα οι προσπάθειες ήταν επιτυχείς. Κατά τη διάρκεια αυτής της περιόδου άρχισαν να ανθίζουν οι παραδοσιακοί υπολογιστές και, όπως έκανε, η έμφαση στον υπολογισμό άφησε την έρευνα του νευρικού συστήματος στο υπόβαθρο.

Ωστόσο, κατά τη διάρκεια αυτής της περιόδου, οι υποστηρικτές των «μηχανών σκέψης» συνέχισαν να υποστηρίζουν τις υποθέσεις τους. Το 1956, το Dartmouth Summer Research Project για την Τεχνητή Νοημοσύνη έδωσε ώθηση τόσο στην τεχνητή νοημοσύνη όσο και στα νευρωνικά δίκτυα. Ένα από τα αποτελέσματα αυτής της διαδικασίας ήταν η τόνωση της έρευνας τόσο στην έξυπνη πλευρά, την τεχνητή νοημοσύνη, όπως είναι γνωστή σε ολόκληρο τον κλάδο, όσο και στο πολύ χαμηλότερο επίπεδο νευρωνικής επεξεργασίας του εγκεφάλου.

Στα χρόνια που ακολούθησαν το Dartmouth Project, και πιο συγκεκριμένα ο John von Neumann πρότεινε να μιμηθούν απλές λειτουργίες νευρώνων χρησιμοποιώντας τηλεγραφικά ρελέ ή σωλήνες κενού. Επίσης, ο Frank Rosenblatt, ένας νευροβιολόγος του Cornell, άρχισε να εργάζεται στο Perceptron. Ήταν ενθουσιασμένος με τη λειτουργία του ματιού μύγας. Μεγάλο μέρος της επεξεργασίας που κάνει την μύγα να πετάει γίνεται στο μάτι της. Το Perceptron δίκτυο, που προέκυψε από αυτήν την έρευνα, κατασκευάστηκε σε υλικό και είναι το παλαιότερο νευρωνικό δίκτυο που χρησιμοποιείται ακόμα σήμερα. Ένα perceptron ενός στρώματος βρέθηκε χρήσιμο για την ταξινόμηση ενός συνόλου εισόδων συνεχούς αξίας σε μία από τις δύο κατηγορίες. Το perceptron υπολογίζει ένα σταθμισμένο άθροισμα των εισόδων, αφαιρεί ένα κατώφλι (threshold) και περνά μία από τις δύο πιθανές τιμές ως αποτέλεσμα. Δυστυχώς, το perceptron είναι περιορισμένο και αποδείχθηκε ως έχει κατά τη διάρκεια των «απογοητευμένων ετών» στο βιβλίο Perceptrons του 1969 του Marvin Minsky και του Seymour Papert.

Το 1959, ο Bernard Widrow και ο Marcian Hoff του Στάνφορντ ανέπτυξαν μοντέλα που ονόμασαν ADALINE και MADALINE. Αυτά τα μοντέλα ονομάστηκαν για τη χρήση του Multiple ADaptive LINear Elements. Το MADALINE ήταν το πρώτο νευρωνικό δίκτυο που εφαρμόστηκε σε πραγματικό πρόβλημα. Είναι ένα προσαρμοστικό φίλτρο που εξαλείφει τις ηχώ στις τηλεφωνικές γραμμές. Αυτό το νευρωνικό δίκτυο βρίσκεται ακόμη σε εμπορική χρήση.

Δυστυχώς, αυτές οι προηγούμενες επιτυχίες προκάλεσαν τους ανθρώπους να υπερβάλλουν τις δυνατότητες των νευρωνικών δικτύων, ιδίως ενόψει του περιορισμού των διαθέσιμων ηλεκτρονικών. Αυτή η υπερβολική διαφημιστική εκστρατεία, που βγήκε έξω από τον ακαδημαϊκό και τεχνικό κόσμο, μολύνει τη γενική βιβλιογραφία της εποχής. Η απογοήτευση τέθηκε ως εκπλήρωση υποσχέσεων. Επίσης, ένας φόβος ξεκίνησε καθώς οι συγγραφείς άρχισαν να συλλογίζονται τι επίδραση θα έχουν οι «μηχανές σκέψης» στον άνθρωπο.

Η σειρά του Asimov στα ρομπότ αποκάλυψε τις επιπτώσεις στα ηθικά και τις αξίες του ανθρώπου, όταν οι μηχανές είναι ικανές να κάνουν όλη τη δουλειά της ανθρωπότητας. Άλλοι συγγραφείς δημιούργησαν περισσότερους απαίσιους υπολογιστές, όπως το HAL από την ταινία 2001.

Αυτοί οι φόβοι, σε συνδυασμό με ανεκπλήρωτους, εξωφρενικούς ισχυρισμούς, προκάλεσαν σεβαστές φωνές να ασκήσουν κριτική στην έρευνα του νευρικού δικτύου. Το αποτέλεσμα ήταν να σταματήσει μεγάλο μέρος της χρηματοδότησης. Αυτή η περίοδος αναστατωμένης ανάπτυξης διήρκεσε μέχρι το 1981.

Το 1982 πολλά γεγονότα προκάλεσαν ένα νέο ενδιαφέρον. Ο Τζον Χόπφιλντ του Caltech παρουσίασε ένα έγγραφο στην Εθνική Ακαδημία Επιστημών. Η προσέγγιση του Χόπφιλντ δεν ήταν απλώς να μοντελοποιήσει το μυαλό αλλά να δημιουργήσει χρήσιμες συσκευές. Με σαφήνεια και μαθηματική ανάλυση, έδειξε πώς αυτά τα δίκτυα θα μπορούσαν να λειτουργήσουν και τι θα μπορούσαν να κάνουν. Ωστόσο, το μεγαλύτερο πλεονέκτημα του Χόπφιλντ ήταν το χάρισμά του. Ήταν αρθρωτός, συμπαθητικός και πρωταθλητής μιας αδρανοποιημένης τεχνολογίας. Ταυτόχρονα, συνέβη ένα άλλο γεγονός. Πραγματοποιήθηκε συνέδριο στο Κιότο της Ιαπωνίας. Αυτό το συνέδριο ήταν η Κοινή Διάσκεψη ΗΠΑ-Ιαπωνίας για Συνεργατικά / Ανταγωνιστικά Νευρωνικά Δίκτυα.

Μέχρι το 1985 το Αμερικανικό Ινστιτούτο Φυσικής ξεκίνησε μια ετήσια συνάντηση - Neural Networks for Computing. Μέχρι το 1987, το πρώτο Διεθνές Συνέδριο για τα Νευρωνικά Δίκτυα του Ινστιτούτου Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών (IEEE) προσέλκυσε περισσότερους από 1.800 συμμετέχοντες.

Μέχρι το 1989 στη συνάντηση Neural Networks for Defense, ο Bernard Widrow είπε στο ακροατήριό του ότι ασχολήθηκαν με τον Δεύτερο Παγκόσμιο Πόλεμο, «ο Δεύτερος Παγκόσμιος Πόλεμος δεν συνέβη ποτέ», όπου τα πεδία της μάχης είναι το παγκόσμιο εμπόριο και η κατασκευή. Το ερευνητικό πρόγραμμα καινοτομίας του Υπουργείου Άμυνας των ΗΠΑ για τις μικρές επιχειρήσεις του 1990 ονόμασε 16 θέματα που στοχεύουν συγκεκριμένα τα νευρωνικά δίκτυα με ένα επιπλέον 13 που αναφέρει την πιθανή χρήση των νευρωνικών δικτύων.

Σήμερα, συζητούνται τα νευρικά δίκτυα παντού. Η υπόσχεσή τους φαίνεται πολύ φωτεινή, καθώς η ίδια η φύση είναι η απόδειξη ότι αυτό το είδος λειτουργεί. Ωστόσο, το μέλλον τους, πράγματι το κλειδί για ολόκληρη την τεχνολογία, έγκειται στην ανάπτυξη υλικού. Επί του παρόντος, η περισσότερη ανάπτυξη νευρωνικών δικτύων αποδεικνύει απλώς ότι το πρότυπο δουλεύει. Αυτή η έρευνα αναπτύσσει νευρωνικά δίκτυα που, λόγω περιορισμών επεξεργασίας, χρειάζονται εβδομάδες για να μάθουν. Οι εταιρείες εργάζονται σε τρεις τύπους neuro chip - ψηφιακά, αναλογικά και οπτικά. Ορισμένες εταιρείες εργάζονται για τη δημιουργία ενός "μεταγλωττιστή πυριτίου" για τη δημιουργία ενός νευρωνικού δικτύου Εφαρμογή συγκεκριμένου ολοκληρωμένου κυκλώματος (ASIC). Αυτά τα ASIC και τα ψηφιακά τσιπ που μοιάζουν με νευρώνες φαίνεται να είναι το κύμα του εγγύς μέλλοντος. Τελικά, τα οπτικά τσιπ φαίνονται πολύ ελπιδοφόρα. Ωστόσο, μπορεί να είναι χρόνια πριν τα οπτικά τσιπ να δουν το φως της ημέρας σε εμπορικές εφαρμογές.

1.3 Βιολογικά Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα αντλούν μεγάλο μέρος της έμπνευσής τους από το βιολογικό νευρικό σύστημα. Είναι επομένως πολύ χρήσιμο να γνωρίζουμε πώς οργανώνεται αυτό το σύστημα.

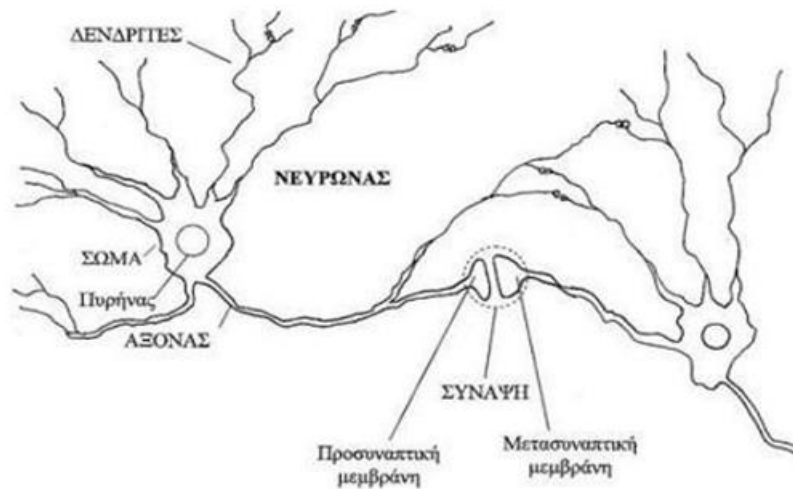
Τα περισσότερα ζωντανά πλάσματα, τα οποία έχουν την ικανότητα να προσαρμοστούν σε ένα μεταβαλλόμενο περιβάλλον, χρειάζονται μια μονάδα ελέγχου που είναι ικανή να μάθει. Τα ανώτερα αναπτυγμένα ζώα και οι άνθρωποι χρησιμοποιούν πολύπλοκα δίκτυα υψηλής εξειδίκευσης νευρώνων για να εκτελέσουν αυτό το έργο.

Η μονάδα ελέγχου - ή ο εγκέφαλος - μπορεί να χωριστεί σε διαφορετικές ανατομικές και λειτουργικές υπομονάδες, καθεμία από τις οποίες έχει συγκεκριμένες εργασίες όπως η όραση, η ακοή, ο κινητήρας και ο αισθητήρας. Ο εγκέφαλος συνδέεται με νεύρα με τους αισθητήρες στο υπόλοιπο σώμα.

Ο εγκέφαλος αποτελείται από έναν πολύ μεγάλο αριθμό νευρώνων, περίπου 1011 κατά μέσο όρο. Αυτά μπορούν να θεωρηθούν ως τα βασικά δομικά τούβλα για το κεντρικό νευρικό σύστημα (CNS). Οι νευρώνες διασυνδέονται σε σημεία που ονομάζονται συνάψεις. Η πολυπλοκότητα του εγκεφάλου οφείλεται στον τεράστιο αριθμό πολύ διασυνδεδεμένων απλών μονάδων που λειτουργούν παράλληλα, με έναν μεμονωμένο νευρώνα να λαμβάνει είσοδο από έως και 10000 άλλες.

Ο νευρώνας περιέχει όλες τις δομές ενός ζωικού κυττάρου. Η πολυπλοκότητα της δομής και των διεργασιών σε ένα απλό κύτταρο είναι τεράστια. Ακόμα και τα πιο εξελιγμένα μοντέλα νευρώνων σε σχέση με τα τεχνητά νευρικά δίκτυα μοιάζουν συγκριτικά με παιχνίδια.

Δομικά, ο νευρώνας μπορεί να χωριστεί σε τρία κύρια μέρη: το κυτταρικό σώμα (soma), τους δένδριτες και τον άξονα, βλ. Σχήμα 1.1 για μια απεικόνιση.



Σχήμα 1: Ένας βιολογικός νευρώνας

Σχήμα 1: Βιολογικός Νευρώνας

Το κυτταρικό σώμα περιέχει τα οργανίδια του νευρώνα και επίσης οι «δενδρίτες» προέρχονται από εκεί. Αυτές είναι λεπτές και ευρέως διακλαδισμένες ίνες, που φτάνουν σε διαφορετικές κατευθύνσεις για να κάνουν συνδέσεις με μεγαλύτερο αριθμό κυττάρων μέσα στο σύμπλεγμα.

Η σύνδεση εισόδου γίνεται από τους άξονες άλλων κυττάρων στους δενδρίτες ή απευθείας στο σώμα του κυττάρου. Αυτές είναι γνωστές ως αξονενδριτικές και αξονωματικές συνάψεις.

Υπάρχει μόνο ένας άξονας ανά νευρώνα. Είναι μια μονή και μακρά ίνα, η οποία μεταφέρει το σήμα εξόδου της κυψέλης ως ηλεκτρικές παλμούς (δυναμικό δράσης) κατά μήκος. Το άκρο του άξονα μπορεί να χωριστεί σε πολλούς κλάδους, οι οποίοι στη συνέχεια συνδέονται με άλλα κύτταρα. Οι κλάδοι έχουν τη λειτουργία να ανεβάζουν το σήμα σε πολλές άλλες εισόδους.

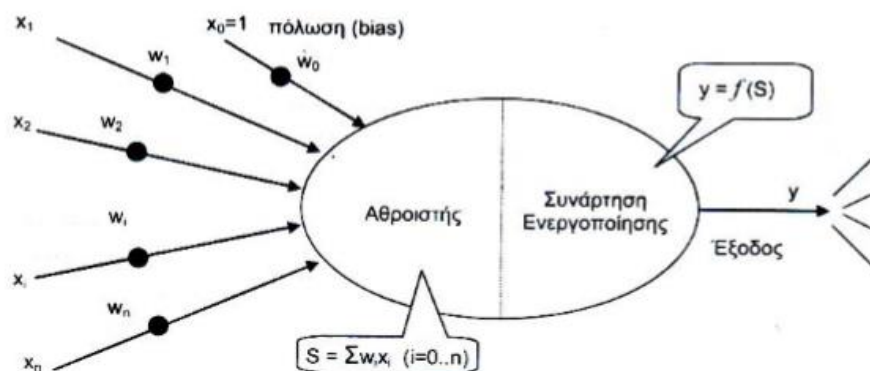
Υπάρχουν πολλοί διαφορετικοί τύποι νευρωνικών κυττάρων που βρίσκονται στο νευρικό σύστημα. Οι διαφορές οφείλονται στη θέση και τη λειτουργία τους. Οι νευρώνες εκτελούν βασικά την ακόλουθη λειτουργία: αθροίζονται όλες οι εισοδοί στο κύτταρο, οι οποίες μπορεί να ποικίλλουν ανάλογα με την ισχύ της σύνδεσης ή τη συχνότητα του εισερχόμενου σήματος. Το άθροισμα εισόδου υποβάλλεται σε επεξεργασία με συνάρτηση κατωφλίου και παράγει σήμα εξόδου.

Ο εγκέφαλος λειτουργεί με παράλληλο και σειριακό τρόπο. Η παράλληλη και σειριακή φύση του εγκεφάλου είναι εμφανής από τη φυσική ανατομία του νευρικού συστήματος. Ότι υπάρχει σειριακή και παράλληλη επεξεργασία μπορεί να φανεί εύκολα από το χρόνο που απαιτείται για την εκτέλεση εργασιών. Για παράδειγμα, ένας άνθρωπος μπορεί να αναγνωρίσει την εικόνα ενός άλλου ατόμου σε περίπου 100 ms. Δεδομένου του χρόνου επεξεργασίας 1 ms για έναν μεμονωμένο νευρώνα, αυτό σημαίνει ότι ένας ορισμένος αριθμός νευρώνων, αλλά μικρότερος από 100, εμπλέκονται σε σειριακό. λαμβάνοντας υπόψη ότι η πολυπλοκότητα της εργασίας αποτελεί απόδειξη για παράλληλη επεξεργασία, επειδή μια δύσκολη εργασία αναγνώρισης δεν μπορεί να εκτελεστεί από έναν τόσο μικρό αριθμό νευρώνων. Αυτό το φαινόμενο είναι γνωστό ως κανόνας 100 βημάτων.

Τα βιολογικά νευρικά συστήματα έχουν συνήθως πολύ υψηλή ανοχή σφαλμάτων. Πειράματα με άτομα με εγκεφαλικά τραύματα έδειξαν ότι η βλάβη των νευρώνων σε ένα ορισμένο επίπεδο δεν επηρεάζει απαραίτητα την απόδοση του συστήματος, αν και εργασίες όπως η γραφή ή η ομιλία μπορεί να πρέπει να μαθευτούν ξανά. Αυτό μπορεί να θεωρηθεί ως επανεκπαίδευση του δικτύου.

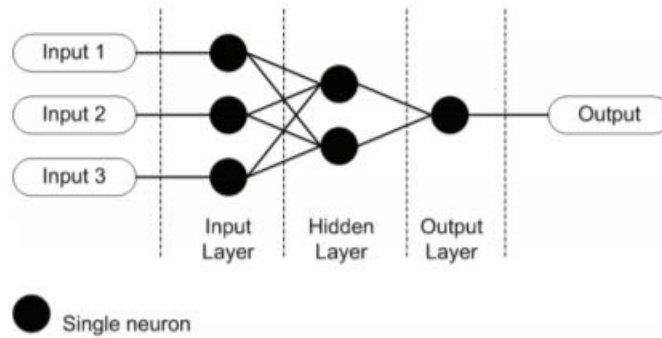
1.4 Τεχνητό Νευρωνικό Δίκτυο

Το Τεχνητό Νευρωνικό Δίκτυο (ANN) είναι ένα μαθηματικό μοντέλο που προσπαθεί να προσομοιώσει τη δομή και τις λειτουργίες των βιολογικών νευρωνικών δικτύων. Βασικό δομικό στοιχείο κάθε τεχνητού νευρικού δικτύου είναι ο τεχνητός νευρώνας, δηλαδή ένα απλό μαθηματικό μοντέλο (λειτουργία). Ένα τέτοιο μοντέλο έχει τρία απλά σύνολα κανόνων: πολλαπλασιασμό, άθροισμα και ενεργοποίηση. Στην είσοδο του τεχνητού νευρώνα οι εισροές σταθμίζονται, πράγμα που σημαίνει ότι κάθε τιμή εισόδου πολλαπλασιάζεται με το ατομικό βάρος. Στο μεσαίο τμήμα του τεχνητού νευρώνα είναι η συνάρτηση αθροίσματος που αθροίζει όλες τις σταθμισμένες εισόδους και πολώσεων. Στην έξοδο του τεχνητού νευρώνα το άθροισμα των προηγούμενων σταθμισμένων εισόδων και πολώσεων περνούν μέσω της συνάρτησης ενεργοποίησης που ονομάζεται επίσης συνάρτηση μεταφοράς (Σχ. 2.).



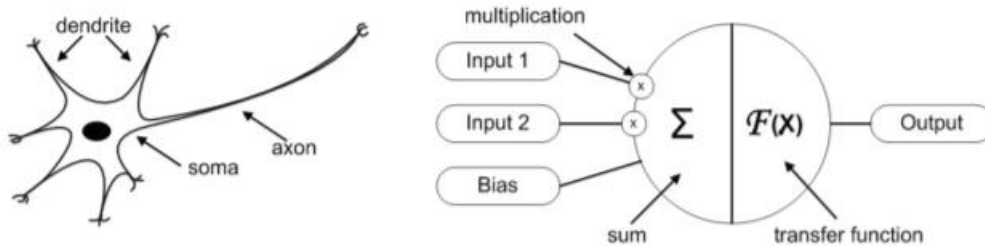
Σχήμα 2: Μοντέλο Τεχνητού Νευρώνα

Παρόλο που οι αρχές λειτουργίας και το απλό σύνολο κανόνων του τεχνητού νευρώνα δεν είναι και τίποτα το ιδιαίτερο, το δυναμικό και η ισχύς υπολογισμού αυτών των μοντέλων ζωντανεύουν όταν αρχίσουμε να τα συνδέουμε σε τεχνητά νευρικά δίκτυα (Εικ. 2.). Αυτά τα τεχνητά νευρικά δίκτυα χρησιμοποιούν απλώς το γεγονός ότι η πολυπλοκότητα μπορεί να αναπτυχθεί από λίγους βασικούς και απλούς κανόνες.



Σχήμα 3: Παράδειγμα απλού τεχνητού νευρωνικού δικτύου

Ο τεχνητός νευρώνας είναι ένα βασικό δομικό στοιχείο κάθε τεχνητού νευρικού δικτύου. Ο σχεδιασμός και οι λειτουργίες του προέρχονται από την παρατήρηση ενός βιολογικού νευρώνα που αποτελεί βασικό δομικό στοιχείο βιολογικών νευρωνικών δικτύων (συστημάτων) που περιλαμβάνει τον εγκέφαλο, τον νωτιαίο μυελό και τα περιφερειακά γάγγλια. Οι ομοιότητες στο σχεδιασμό και τις λειτουργίες φαίνονται στο Σχ.3. όπου η αριστερή πλευρά ενός σχήματος αντιπροσωπεύει έναν βιολογικό νευρώνα με τους soma, τους δενδρίτες και τον άξονα και όπου η δεξιά πλευρά ενός σχήματος αντιπροσωπεύει έναν τεχνητό νευρώνα με τις εισόδους, τα βάρη, τη λειτουργία μεταφοράς, την πόλωση και τις εξόδους.



Σχήμα 4: Βιολογικό και τεχνητό νευρωνικό δίκτυο

Σε περίπτωση που οι πληροφορίες βιολογικών νευρώνων εισέρχονται στον νευρώνα μέσω του δενδρίτη, το κυτταρικό σώμα επεξεργάζεται τις πληροφορίες και τις μεταδίδει μέσω του άξονα. Σε περίπτωση τεχνητού νευρώνα οι πληροφορίες εισέρχονται στο σώμα ενός τεχνητού νευρώνα μέσω εισόδων που είναι σταθμισμένες (κάθε είσοδος μπορεί να πολλαπλασιαστεί μεμονωμένα με το βάρος). Στη συνέχεια, το σώμα ενός τεχνητού νευρώνα αθροίζει τις σταθμισμένες εισόδους και τις πολώσεις και «επεξεργάζεται» το άθροισμα με μια λειτουργία μεταφοράς. Στο τέλος ένας τεχνητός νευρώνας περνά τις επεξεργασμένες πληροφορίες μέσω εξόδου (εξ). Το όφελος της απλότητας του τεχνητού νευρώνα φαίνεται στη μαθηματική περιγραφή του παρακάτω :

$$, (1) \quad y(k) = F \left(\sum_{i=0}^m w_i(k) \cdot x_i(k) + b \right)$$

Όπου:

- $x_i(k)$ είναι η τιμή εισαγωγής σε διακριτό χρόνοό όπουό πηγαίνει από 0 σε m
- $w(k)$ είναι η τιμή βάρους σε διακριτό χρόνοό όπουό πηγαίνει από 0 σεδ m
- b είναι η πόλωση,
- F είναι μια συνάρτηση μεταφοράς,
- $y(k)$ είναι η τιμή εξόδου σε διακριτό χρόνοό .

Όπως φαίνεται από ένα μοντέλο τεχνητού νευρώνα και την εξίσωσή του (1) η κύρια άγνωστη μεταβλητή του μοντέλου μας είναι η λειτουργία μεταφοράς του. Η λειτουργία μεταφοράς καθορίζει τις ιδιότητες του τεχνητού νευρώνα και μπορεί να είναι οποιαδήποτε μαθηματική συνάρτηση. Επιλέγεται με βάση το πρόβλημα που πρέπει να λύσει ο τεχνητός νευρώνας (τεχνητό νευρωνικό δίκτυο) και στις περισσότερες περιπτώσεις επιλέγεται από το ακόλουθο σύνολο συναρτήσεων: συνάρτηση βημαίων, γραμμική συνάρτηση και μη γραμμική (Sigmoid) συνάρτηση.

Η συνάρτηση βημάτων είναι δυαδική συνάρτηση που έχει μόνο δύο πιθανές τιμές εξόδου (π.χ. μηδέν και μία). Αυτό σημαίνει ότι εάν η τιμή εισόδου πληροί συγκεκριμένο όριο, η τιμή εξόδου οδηγεί σε μία τιμή και εάν το συγκεκριμένο κατώτατο όριο δεν πληροί που έχει ως αποτέλεσμα διαφορετική τιμή εξόδου. Η κατάσταση μπορεί να περιγραφεί με την εξίσωση (2).

$$y = \begin{cases} 1 & \text{if } w_i x_i \geq \text{threshold} \\ 0 & \text{if } w_i x_i < \text{threshold} \end{cases} \quad (2)$$

Όταν αυτός ο τύπος λειτουργίας μεταφοράς χρησιμοποιείται στον τεχνητό νευρώνα, ονομάζουμε αυτόν τον τεχνητό νευρώνα perceptron. Το Perceptron χρησιμοποιείται για την επίλυση προβλημάτων ταξινόμησης και ως εκ τούτου μπορεί να βρεθεί συνήθως στο τελευταίο στρώμα τεχνητών νευρικών δικτύων.

Σε περίπτωση γραμμικής μεταφοράς, ο τεχνητός νευρώνας κάνει απλό γραμμικό μετασχηματισμό πάνω από το άθροισμα των σταθμισμένων εισόδων και πόλωσης. Ένας τέτοιος τεχνητός νευρώνας έρχεται σε αντίθεση με το perceptron που χρησιμοποιείται πιο συχνά στο στρώμα εισόδου τεχνητών νευρικών δικτύων. Όταν χρησιμοποιούμε μη γραμμική συνάρτηση, η σιγμοειδής συνάρτηση είναι η πιο συχνά χρησιμοποιούμενη. Η λειτουργία Sigmoid έχει υπολογίσει εύκολα παράγωγο, το οποίο μπορεί να είναι σημαντικό κατά τον υπολογισμό των ενημερώσεων βάρους στο τεχνητό νευρικό δίκτυο.

1.5 Πλεονεκτήματα και μειονεκτήματα τεχνητών νευρωνικών δικτύων

Στο παρόν υποκεφάλαιο θα αναλυθούν τα βασικά πλεονεκτήματα και μειονεκτήματα των τεχνητών νευρωνικών δικτύων.

Πλεονεκτήματα Τεχγικών Νευρωνικών Δικτύων

- ▶ Αποθήκευση πληροφοριών σε ολόκληρο το δίκτυο: Πληροφορίες όπως στον παραδοσιακό προγραμματισμό αποθηκεύονται σε ολόκληρο το δίκτυο και όχι σε μια βάση δεδομένων. Η εξαφάνιση μερικών πληροφοριών σε ένα μέρος δεν εμποδίζει τη λειτουργία του δικτύου.
- ▶ Δυνατότητα εργασίας με ελλιπείς γνώσεις: Μετά την εκπαίδευση ANN, τα δεδομένα ενδέχεται να παράγουν έξοδο ακόμη και με ελλιπείς πληροφορίες. Η απώλεια απόδοσης εδώ εξαρτάται από τη σημασία των πληροφοριών που λείπουν.
- ▶ Ανοχή σφαλμάτων: Η καταστροφή ενός ή περισσότερων κυττάρων του ANN δεν το εμποδίζει να παράγει έξοδο. Αυτή η δυνατότητα κάνει τα δίκτυα ανεκτικά σε σφάλματα.
- ▶ Κατανεμημένη μνήμη: Για να είναι σε θέση να μάθει το ANN, είναι απαραίτητο να προσδιοριστούν τα παραδείγματα και να διδαχθεί το δίκτυο σύμφωνα με την επιθυμητή έξοδο, δείχνοντας αυτά τα παραδείγματα στο δίκτυο. Η επιτυχία του δικτύου είναι ευθέως ανάλογη με τις επιλεγμένες παρουσίες και εάν το συμβάν δεν μπορεί να εμφανιστεί στο δίκτυο από όλες τις πτυχές του, το δίκτυο μπορεί να παράγει ψευδή έξοδο.
- ▶ Σταδιακή καταστροφή: Ένα δίκτυο επιβραδύνεται με την πάροδο του χρόνου και υφίσταται σχετική υποβάθμιση. Το πρόβλημα του δικτύου δεν διαβρώνεται αμέσως.
- ▶ Δυνατότητα μηχανικής μάθησης: Τα τεχνητά νευρικά δίκτυα μαθαίνουν γεγονότα και λαμβάνουν αποφάσεις σχολιάζοντας παρόμοια γεγονότα.
- ▶ Δυνατότητα παράλληλης επεξεργασίας: Τα τεχνητά νευρικά δίκτυα έχουν αριθμητική ισχύ που μπορεί να εκτελέσει περισσότερες από μία εργασίες ταυτόχρονα.

Μειονεκτήματα Τεχνητών Νευρωνικών Δικτύων

- ▶ **Εξάρτηση υλικού:** Τα τεχνητά νευρικά δίκτυα απαιτούν επεξεργαστές με παράλληλη ισχύ επεξεργασίας, σύμφωνα με τη δομή τους. Για το λόγο αυτό, εξαρτάται η πραγματοποίηση του εξοπλισμού.
- ▶ **Ανεξήγητη συμπεριφορά του δικτύου:** Αυτό είναι το πιο σημαντικό πρόβλημα του ANN. Όταν το ANN παράγει μια διερευνητική λύση, δεν δίνει ιδέα για το γιατί και το πώς. Αυτό μειώνει την εμπιστοσύνη στο δίκτυο.
- ▶ **Προσδιορισμός της σωστής δομής δικτύου:** Δεν υπάρχει συγκεκριμένος κανόνας για τον προσδιορισμό της δομής των τεχνητών νευρικών δικτύων. Η κατάλληλη δομή δικτύου επιτυγχάνεται μέσω εμπειρίας και δοκιμής και σφάλματος.
- ▶ **Δυσκολία εμφάνισης του προβλήματος στο δίκτυο:** Τα ANN μπορούν να λειτουργήσουν με αριθμητικές πληροφορίες. Τα προβλήματα πρέπει να μεταφραστούν σε αριθμητικές τιμές πριν εισαχθούν στο ANN. Ο μηχανισμός εμφάνισης που θα καθοριστεί εδώ θα επηρεάσει άμεσα την απόδοση του δικτύου. Αυτό εξαρτάται από την ικανότητα του χρήστη.
- ▶ **Η διάρκεια του δικτύου είναι άγνωστη:** Το δίκτυο μειώνεται σε μια συγκεκριμένη τιμή του σφάλματος στο δείγμα, αυτό σημαίνει ότι η εκπαίδευση έχει ολοκληρωθεί. Αυτή η τιμή δεν μας δίνει βέλτιστα αποτελέσματα.

Τα επιστημονικά τεχνητά νευρικά δίκτυα που έχουν εισέλθει στον κόσμο στα μέσα του 20ού αιώνα αναπτύσσονται γρήγορα. Στη σημερινή μας εποχή, έχουν εξεταστεί τα πλεονεκτήματα των τεχνητών νευρικών δικτύων και τα προβλήματα που συναντήθηκαν κατά τη χρήση τους. Δεν πρέπει να ξεχνάμε ότι τα μειονεκτήματα των δικτύων ANN, τα οποία αποτελούν αναπτυσσόμενο κλάδο της επιστήμης, εξαλείφονται ένα προς ένα και τα πλεονεκτήματά τους αυξάνονται μέρα με τη μέρα. Αυτό σημαίνει ότι τα τεχνητά νευρικά δίκτυα θα γίνουν αναπόσπαστο κομμάτι της ζωής μας μέρα με την μέρα.

1.6 Convolutional Neural Networks

Το συνελικτικό νευρωνικό δίκτυο (CNN) είναι ένας τύπος τεχνητού νευρικού δικτύου που χρησιμοποιείται στην αναγνώριση και επεξεργασία εικόνας που έχει σχεδιαστεί ειδικά για την επεξεργασία δεδομένων pixel.

Τα CNN είναι ισχυρή επεξεργασία εικόνας, τεχνητή νοημοσύνη (AI) που χρησιμοποιούν βαθιά μάθηση (deep learning) για να εκτελούν τόσο γενετικές όσο και περιγραφικές εργασίες, συχνά χρησιμοποιώντας μηχανική όραση που περιλαμβάνει αναγνώριση εικόνας και βίντεο, μαζί με συστήματα σύστασης και επεξεργασία φυσικής γλώσσας (NLP).

Ένα νευρωνικό δίκτυο είναι ένα σύστημα υλικού ή / και λογισμικού που έχει σχεδιαστεί μετά τη λειτουργία νευρώνων στον ανθρώπινο εγκέφαλο. Τα παραδοσιακά νευρωνικά δίκτυα δεν είναι ιδανικά για επεξεργασία εικόνας και πρέπει να τροφοδοτούνται εικόνες σε κομμάτια μειωμένης ανάλυσης. Το CNN έχει τους «νευρώνες» τους διευθετημένους περισσότερο όπως εκείνους του μετωπιαίου λοβού, την περιοχή που είναι υπεύθυνη για την επεξεργασία οπτικών ερεθισμάτων σε ανθρώπους και άλλα ζώα. Τα στρώματα των νευρώνων είναι διατεταγμένα με τέτοιο τρόπο ώστε να καλύπτουν ολόκληρο το οπτικό πεδίο αποφεύγοντας το αποσπασματικό πρόβλημα επεξεργασίας εικόνας των παραδοσιακών νευρικών δικτύων.

Ένα CNN χρησιμοποιεί ένα σύστημα σαν ένα πολυεπίπεδο perceptron που έχει σχεδιαστεί για μειωμένες απαιτήσεις επεξεργασίας. Τα επίπεδα ενός CNN αποτελούνται από ένα στρώμα εισόδου, ένα στρώμα εξόδου και ένα κρυφό στρώμα που περιλαμβάνει πολλαπλά συνελικτικά στρώματα, συγκεντρωτικά στρώματα, πλήρως συνδεδεμένα στρώματα και στρώματα κανονικοποίησης. Η κατάργηση των περιορισμών και η αύξηση της αποτελεσματικότητας για την επεξεργασία εικόνας οδηγεί σε ένα σύστημα που είναι πολύ πιο αποτελεσματικό.

1.7 Πλεονεκτήματα Convolutional Neural Networks

Παρακάτω γίνεται μια παρουσίαση των κυριότερων πλεονεκτημάτων των συνελικτικών νευρωνικών δικτύων.

- Το δίκτυο μπορεί να έχει οποιοδήποτε αριθμό κρυφών επιπέδων.
- Η προσθήκη ενός επιπέδου ενσωμάτωσης, όπως τα πλήρως συνδεδεμένα δίκτυα, μπορεί να είναι πολύ ισχυρή και σας επιτρέπει να συνεχίσετε να χρησιμοποιείτε τον αλγόριθμο δικτύου ταυτόχρονα.
- Η επεξεργασία φυσικής γλώσσας είναι ικανή, αλλά δεν θα νικήσει την ομιλία και την αναγνώριση εικόνας.
- Τα δίκτυα μπορούν να μάθουν να λαμβάνουν κακές αποφάσεις χωρίς υπερβολική προσαρμογή - εάν μια επανάληψη προπόνησης δεν παράγει καμία βελτίωση κατά την τελευταία, το δίκτυο θα επανέλθει στην αρχική του γραμμή. Αυτό ήταν ένα γνωστό πρόβλημα με τα μοντέλα της προηγούμενης τέχνης.
- Τα νευρωνικά δίκτυα πολλαπλών επιπέδων είναι πρακτικά για πολλές εργασίες.
- Η πολυπλοκότητα του μοντέλου, τα ποσοστά εκμάθησης, η εκτίμηση της μέγιστης πιθανότητας και η ακρίβεια των προβλέψεων είναι όλα γραμμικά με την πολυπλοκότητα του μοντέλου και η εκπαίδευση είναι συνήθως γρήγορη και εύκολη.
- Έχουν καλύτερη απόδοση σε πολλές δυνατότητες ταυτόχρονα σε μεγάλες εικόνες.
- Τα Convolutional Neural Networks δεν είναι απλά στην θεωρία, αλλά στην πράξη είναι πολύ αποτελεσματικά.
- Τα δίκτυα πολλαπλών επιπέδων είναι καλύτερα για εκπαίδευση επειδή μπορούν να επαναχρησιμοποιήσουν μεγάλο μέρος του διαθέσιμου υπολογισμού σε κάθε επίπεδο.
- Εφαρμογές που απαιτούν ένα ευρύ φάσμα μοντέλων όπως αναγνώριση προσώπου, σημασιολογική τμηματοποίηση κ.λπ., μπορούν συχνά να γίνουν μόνο με ένα μοντέλο.

- Δεν απαιτούν πολύ μεγάλα σύνολα δεδομένων.
- Μπορούν να το επιτύχουν μέσω μιας ποικιλίας αλγορίθμων μηχανικής μάθησης που είναι σχετικά εξειδικευμένες για εργασίες μηχανικής μάθησης.
- Έχουν παρόμοιες με τις διεπαφές Φυσικής Επεξεργασίας Γλώσσας (NLP) και Μηχανικής Μετάφρασης (MT) με την έννοια ότι επιτρέπουν στον χρήστη να καθορίσει ποιες δυνατότητες θα πάρει τα αποτελέσματα.
- Τα Convolutional Neural Networks μπορούν επίσης να χρησιμοποιηθούν ως παράγοντες κλιμάκωσης. Δεν λειτουργούν όπως και άλλοι τύποι δικτύου όταν το μέγεθος εισόδου μεγαλώνει πολύ.
- Είναι πολύ γρήγορα, μπορούμε να τα χρησιμοποιήσουμε για εργασίες όπως ανίχνευση προσώπων, αναγνώριση κειμένου, διάκριση αριθμών, αναγνώριση περιεχομένου, ανίχνευση προσώπου και πολλά άλλα.
- Υπάρχει ένα μεγαλύτερο εύρος πιθανών παραμέτρων συντονισμού, όπως φαίνεται από την εργασία που δείχνει ότι τα συνελκτικά δίκτυα παράγουν χρήσιμα αποτελέσματα.
- Δεν υπάρχει περιορισμός μάθησης (ακόμη και μη γραμμικότητες) στο μέγεθος του δεκτικού πεδίου, έτσι ώστε τα αποτελέσματα να μπορούν να εφαρμοστούν σε οποιοδήποτε μέγεθος πεδίου ενδιαφέροντος.
- Η χρήση των Convolutional Neural Networks (CNN) είναι εύκολη επειδή χρησιμοποιούνται ευρέως σε πολλούς άλλους τομείς. Είναι λιγότερο επιρρεπείς στο να έχουν υπερβολική τοποθέτηση (βάρους στο θόρυβο) και υπερβολική τοποθέτηση (υπερβολικό βάρος στο θόρυβο), κάτι που είναι καλό για τη μηχανική εκμάθηση.

ΕΠΙΛΟΓΟΣ

Συμπερασματικά στο παρόν και εισαγωγικό κεφάλαιο γίνεται μια εισαγωγή στα τεχνητά νευρωνικά δίκτυα και στα συνελκτικά νευρωνικά δίκτυα. Αρχικά υπάρχει μια ιστορική αναδρομή , που δείχνει πως οδηγήθηκε η ανθρωπότητα στην μελέτη του βιολογικού νευρωνικού δικτύου και έφτασε στο σημείο της δημιουργίας τεχνητών νευρωνικών δικτύων. Στη συνέχεια παρουσιάζεται το βιολογικό νευρωνικό δίκτυο και έπειτα στο επόμενο υποκεφάλαιο αναλύεται το τεχνητό νευρωνικό δίκτυο. Εν συνεχεία στο επόμενο υποκεφάλαιο παρουσιάζονται τα πλεονεκτήματα που έχουν τα τεχνητά νευρωνικά δίκτυα. Τέλος γίνεται μια σύντομη εισαγωγή στα συνελκτικά νευρωνικά δίκτυα (Convolution Neural Networks) που θα μας απασχολήσουν στην παρούσα πτυχιακή εργασία.

ΚΕΦΑΛΑΙΟ 2 - Residual Networks (ResNet)

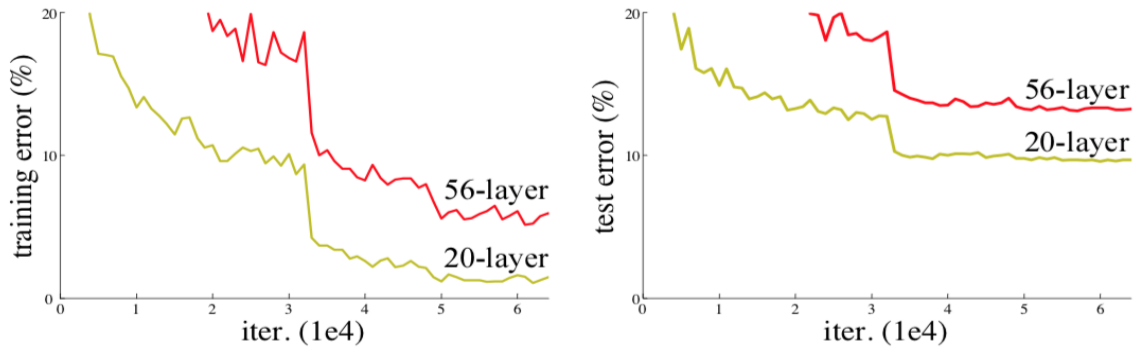
2.1 ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια, τα νευρωνικά δίκτυα έχουν γίνει βαθύτερα, με τα υπερσύγχρονα δίκτυα να ξεκινούν από λίγα επίπεδα (π.χ. AlexNet) σε πάνω από εκατό επίπεδα.

Ένα από τα σημαντικότερα οφέλη ενός πολύ βαθύ δικτύου είναι ότι μπορεί να αντιπροσωπεύει πολύπλοκες λειτουργίες. Ωστόσο, ένα τεράστιο εμπόδιο για την εκπαίδευσή τους είναι η εξαφάνιση των διαβαθμίσεων: τα πολύ βαθιά δίκτυα έχουν συχνά ένα σήμα διαβάθμισης που πηγαίνει στο μηδέν γρήγορα, κάνοντας έτσι την κάθοδο της κλίσης απαγορευτικά αργή. Πιο συγκεκριμένα, κατά την κάθοδο της κλίσης, καθώς προχωρούμε πίσω από το τελικό στρώμα πίσω στο πρώτο στρώμα, πολλαπλασιάζουμε με τη μήτρα βάρους σε κάθε βήμα. Εάν οι κλίσεις είναι μικρές, λόγω μεγάλου αριθμού πολλαπλασιασμών, η κλίση μπορεί να μειωθεί εκθετικά γρήγορα στο μηδέν (ή, σε σπάνιες περιπτώσεις, να αυξηθεί εκθετικά γρήγορα και να «εκραγεί» για να πάρει πολύ μεγάλες τιμές).

Τα κανονικά δίκτυα όπως το VGG-16 ονομάζονται «απλά» δίκτυα.

Σε απλά δίκτυα, καθώς ο αριθμός των επιπέδων αυξάνεται από 20 σε 56 (όπως φαίνεται παρακάτω), ακόμη και μετά από χιλιάδες επαναλήψεις, το σφάλμα εκπαίδευσης ήταν χειρότερο για ένα επίπεδο 56 σε σύγκριση με ένα δίκτυο 20 επιπέδων.



Σχήμα 5: Training Error and Test Error on CIFAR-10 with 20-layer and 56-layer "plain" networks

Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer "plain" networks. The deeper network has higher training error, and thus test error.

Θεωρητικά, περιμένουμε ότι ένα βαθύτερο δίκτυο θα βοηθούσε μόνο, αλλά στην πραγματικότητα, το βαθύτερο δίκτυο έχει υψηλότερο σφάλμα εκπαίδευσης.

Όταν τα βαθύτερα δίκτυα είναι σε θέση να αρχίσουν να συγκλίνουν, ένα πρόβλημα υποβάθμισης έχει εκτεθεί: με το βάθος του δικτύου να αυξάνεται, η ακρίβεια γίνεται κορεσμένη (κάτι που μπορεί να μην προκαλεί έκπληξη) και στη συνέχεια υποβαθμίζεται γρήγορα.

Η χρήση βαθύτερων δικτύων υποβαθμίζει την απόδοση του μοντέλου. Το έγγραφο Microsoft Research προσπαθεί να λύσει αυτό το πρόβλημα χρησιμοποιώντας Deep Residual Learning .

2.1 Residual Network (ResNet)

Τα υπολειπόμενα δίκτυα (**Residual Networks**) είναι νευρωνικά δίκτυα στα οποία κάθε στρώμα αποτελείται από μια υπολειμματική μονάδα f_i και μια παράλειψη σύνδεσης παρακάμπτοντας το f_i . Δεδομένου ότι τα επίπεδα σε υπολειμματικά δίκτυα μπορούν να περιλαμβάνουν πολλαπλά συνελκτικά στρώματα, τα αναφερόμαστε ως υπολειμματικά μπλοκ στο υπόλοιπο αυτής της εργασίας. Για λόγους σαφήνειας, παραλείπουμε τα αρχικά βήματα προεπεξεργασίας και τελικής ταξινόμησης. Με y_{i-1} όπως είναι η είσοδος, η έξοδος του μπλοκ ορίζεται αναδρομικά ως :

$$y_i \equiv f_i(y_{i-1}) + y_{i-1}, \quad (1)$$

όπου $f_i(x)$ είναι κάποια ακολουθία συνεπειών, ομαλοποίηση παρτίδας και Rectified Linear Units (ReLU) ως μη γραμμικότητες. Το σχήμα 1 (α) δείχνει μια σχηματική όψη αυτής της αρχιτεκτονικής. Στην πιο πρόσφατη διαμόρφωση υπολειμματικών δικτύων [6], το $f_i(x)$ ορίζεται από :

$$f_i(x) \equiv W_i \cdot \sigma(B(W'_i \cdot \sigma(B(x))))), \quad (2)$$

όπου τα W είναι πίνακες βάρους, το $B(x)$ είναι ομαλοποίηση παρτίδας και $\sigma(x) \equiv \max(x, 0)$. Άλλες τυποποιήσεις αποτελούνται συνήθως από τις ίδιες λειτουργίες, αλλά μπορεί να διαφέρουν στη σειρά τους.

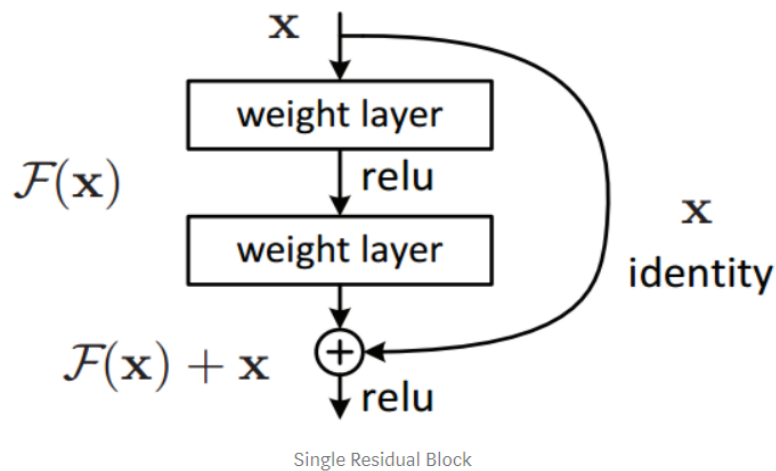
Η ιδέα της διακλάδωσης των διαδρομών στα νευρικά δίκτυα δεν είναι καινούργια. Για παράδειγμα, στο καθεστώς των συνελκτικών νευρικών δικτύων, τα μοντέλα που βασίζονται σε μονάδες έναρξης ήταν από τα πρώτα που τακτοποίησαν στρώματα σε μπλοκ με παράλληλες διαδρομές και όχι αυστηρή διαδοχική σειρά. Επιλέγουμε υπολειμματικά δίκτυα για αυτήν τη μελέτη λόγω της απλής αρχής σχεδιασμού τους.

Ποια είναι η ανάγκη για υπολειμματική μάθηση (Residual learning) ;

Τα βαθιά συνελκτικά νευρωνικά δίκτυα έχουν οδηγήσει σε μια σειρά ανακαλύψεων για την ταξινόμηση της εικόνας. Πολλές άλλες εργασίες οπτικής αναγνώρισης έχουν επίσης επωφεληθεί σε μεγάλο βαθμό από πολύ βαθιά μοντέλα. Έτσι, με την πάροδο των ετών υπάρχει μια τάση να πηγαίνουμε πιο βαθιά, να επιλύουμε πιο περίπλοκα καθήκοντα και επίσης να αυξάνουμε / βελτιώνουμε την ακρίβεια της ταξινόμησης / αναγνώρισης. Όμως, καθώς πηγαίνουμε βαθύτερα, η εκπαίδευση του νευρικού δικτύου γίνεται δύσκολη και επίσης η ακρίβεια αρχίζει να κορεστεί και στη συνέχεια υποβαθμίζεται επίσης. Το Residual Learning προσπαθεί να λύσει και τα δύο αυτά προβλήματα.

2.2 Residual μπλοκ - Δομικά στοιχεία του ResNet

Η κατανόηση ενός υπολειμματικού μπλοκ (residual block) είναι αρκετά εύκολη. Στα παραδοσιακά νευρωνικά δίκτυα, κάθε στρώμα τροφοδοτείται στο επόμενο στρώμα. Σε ένα δίκτυο με υπολειμματικά μπλοκ, κάθε στρώμα τροφοδοτείται στο επόμενο στρώμα και απευθείας στα επίπεδα περίπου 2-3 λυκίσκους (hops) μακριά. Όμως, κατανοώντας τη διαίσθηση πίσω από το γιατί ήταν απαραίτητο, γιατί είναι τόσο σημαντικό και πόσο παρόμοιο με κάποιες άλλες αρχιτεκτονικές τελευταίας τεχνολογίας είναι εκεί όπου πρόκειται να επικεντρωθούμε. Υπάρχουν περισσότερες από μία ερμηνείες για το γιατί τα υπολειμματικά μπλοκ είναι καταπληκτικά και πώς & γιατί είναι μία από τις βασικές ιδέες που μπορούν να κάνουν μια νευρωνική εμφάνιση δικτύου για τις τελευταίες εξελίξεις σε ένα ευρύ φάσμα εργασιών. Η εικόνα για το πως μοιάζει ένα υπολειμματικό μπλοκ παρουσιάζεται παρακάτω:



Σχήμα 6: Single Residual Block

Η προσθήκη επιπρόσθετων / νέων επιπέδων δεν θα βλάψει την απόδοση του μοντέλου, καθώς η κανονικοποίηση θα τα παρακάμψει εάν αυτά τα επίπεδα δεν ήταν χρήσιμα.

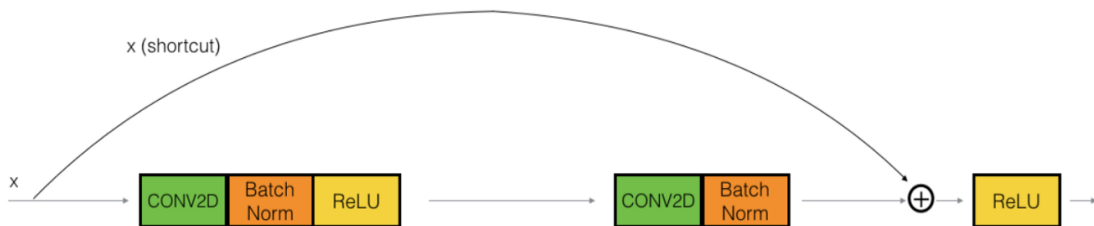
Εάν τα πρόσθετα / νέα στρώματα ήταν χρήσιμα, ακόμη και με την παρουσία κανονικοποίησης, τα βάρη ή οι πυρήνες των στρωμάτων θα είναι μηδενικά και η απόδοση του μοντέλου θα μπορούσε να αυξηθεί ελαφρώς.

Επομένως, με την προσθήκη νέων επιπέδων, λόγω της "Παράλειψη σύνδεσης" / "υπολειπόμενη σύνδεση", είναι εγγυημένο ότι η απόδοση του μοντέλου δεν μειώνεται αλλά θα μπορούσε να αυξηθεί ελαφρώς.

Τοποθετώντας αυτά τα μπλοκ ResNet το ένα πάνω στο άλλο, μπορείτε να δημιουργήσετε ένα πολύ βαθύ δίκτυο. Έχοντας μπλοκ ResNet με τη συντόμευση καθιστά επίσης πολύ εύκολο για ένα από τα μπλοκ να μάθει μια συνάρτηση ταυτότητας.

Δύο βασικοί τύποι μπλοκ χρησιμοποιούνται σε ένα ResNet, ανάλογα κυρίως με το αν οι διαστάσεις εισόδου / εξόδου είναι ίδιες ή διαφορετικές.

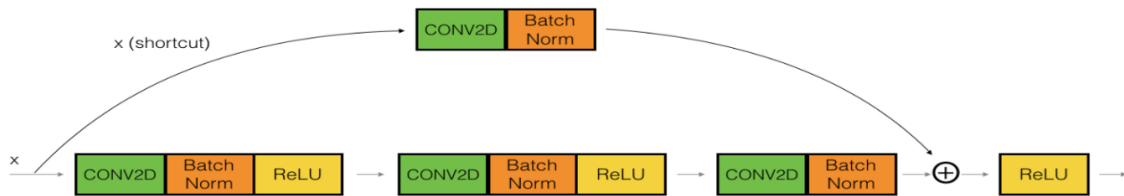
1. Το μπλοκ ταυτότητας - ίδιο με αυτό που είδαμε παραπάνω. Το μπλοκ ταυτότητας είναι το τυπικό μπλοκ που χρησιμοποιείται στο ResNets και αντιστοιχεί στην περίπτωση όπου η ενεργοποίηση εισόδου έχει την ίδια διάσταση με την ενεργοποίηση εξόδου.



Σχήμα 7: Identity block

Identity block. Skip connection “skips over” 2 layers

2. Το Convolutional block - Μπορούμε να χρησιμοποιήσουμε αυτόν τον τύπο μπλοκ όταν οι διαστάσεις εισόδου και εξόδου δεν ταιριάζουν. Η διαφορά με το μπλοκ ταυτότητας είναι ότι υπάρχει ένα επίπεδο CONV2D στη διαδρομή συντόμευσης.



Σχήμα 8: Convolutional block

Convolutional block

Όπως είναι γνωστό τα νευρωνικά δίκτυα είναι προσεγγιστικά καθολικής λειτουργίας και η ακρίβεια αυξάνεται με τον αυξανόμενο αριθμό επιπέδων. Αλλά υπάρχει ένα όριο στον αριθμό των επιπέδων που προστίθενται που έχουν ως αποτέλεσμα τη βελτίωση της ακρίβειας. Έτσι, εάν τα νευρωνικά δίκτυα ήταν καθολικοί συντελεστές λειτουργίας τότε θα έπρεπε να ήταν σε θέση να μάθουν οποιαδήποτε απλή ή πολύπλοκη λειτουργία. Αλλά αποδεικνύεται ότι, χάρη σε ορισμένα προβλήματα, όπως η εξαφάνιση των διαβαθμίσεων και η κατάρρα της διαστατικότητας, εάν έχουμε αρκετά βαθιά δίκτυα, μπορεί να μην είναι σε θέση να μάθουν απλές λειτουργίες όπως μια λειτουργία ταυτότητας. Τώρα αυτό είναι σαφώς ανεπιθύμητο.

Επίσης, εάν συνεχιστεί να αυξάνεται ο αριθμός των επιπέδων, θα φανεί ότι η ακρίβεια θα αρχίσει να κορεστεί σε ένα σημείο και τελικά θα υποβαθμιστεί. Και αυτό συνήθως δεν προκαλείται λόγω υπερβολικής τοποθέτησης. Άρα, φαίνεται ότι τα ρηχά δίκτυα μαθαίνουν καλύτερα από τα βαθύτερα αντίστοιχά τους και αυτό είναι αρκετά αντιδραστικό. Αυτό όμως φαίνεται στην πράξη και είναι γνωστό ως το πρόβλημα της υποβάθμισης.

Χωρίς ρίζα που προκαλεί το πρόβλημα της υποβάθμισης και την αδυναμία ενός βαθιού νευρικού δικτύου να μάθει λειτουργίες ταυτότητας, πρέπει να αρχίσουν να βρίσκονται μερικές λύσεις. Σε πρόβλημα υποβάθμισης, δεδομένο είναι ότι τα ρηγά δίκτυα αποδίδουν καλύτερα από τα βαθύτερα αντίστοιχα που έχουν προσθέσει λίγα περισσότερα επίπεδα σε αυτά. Τότε μια λύση θα μπορούσε να είναι η παράλειψη αυτών των επιπλέον επιπέδων, το θέμα συνεπώς είναι πως θα γίνει αυτή η παράλειψη.

Μια λύση για αυτά τα προβλήματα είναι η παράλειψη της εκπαίδευσης λίγων επιπέδων χρησιμοποιώντας παραλείψεις – συνδέσεις ή υπολειπόμενες συνδέσεις. Αυτό φαίνεται στην παραπάνω εικόνα. Στην πραγματικότητα, μάθουμε άμεσα μια συνάρτηση ταυτότητας βασιζόμενοι μόνο σε παραλείψεις συνδέσεων. Αυτός είναι ο ακριβής λόγος για τον οποίο οι παραλείψεις συνδέσεων καλούνται επίσης ως συνδέσεις συντόμευσης ταυτότητας.

Αν εξεταστεί ένα μπλοκ νευρικού δικτύου, του οποίου η είσοδος είναι x και θα θέλαμε να μάθουμε την πραγματική κατανομή $H(x)$. Η διαφορά (ή το υπόλοιπο) μεταξύ αυτού είναι :

$$R(x) = \text{Output} - \text{Input} = H(x) - x$$

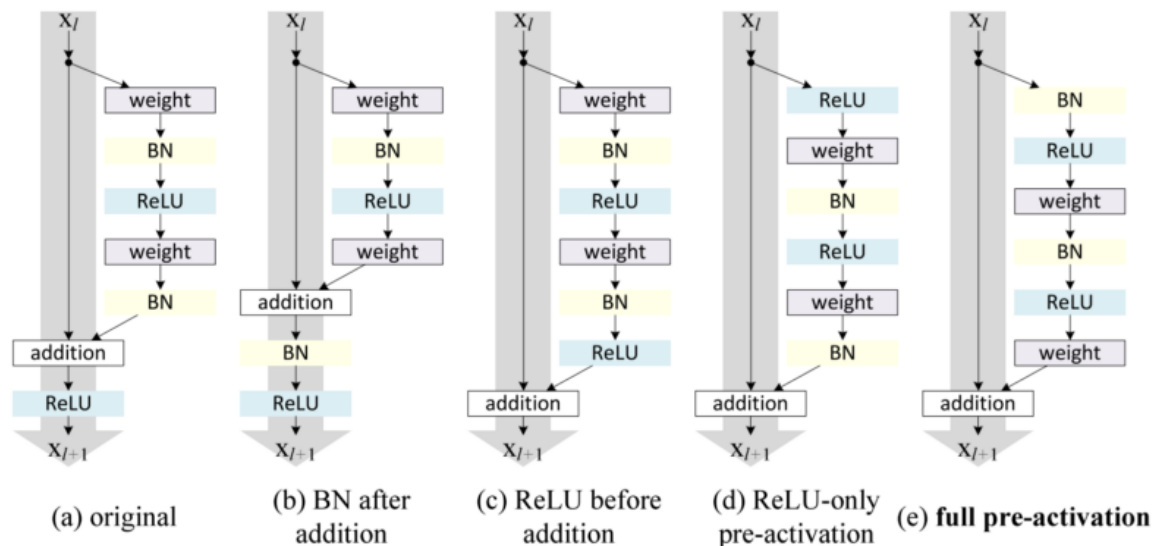
και με αναδιάταξη είναι :

$$H(x) = R(x) + x$$

Το υπολειμματικό (residual) μπλοκ προσπαθεί να μάθει την πραγματική έξοδο, $H(x)$ και δεδομένου ότι υπάρχει μια σύνδεση ταυτότητας που οφείλεται στο x , τα στρώματα προσπαθούν πραγματικά να μάθουν το υπόλοιπο, $R(x)$. Συνοψίζοντας, τα επίπεδα σε ένα παραδοσιακό δίκτυο μαθαίνουν την πραγματική έξοδο ($H(x)$) ενώ τα επίπεδα σε ένα υπολειμματικό (residual) δίκτυο μαθαίνουν τα υπόλοιπα ($R(x)$). Ως εκ τούτου, το όνομα: Residual Block.

Έχει επίσης παρατηρηθεί ότι είναι ευκολότερη η μάθηση των υπολοίπων της εξόδου και της εισόδου, παρά μόνο της εισόδου. Ως πρόσθετο πλεονέκτημα, το δίκτυο μπορεί τώρα να μάθει τη λειτουργία ταυτότητας ορίζοντας απλώς το υπόλοιπο ως μηδέν. Λόγω των συνδέσεων παράλειψης, υπάρχει η δυνατότητα να διαδοθούν μεγαλύτερες κλίσεις στα αρχικά επίπεδα, και συνεπώς αυτά τα επίπεδα θα μπορούσαν επίσης να μάθουν τόσο γρήγορα όσο τα τελικά στρώματα, δίνοντας έτσι την δυνατότητα για εκπαίδευση βαθύτερων δικτύων.

Και αν καταλαβαίνετε αληθινά την αναδρομή και πόσο σοβαρό γίνεται το πρόβλημα της εξαφάνισης της κλίσης με τον αυξανόμενο αριθμό επιπέδων, τότε μπορείτε να δείτε ξεκάθαρα ότι λόγω αυτών των συνδέσεων παράλειψης, μπορούμε να διαδώσουμε μεγαλύτερες κλίσεις στα αρχικά επίπεδα και αυτά τα επίπεδα θα μπορούσαν επίσης να μάθουν τόσο γρήγορα όσο τα τελικά στρώματα, δίνοντάς μας τη δυνατότητα να εκπαιδεύσουμε βαθύτερα δίκτυα.



Σχήμα 9: Types of Residual Block

Η παρακάτω εικόνα δείχνει πώς μπορούμε να τακτοποιήσουμε το υπόλοιπο μπλοκ και τις συνδέσεις ταυτότητας για τη βέλτιστη ροή κλίσης. Έχει παρατηρηθεί ότι οι προ-ενεργοποιήσεις με ομαλοποιήσεις παρτίδων δίνουν γενικά τα καλύτερα αποτελέσματα (δηλ. Το δεξί- υπολειμματικό (residual) μπλοκ στην παρακάτω εικόνα δίνει τα πολλά υποσχόμενα αποτελέσματα).

Υπάρχουν ακόμη περισσότερες ερμηνείες των υπολειμματικών μπλοκ και του ResNets, εκτός από αυτές που συζητήθηκαν παραπάνω. Κατά την εκπαίδευση του ResNets, είτε εκπαιδύουμε τα επίπεδα σε υπολειμματικά μπλοκ είτε παραλείπουμε την εκπαίδευση για αυτά τα επίπεδα χρησιμοποιώντας συνδέσεις παράλειψης. Έτσι, για διαφορετικά σημεία δεδομένων εκπαίδευσης, διαφορετικά τμήματα των δικτύων θα εκπαιδευτούν με διαφορετικούς ρυθμούς ανάλογα με το πώς το σφάλμα ρέει προς τα πίσω στο δίκτυο. Αυτό μπορεί να θεωρηθεί ότι εκπαιδύει ένα σύνολο διαφορετικών μοντέλων στο σύνολο δεδομένων και εξασφαλίζει την καλύτερη δυνατή ακρίβεια.

Η παράλειψη της προπόνησης σε ορισμένα υπολειπόμενα στρώματα μπλοκ μπορεί να εξεταστεί και από μια αισιόδοξη άποψη. Γενικά, δεν γνωρίζουμε τον βέλτιστο αριθμό επιπέδων (ή υπολειμματικών μπλοκ) που απαιτούνται για ένα νευρωνικό δίκτυο που μπορεί να εξαρτάται από την πολυπλοκότητα του συνόλου δεδομένων. Αντί να αντιμετωπίζουμε τον αριθμό των επιπέδων ένα σημαντικός υπερπαραμέτρος για συντονισμό, προσθέτοντας παράλειψη συνδέσεων στο δίκτυό μας, επιτρέπουμε στο δίκτυο να παραλείψει την εκπαίδευση για τα επίπεδα που δεν είναι χρήσιμα και δεν προσθέτουν αξία στη συνολική ακρίβεια. Κατά κάποιον τρόπο, η παράλειψη συνδέσεων κάνει τα νευρωνικά δίκτυα δυναμικά, έτσι ώστε να βελτιστοποιεί τον αριθμό των επιπέδων κατά τη διάρκεια της προπόνησης.

Η παρακάτω εικόνα δείχνει πολλαπλές ερμηνείες ενός υπολειμματικού (residual) μπλοκ.

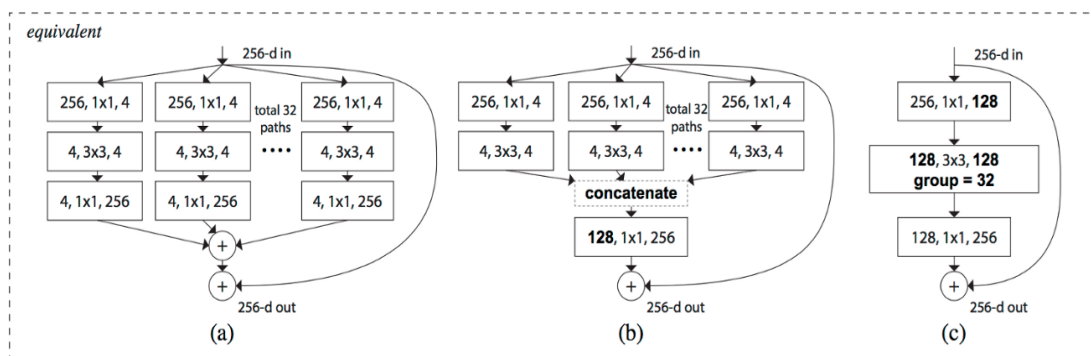


Figure 3. Equivalent building blocks of ResNeXt. (a): Aggregated residual transformations, the same as Fig. 1 right. (b): A block equivalent to (a), implemented as early concatenation. (c): A block equivalent to (a,b), implemented as grouped convolutions [24]. Notations in **bold** text highlight the reformulation changes. A layer is denoted as (# input channels, filter size, # output channels).

Σχήμα 10: Διαφορετικές ερμηνείες του Residual Block

Η ιδέα της παράλειψης συνδέσεων μεταξύ των επιπέδων παρουσιάστηκε για πρώτη φορά στα Highway Networks. Τα δίκτυα αυτοκινητόδρομων είχαν παραλείψει συνδέσεις με πύλες που ελέγχουν πόσες πληροφορίες περνούν μέσω αυτών και αυτές οι πύλες μπορούν να εκπαιδευτούν για να ανοίγουν επιλεκτικά. Αυτή η ιδέα φαίνεται επίσης στα δίκτυα LSTM που ελέγχουν πόσες πληροφορίες ρέουν από τα προηγούμενα σημεία δεδομένων που έχουν δει το δίκτυο.

Τα υπόλοιπα μπλοκ είναι βασικά μια ειδική περίπτωση δικτύων αυτοκινητόδρομων χωρίς πύλες στις συνδέσεις παράλειψης. Ουσιαστικά, τα υπόλοιπα μπλοκ επιτρέπουν τη ροή της μνήμης (ή των πληροφοριών) από τα αρχικά στρώματα στα τελευταία στρώματα. Παρά την απουσία πυλών στις συνδέσεις παράλειψης, τα υπόλοιπα δίκτυα αποδίδουν τόσο καλά όσο οποιοδήποτε άλλο δίκτυο αυτοκινητοδρόμων στην πράξη.

2.3 Το ResNet και οι παραλλαγές του

Μετά τη διάσημη νίκη του AlexNet στο διαγωνισμό κατάταξης LSVRC2012, το Deep Residual Network ήταν αναμφισβήτητο το πιο πρωτοποριακό έργο στην κοινότητα του οράματος / της βαθιάς μάθησης των υπολογιστών τα τελευταία χρόνια. Το ResNet καθιστά δυνατή την εκπαίδευση έως και εκατοντάδων ή και χιλιάδων επιπέδων και επιτυγχάνει ακόμα εντυπωσιακή απόδοση.

Αξιοποιώντας την ισχυρή αντιπροσωπευτική του ικανότητα, ενισχύθηκε η απόδοση πολλών εφαρμογών όρασης υπολογιστή εκτός από την ταξινόμηση εικόνας, όπως η ανίχνευση αντικειμένων και η αναγνώριση προσώπου.

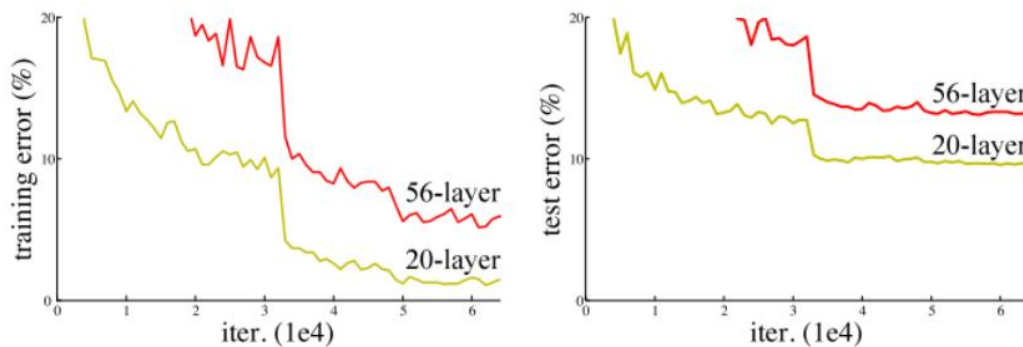
Από τότε που το ResNet μπήκε στις ζώες των ανθρώπων το 2015, πολλοί στην ερευνητική κοινότητα έχουν βυθιστεί στα μυστικά της επιτυχίας του, έχουν γίνει πολλές βελτιώσεις στην αρχιτεκτονική. Αυτό το υποκεφάλαιο χωρίζεται σε δύο μέρη, στο πρώτο μέρος θα γίνει προσπάθεια για εξοικείωση με το ResNet, στο δεύτερο θα εξετάστούν μερικά από τα διαφορετικές παραλλαγές και ερμηνείες την αρχιτεκτονική ResNet.

Επανεξέταση του ResNet

Σύμφωνα με το θεώρημα της καθολικής προσέγγισης, δεδομένης της αρκετής χωρητικότητας, γνωρίζουμε ότι ένα δίκτυο τροφοδοσίας με ένα μόνο στρώμα είναι αρκετό για να αντιπροσωπεύει οποιαδήποτε λειτουργία. Ωστόσο, το επίπεδο μπορεί να είναι τεράστιο και το δίκτυο είναι επιρρεπές σε υπερβολική προσαρμογή των δεδομένων. Επομένως, υπάρχει μια κοινή τάση στην ερευνητική κοινότητα ότι η αρχιτεκτονική του δικτύου πρέπει να προχωρήσει βαθύτερα.

Από το AlexNet, η υπερσύγχρονη αρχιτεκτονική CNN πηγαίνει όλο και πιο βαθιά. Ενώ το AlexNet είχε μόνο 5 συνεχόμενα επίπεδα, το δίκτυο VGG και το GoogleNet (επίσης με κωδικό όνομα Inception_v1) είχαν 19 και 22 επίπεδα αντίστοιχα.

Ωστόσο, η αύξηση του βάθους δικτύου δεν λειτουργεί απλώς με το στοίβαγμα επιπέδων μαζί. Τα βαθιά δίκτυα είναι δύσκολο να εκπαιδευτούν εξαιτίας του διαβόητου προβλήματος της διαβάθμισης που εξαφανίζεται - καθώς η κλίση αναπαράγεται πίσω σε προηγούμενα στρώματα, ο επαναλαμβανόμενος πολλαπλασιασμός μπορεί να κάνει την κλίση απεριόριστα μικρή. Ως αποτέλεσμα, καθώς το δίκτυο πηγαίνει βαθύτερα, η απόδοσή του εμφανίζει κορεσμό ή ακόμη και αρχίζει να υποβαθμίζεται γρήγορα.

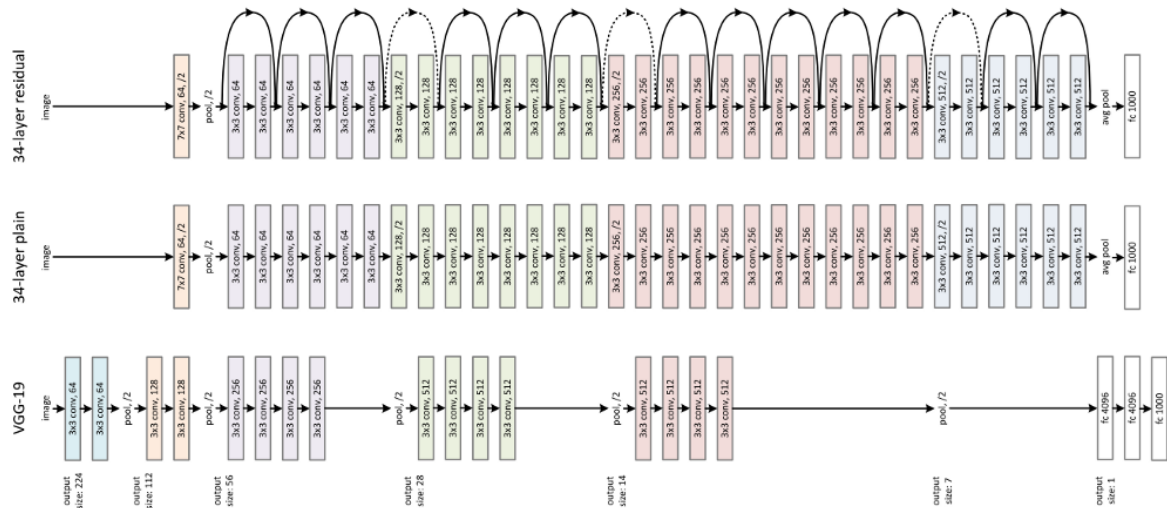


increasing network depth leads to worse performance

Σχήμα 11: Increasing network depth leads to worse performance

Πριν από το ResNet, υπήρχαν αρκετοί τρόποι για να αντιμετωπιστεί το ζήτημα της διαβάθμισης που εξαφανίστηκε, για παράδειγμα, προσθέτοντας μια βοηθητική απώλεια σε ένα μεσαίο στρώμα ως επιπλέον επίβλεψη, αλλά κανένας τρόπος δεν φάνηκε να αντιμετωπίζει πραγματικά το πρόβλημα μια για πάντα.

Η βασική ιδέα του ResNet είναι η εισαγωγή μιας λεγόμενης «σύνδεσης συντόμευσης ταυτότητας» που παραλείπει ένα ή περισσότερα επίπεδα, όπως φαίνεται στο παρακάτω σχήμα:



Σχήμα 12: Residual Architecture

Οι συγγραφείς του [2] υποστηρίζουν ότι η στοιβαξη επιπέδων δεν πρέπει να υποβαθμίσει την απόδοση του δικτύου, επειδή θα μπορούσαμε απλώς να στοιβαξουμε αντιστοιχισεις ταυτότητας (επίπεδο που δεν κάνει τίποτα) στο τρέχον δίκτυο και η προκύπτουσα αρχιτεκτονική θα είχε την ίδια απόδοση. Αυτό υποδηλώνει ότι το βαθύτερο μοντέλο δεν πρέπει να προκαλεί σφάλμα εκπαίδευσης υψηλότερο από τα ρηγά αντίστοιχα. Υποθέτουν ότι το να αφήσετε τα στοιβαγμένα στρώματα να ταιριάζουν σε εναπομένονσα χαρτογράφηση είναι ευκολότερο από το να τα αφήσετε να ταιριάζουν απευθείας στην επιθυμητή χαρτογράφηση υποστρώματος. Και το υπόλοιπο μπλοκ παραπάνω επιτρέπει ρηγά να το κάνει ακριβώς αυτό.

Στην πραγματικότητα, το ResNet δεν ήταν το πρώτο δίκτυο που έκανε χρήση συνδέσεων συντόμευσης, το Highway Network εισήγαγε πύλες συντομεύσεων. Αυτές οι παραμετροποιημένες πύλες ελέγχουν πόσες πληροφορίες επιτρέπεται να ρέουν κατά μήκος της συντόμευσης. Παρόμοια ιδέα μπορεί να βρεθεί στο κύτταρο Long Term Short Memory (LSTM), στο οποίο υπάρχει μια παραμετροποιημένη πύλη (forget gate) που ελέγχει πόσες πληροφορίες θα ρέουν στο επόμενο βήμα. Επομένως, το ResNet μπορεί να θεωρηθεί ως ειδική περίπτωση του Highway Network.

Ωστόσο, τα πειράματα δείχνουν ότι το Highway Network δεν αποδίδει καλύτερα από το ResNet, κάτι που είναι περίεργο επειδή ο χώρος λύσης του Highway Network περιέχει ResNet, επομένως θα πρέπει να είναι τουλάχιστον το ίδιο καλό με το ResNet. Αυτό υποδηλώνει ότι είναι πιο σημαντικό να διατηρηθούν αυτοί οι «αυτοκινητόδρομοι κλίσης» καθαροί από το να αναζητηθεί μεγαλύτερος χώρος λύσης.

Μετά από αυτή τη διαίσθηση, οι συγγραφείς του [2] βελτίωσαν το υπόλοιπο μπλοκ και πρότειναν μια παραλλαγή προ-ενεργοποίησης του υπολειμματικού μπλοκ, στην οποία οι κλίσεις μπορούν να ρέουν μέσω των συνδέσεων συντόμευσης σε οποιοδήποτε άλλο προηγούμενο στρώμα ανεμπόδιστα. Στην πραγματικότητα, χρησιμοποιώντας το αρχικό υπόλοιπο μπλοκ στο [2], η κατάρτιση ενός ResNet 1202 επιπέδων είχε ως αποτέλεσμα χειρότερη απόδοση από το αντίστοιχο των 110 επιπέδων.

Οι συγγραφείς του [7] απέδειξαν με πειράματα ότι τώρα μπορούν να εκπαιδεύσουν ένα ResNet βάθους 1001 επιπέδων για να ξεπεράσουν τα ρηγά αντίστοιχα. Λόγω των συναρπαστικών αποτελεσμάτων του, το ResNet έγινε γρήγορα μια από τις πιο δημοφιλείς αρχιτεκτονικές σε διάφορες εργασίες όρασης υπολογιστή.

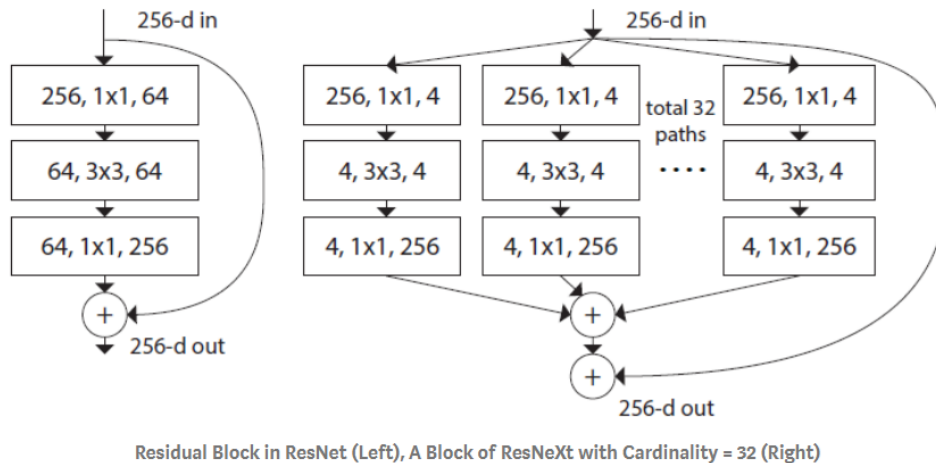
Πρόσφατες παραλλαγές και ερμηνείες του ResNet

Καθώς το ResNet αποκτά όλο και μεγαλύτερη δημοτικότητα στην ερευνητική κοινότητα, η αρχιτεκτονική του μελετάται σε μεγάλο βαθμό. Σε αυτήν την υποενότητα, θα παρουσιαστούν πρώτα αρκετές νέες αρχιτεκτονικές που βασίζονται στο ResNet και μετά θα παρουσιαστεί ένα έγγραφο που παρέχει μια ερμηνεία της αντιμετώπισης του ResNet ως συνόλου πολλών μικρότερων δικτύων.

2.3.1 ResNeXt

Οι Xie et al. [8] πρότεινε μια παραλλαγή του ResNet που κωδικοποιείται ResNeXt με το ακόλουθο δομικό στοιχείο:

Το όνομα του μοντέλου, ResNeXt, περιέχει το Next. Σημαίνει την επόμενη διάσταση, πάνω από το ResNet. Αυτή η επόμενη διάσταση ονομάζεται διάσταση «καρδινιτότητα». Και το ResNeXt γίνεται ο 1ος Runner Up της εργασίας ταξινόμησης ILSVRC.



Σχήμα 13: Residual Block and ResNeXt Block

Σε σύγκριση με το ResNet (Ο νικητής στο ILSVRC 2015, 3,57%) και το PolyNet (2ος Runner Up, 3,04%, Όνομα ομάδας CU-DeepLink), το ResNeXt πήρε 3,03% ποσοστό σφάλματος Top-5, το οποίο είναι μια μεγάλη σχετική βελτίωση περίπου 15%.

Densely Connected CNN

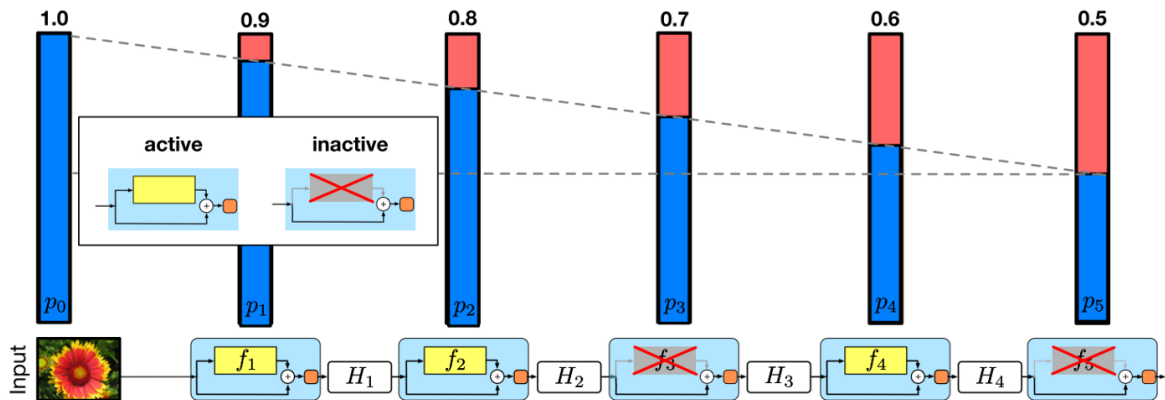
Οι Huang et al. [9] πρότεινε μια νέα αρχιτεκτονική που ονομάζεται DenseNet που εκμεταλλεύεται περαιτέρω τα αποτελέσματα των συνδέσεων συντόμευσης - συνδέει όλα τα επίπεδα απευθείας μεταξύ τους. Σε αυτήν τη νέα αρχιτεκτονική, η είσοδος κάθε επιπέδου αποτελείται από τους χάρτες χαρακτηριστικών όλων των προηγούμενων επιπέδων και η έξοδος του μεταφέρεται σε κάθε επόμενο επίπεδο. Οι χάρτες χαρακτηριστικών συγκεντρώνονται με βάθος-συνένωση.

Εκτός από την αντιμετώπιση του προβλήματος των βαθμίδων που εξαφανίζονται, οι συγγραφείς του [8] υποστηρίζουν ότι αυτή η αρχιτεκτονική ενθαρρύνει επίσης την επαναχρησιμοποίηση χαρακτηριστικών, καθιστώντας το δίκτυο εξαιρετικά αποδοτικό στις παραμέτρους. Μια απλή ερμηνεία αυτού είναι ότι, το αποτέλεσμα της χαρτογράφησης ταυτότητας προστέθηκε στο επόμενο μπλοκ, το οποίο θα μπορούσε να εμποδίσει τη ροή πληροφοριών εάν οι χάρτες χαρακτηριστικών δύο επιπέδων έχουν πολύ διαφορετικές κατανομές. Επομένως, οι συνδυαστικοί χάρτες χαρακτηριστικών μπορούν να τους διατηρήσουν όλους και να αυξήσουν τη διακύμανση των εξόδων, ενθαρρύνοντας την επαναχρησιμοποίηση χαρακτηριστικών.

2.3.2 Deep Network with Stochastic Depth

Αν και το ResNet έχει αποδειχθεί ισχυρό σε πολλές εφαρμογές, ένα σημαντικό μειονέκτημα είναι ότι το βαθύτερο δίκτυο απαιτεί συνήθως εβδομάδες για εκπαίδευση, καθιστώντας το πρακτικά ανέφικτο σε πραγματικές εφαρμογές. Για να αντιμετωπιστεί αυτό το ζήτημα, οι Huang et al. [10] εισήγαγε μια μέθοδο για την τυχαία πτώση των επιπέδων κατά τη διάρκεια της προπόνησης και τη χρήση του πλήρους δικτύου στις δοκιμές.

Οι συγγραφείς χρησιμοποίησαν το υπολειμματικό μπλοκ ως δομικό στοιχείο του δικτύου τους, επομένως, κατά τη διάρκεια της εκπαίδευσης, όταν είναι ενεργοποιημένο ένα συγκεκριμένο υπολειμματικό μπλοκ, η είσοδος του ρέει τόσο από τη συντόμευση ταυτότητας όσο και από τα επίπεδα βάρους, διαφορετικά η είσοδος ρέει μόνο μέσω της συντόμευσης ταυτότητας. Κατά την προπόνηση, κάθε στρώμα έχει «πιθανότητα επιβίωσης» και πέφτει τυχαία. Στον χρόνο δοκιμής, όλα τα μπλοκ διατηρούνται ενεργά και βαθμονομούνται εκ νέου σύμφωνα με την πιθανότητα επιβίωσής του κατά τη διάρκεια της προπόνησης.



Σχήμα 14: Probability of being disabled during training, each layer has a probability of being disabled

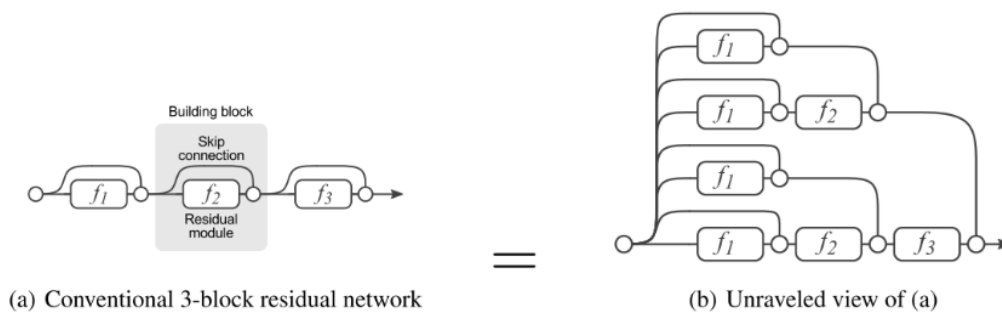
Παρόμοια με το Dropout [11], η εκπαίδευση ενός βαθιού δικτύου με στοχαστικό βάθος μπορεί να θεωρηθεί ως εκπαίδευση ενός συνόλου πολλών μικρότερων ResNets. Η διαφορά είναι ότι αυτή η μέθοδος ρίχνει τυχαία ένα ολόκληρο επίπεδο ενώ το Dropout ρίχνει μόνο μέρος των κρυφών μονάδων σε ένα επίπεδο κατά τη διάρκεια της προπόνησης.

Τα πειράματα δείχνουν ότι η εκπαίδευση ενός ResNet 110 στρωμάτων με στοχαστικό βάθος οδηγεί σε καλύτερη απόδοση από την προπόνηση ενός ResNet 110 στρωμάτων σταθερού βάθους, ενώ μειώνει δραματικά τον χρόνο προπόνησης. Αυτό υποδηλώνει ότι ορισμένα από τα επίπεδα (διαδρομές) στο ResNet ενδέχεται να είναι περιττά.

ResNet ως σύνολο μικρότερων δικτύων

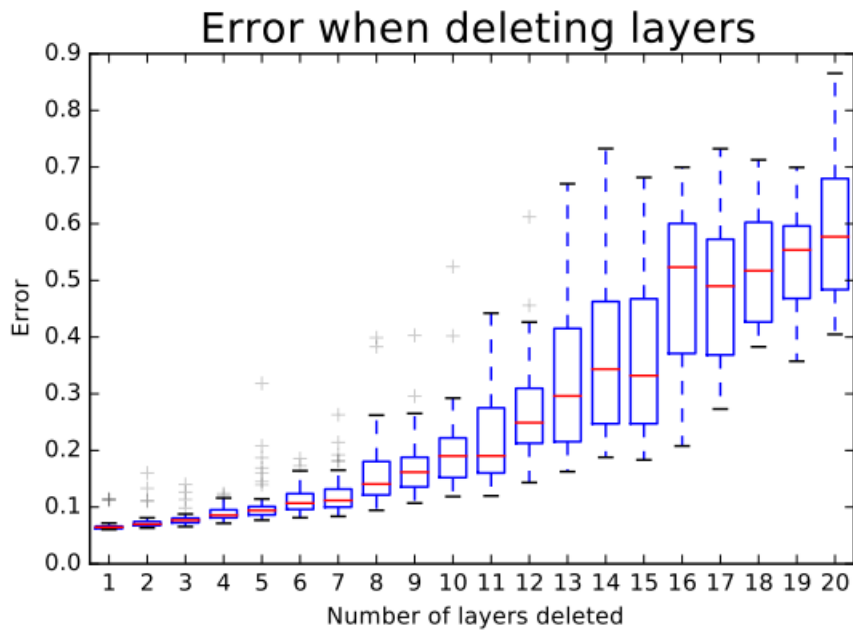
Οι Huang et al. πρότειναν έναν τρόπο εκπαίδευσης ενός πολύ βαθύ δικτύου, ρίχνοντας τυχαία τα στρώματά του κατά τη διάρκεια της προπόνησης και χρησιμοποιώντας το πλήρες δίκτυο στον χρόνο δοκιμής. Οι Veit et al. [14] είχε ένα ακόμη πιο διαισθητικό εύρημα: μπορούμε πραγματικά να ρίξουμε μερικά από τα επίπεδα ενός εκπαιδευμένου ResNet και να έχουμε ακόμα συγκρίσιμη απόδοση. Αυτό καθιστά την αρχιτεκτονική ResNet ακόμη πιο ενδιαφέρουσα καθώς υποβάθμισε δραματικά την απόδοσή του.

Οι Veit et al. [14] παρέχει αρχικά μια αναλυτική προβολή του ResNet για να κάνει τα πράγματα πιο ξεκάθαρα. Αφού ξετυλίξουμε την αρχιτεκτονική δικτύου, είναι πολύ σαφές ότι μια αρχιτεκτονική ResNet με υπολειμματικά μπλοκ έχει $2 \times i$ διαφορετικές διαδρομές (επειδή κάθε υπολειμματικό μπλοκ παρέχει δύο ανεξάρτητες διαδρομές).



Λαμβάνοντας υπόψη το παραπάνω εύρημα, είναι απολύτως σαφές γιατί η αφαίρεση μερικών επιπέδων σε μια αρχιτεκτονική ResNet δεν θέτει σε κίνδυνο την απόδοσή της πάρα πολύ - η αρχιτεκτονική έχει πολλές ανεξάρτητες αποτελεσματικές διαδρομές και η πλειονότητα αυτών παραμένει ανέπαφη μετά την αφαίρεση δύο επιπέδων. Αντίθετα, το δίκτυο VGG έχει μόνο μία αποτελεσματική διαδρομή, οπότε η αφαίρεση ενός στρώματος θέτει σε κίνδυνο αυτή τη μοναδική διαδρομή.

Οι συγγραφείς διεξήγαγαν επίσης πειράματα για να δείξουν ότι η συλλογή των διαδρομών στο ResNet έχει συμπεριφορά σαν σύνολο. Αυτό το κάνουν διαγράφοντας διαφορετικό αριθμό επιπέδων κατά τη διάρκεια της δοκιμής και βλέπουν εάν η απόδοση του δικτύου συσχετίζεται ομαλά με τον αριθμό των διαγραμμένων επιπέδων. Τα αποτελέσματα δείχνουν ότι το δίκτυο συμπεριφέρεται πράγματι σαν σύνολο, όπως φαίνεται στο παρακάτω σχήμα:

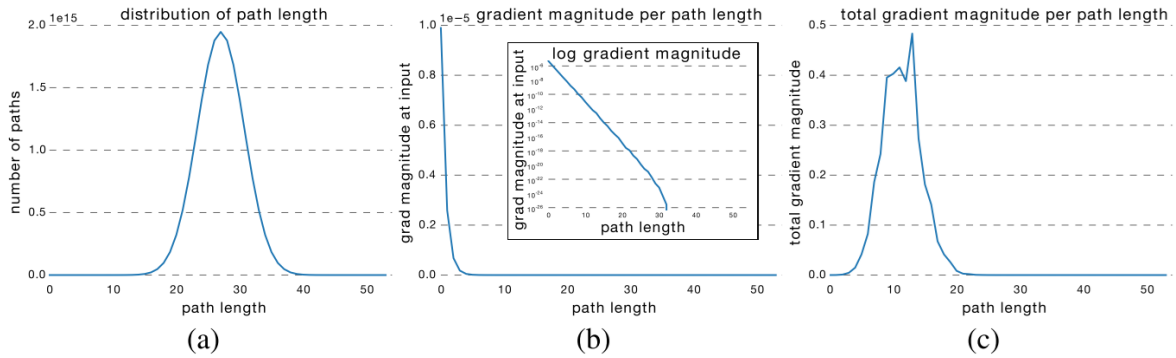


Σχήμα 15: Error Increases smoothly as the number of deleted layers increases

error increases smoothly as the the number of deleted layers increases

Τέλος, οι συγγραφείς εξέτασαν τα χαρακτηριστικά των διαδρομών στο ResNet:

Είναι προφανές ότι η κατανομή όλων των πιθανών μηκών διαδρομής ακολουθεί μια διωνυμική κατανομή, όπως φαίνεται στο (α) του σχήματος εμφύσησης. Η πλειονότητα των διαδρομών περνά από 19 έως 35 υπολειμματικά μπλοκ.



Σχήμα 16: Χαρακτηριστικά των διαδρομών - στο ResNet

Για να διερευνήσει τη σχέση μεταξύ του μήκους της διαδρομής και του μεγέθους των βαθμίδων που ρέουν μέσω αυτής. Για να πάρουν το μέγεθος των διαβαθμίσεων στη διαδρομή του μήκους k , οι συγγραφείς τροφοδότησαν πρώτα μια παρτίδα δεδομένων στο δίκτυο και δειγματοληψία τυχαία k υπολειμμάτων μπλοκ. Όταν διαδίδονταν πίσω οι κλίσεις, διαδόθηκαν μέσω του στρώματος βάρους μόνο για τα υπολειμματικά μπλοκ του δείγματος. (β) δείχνει ότι το μέγεθος των διαβαθμίσεων μειώνεται γρήγορα καθώς η διαδρομή γίνεται μεγαλύτερη.

Μπορούμε τώρα να πολλαπλασιάσουμε τη συχνότητα κάθε μήκους διαδρομής με το αναμενόμενο μέγεθος διαβάθμισης για να έχουμε μια αίσθηση του πόσο διαδρομές κάθε μήκους συμβάλλουν στην προπόνηση, όπως στο (γ). Παραδόξως, οι περισσότερες συνεισφορές προέρχονται από διαδρομές μήκους 9 έως 18, αλλά αποτελούν μόνο ένα μικρό μέρος των συνολικών διαδρομών, όπως στο (α). Αυτό είναι ένα πολύ ενδιαφέρον εύρημα, καθώς υποδηλώνει ότι το ResNet δεν έλυσε το πρόβλημα των εξαφανισμένων βαθμίδων για πολύ μεγάλες διαδρομές, και ότι το ResNet επιτρέπει πραγματικά την εκπαίδευση πολύ βαθύ δίκτυο μειώνοντας τις αποτελεσματικές διαδρομές του.

Ήταν επιτυχές το ResNet; - Ναι.

- Κέρδισε την 1η θέση στον διαγωνισμό ταξινόμησης ILSVRC 2015 με κορυφαίο ποσοστό σφάλματος 3,57% (Ένα μοντέλο συνόλου) .
- Κέρδισε την 1η θέση στο διαγωνισμό ILSVRC και COCO 2015 στον εντοπισμό ImageNet, την ανίχνευση Coco και την τμηματοποίηση της Coco.
- Αντικατάσταση επιπέδων VGG-16 στο Faster R-CNN με το ResNet-101. Παρατήρησαν σχετικές βελτιώσεις 28%
- Αποτελεσματικά εκπαιδευμένα δίκτυα με 100 και 1000 επίπεδα .

ΕΠΙΛΟΓΟΣ

Στο πρώτο κεφάλαιο της πτυχιακής εργασίας , γίνεται παρουσίαση του Residual Network (ResNet. Αρχικά γίνεται μια εισαγωγή στο πως καταλήξαμε στην ανάγκη για την δημιουργία των Residual Networks (ResNet) . Αρχικά γίνεται μια παρουσίαση των Residual Networks. Έπειτα γίνεται εξέταση του Residual Block , και γενικότερα υπάρχει παρουσίαση των δομικών στοιχείων του ResNet. Έπειτα στο τέλος του κεφαλαίου , γίνεται αναφορά στις γενικότερες παραλλαγές του δικτύου ResNet, όπως το ResNeXt, το Deep Network with Stochastic Depth, το ResNet ως σύνολο μικρότερων δικτύων και το Densely Connected CNN. Τέλος βγαίνουν κάποια συμπεράσματα στο προς αν είναι επιτυχές το ResNet

ΚΕΦΑΛΑΙΟ 3 - Densely Connected Convolutional Networks (DenseNet)

ΕΙΣΑΓΩΓΗ

Η αρχιτεκτονική ResNet έχει ένα θεμελιώδες δομικό στοιχείο (Identity) όπου συγχωνεύετε (πρόσθετο) ένα προηγούμενο επίπεδο σε ένα μελλοντικό επίπεδο. Η λογική εδώ είναι με την προσθήκη συγχωνεύσεων πρόσθετων που αναγκάζουμε το δίκτυο να μάθει υπολείμματα (σφάλματα, δηλαδή διαφορά μεταξύ κάποιου προηγούμενου επιπέδου και τρέχοντος). Αντίθετα, το DenseNet προτείνει συνδυασμένες εξόδους από τα προηγούμενα επίπεδα αντί να χρησιμοποιεί το άθροισμα.

Ενώ τα παραδοσιακά συνελκτικά δίκτυα με στρώματα L έχουν L συνδέσεις - μία μεταξύ κάθε επιπέδου και της επόμενης στρώσης - το δίκτυό μας έχει $L(L + 1) / 2$ άμεσες συνδέσεις. Για κάθε επίπεδο, οι χάρτες χαρακτηριστικών όλων των προηγούμενων επιπέδων χρησιμοποιούνται ως είσοδοι και οι δικόι του χάρτες χαρακτηριστικών χρησιμοποιούνται ως είσοδοι σε όλα τα επόμενα επίπεδα.

Τα DenseNets έχουν αρκετά συναρπαστικά πλεονεκτήματα: ανακουφίζουν το πρόβλημα διαβάθμισης που εξαφανίζεται, ενισχύουν τη διάδοση χαρακτηριστικών, ενθαρρύνουν την επαναχρησιμοποίηση χαρακτηριστικών και μειώνουν ουσιαστικά τον αριθμό των παραμέτρων.

Τα DenseNets επιτυγχάνουν σημαντικές βελτιώσεις σε σχέση με την τελευταία λέξη της τεχνολογίας στα περισσότερα από αυτά, ενώ απαιτούν λιγότερη μνήμη και υπολογισμό για την επίτευξη υψηλών επιδόσεων

3.1 Densely Connected Convolutional Networks (DenseNet)

Τα Densely Connected Convolutional Networks ή αλλιώς τα DenseNets, είναι το επόμενο βήμα για να συνεχιστεί η αύξηση του βάθους των συνελκτικών δικτύων.

Έχουμε δει πώς πήγαμε από το LeNet με 5 επίπεδα, στο VGG με 19 επίπεδα και το ResNets ξεπερνώντας 100 ακόμη και 1000 επίπεδα.

Τα προβλήματα προκύπτουν με τα CNN όταν πηγαίνουν βαθύτερα. Αυτό συμβαίνει επειδή η διαδρομή για πληροφορίες από το επίπεδο εισόδου έως το επίπεδο εξόδου (και για την κλίση στην αντίθετη κατεύθυνση) γίνεται τόσο μεγάλη, ώστε να εξαφανιστούν πριν φτάσουν στην άλλη πλευρά.

Το DenseNets απλοποιεί το μοτίβο συνδεσιμότητας μεταξύ επιπέδων που εισάγονται σε άλλες αρχιτεκτονικές όπως :

- Highway Networks
- Residual Networks
- Fractal Networks

Το πρόβλημα λύνεται διασφαλίζοντας τη μέγιστη ροή πληροφοριών (και την κλίση). Για γίνει αυτό, απλώς συνδέεται κάθε στρώμα απευθείας μεταξύ τους.

Αντί να αντλήσει αντιπροσωπευτική δύναμη από εξαιρετικά βαθιές ή ευρείες αρχιτεκτονικές, το DenseNet εκμεταλλεύεται τις δυνατότητες του δικτύου μέσω επαναχρησιμοποίησης χαρακτηριστικών.

Τι πρόβλημα επιλύουν τα DenseNets;

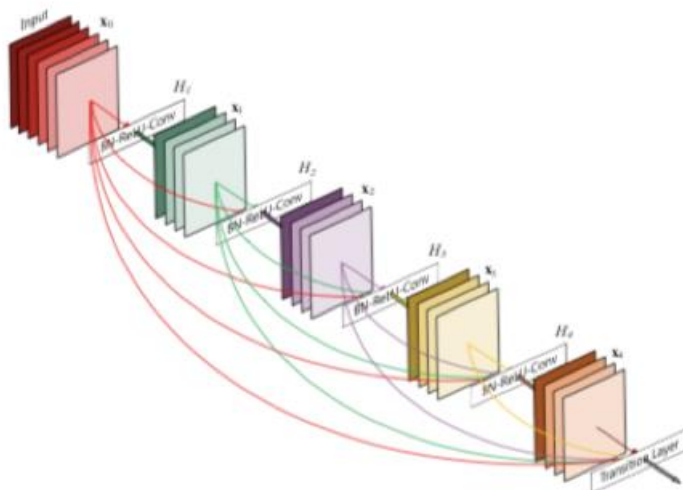
Αντίθετα, συνδέοντας αυτόν τον τρόπο, τα DenseNets απαιτούν λιγότερες παραμέτρους από ένα αντίστοιχο παραδοσιακό CNN, καθώς δεν χρειάζεται να μάθουμε περιττούς χάρτες χαρακτηριστικών.

Επιπλέον, ορισμένες παραλλαγές του ResNets έχουν αποδείξει ότι πολλά επίπεδα συνεισφέρουν μετά βίας οπότε μπορούν να απορριφθούν. Στην πραγματικότητα, ο αριθμός των παραμέτρων του ResNets είναι μεγάλος επειδή κάθε επίπεδο έχει τα βάρη του να μάθει. Αντίθετα, τα επίπεδα DenseNets είναι πολύ στενά (π.χ. 12 φίλτρα) και προσθέτουν μόνο ένα μικρό σύνολο νέων χαρτών χαρακτηριστικών.

Ένα άλλο πρόβλημα με πολύ βαθιά δίκτυα ήταν τα προβλήματα εκπαίδευσης, λόγω της αναφερόμενης ροής πληροφοριών και διαβαθμίσεων. Το DenseNets επιλύει αυτό το ζήτημα, καθώς κάθε επίπεδο έχει άμεση πρόσβαση στις διαβαθμίσεις από τη λειτουργία απώλειας και την αρχική εικόνα εισόδου.

3.2 Δομή του DenseNet

Τα παραδοσιακά νευρωνικά δίκτυα τροφοδοσίας συνδέουν την έξοδο του επιπέδου με το επόμενο επίπεδο μετά την εφαρμογή σύνθετων λειτουργιών.



Σχήμα 17: DenseNet with 5 layers with expansion of 4

Συνήθως αυτό το σύνθετο περιλαμβάνει μια λειτουργία συνέλιξης ή στρώματα συγκέντρωσης, μια ομαλοποίηση παρτίδας και μια λειτουργία ενεργοποίησης.

Η εξίσωση για αυτό θα ήταν:

$$x_l = H_l(x_{l-1})$$

Τα ResNets επέκτειναν αυτήν τη συμπεριφορά, συμπεριλαμβανομένης της παράλειψης σύνδεσης, αναδιατυπώνοντας αυτήν την εξίσωση σε:

$$x_l = H_l(x_{l-1}) + x_{l-1}$$

Το DenseNets κάνει την πρώτη διαφορά με το ResNets εδώ. Τα DenseNets δεν αθροίζουν τους χάρτες χαρακτηριστικών εξόδου του επιπέδου με τους εισερχόμενους χάρτες χαρακτηριστικών, αλλά τους συνδυάζουν.

Κατά συνέπεια, η εξίσωση αναμορφώνεται ξανά σε:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}])$$

Το ίδιο πρόβλημα που αντιμετωπίστηκε στην εργασία με το ResNet, αυτή η ομαδοποίηση χαρτών χαρακτηριστικών δεν μπορεί να γίνει όταν τα μεγέθη τους είναι διαφορετικά. Ανεξάρτητα από το αν η ομαδοποίηση είναι προσθήκη ή συνένωση. Επομένως, και με τον ίδιο τρόπο που χρησιμοποιήθηκε για το ResNet, το DenseNet χωρίζεται σε DenseBlocks, όπου οι διαστάσεις των χαρτών χαρακτηριστικών παραμένουν σταθερές εντός ενός μπλοκ, αλλά ο αριθμός των φίλτρων αλλάζει μεταξύ τους. Αυτά τα επίπεδα μεταξύ τους ονομάζονται στρώματα μετάβασης και φροντίζουν για την κάτω δειγματοληψία εφαρμόζοντας μια ομαλοποίηση παρτίδας, μια συνέλιξη 1x1 και μια στρώση συγκέντρωσης (pooling layer) 2x2.

Τώρα είμαστε έτοιμοι να μιλήσουμε για το ρυθμό ανάπτυξης. Δεδομένου ότι συνδυάζουμε χάρτες χαρακτηριστικών, αυτή η διάσταση καναλιού αυξάνεται σε κάθε επίπεδο. Εάν κάνουμε το H_l να παράγει χάρτες χαρακτηριστικών k κάθε φορά, τότε μπορούμε να γενικεύσουμε για το επίπεδο l -th:

$$k_l = k_0 + k * (l - 1)$$

Με το k ορίζεται ο ρυθμός ανάπτυξης. Ο ρυθμός ανάπτυξης ρυθμίζει πόσες πληροφορίες προστίθενται στο δίκτυο σε κάθε επίπεδο.

Οι χάρτες χαρακτηριστικών θα μπορούσαν να φανούν ως πληροφορίες του δικτύου. Κάθε επίπεδο έχει πρόσβαση στους προηγούμενους χάρτες χαρακτηριστικών του, και ως εκ τούτου, στη συλλογική γνώση. Στη συνέχεια, κάθε στρώμα προσθέτει μια νέα πληροφορία σε αυτήν τη συλλογική γνώση, σε συγκεκριμένους χάρτες πληροφοριών.

DenseNets - B

Τα DenseNets-B είναι απλώς DenseNets που εκμεταλλεύονται την ανάλυση (convolution) 1x1 για να μειώσουν το μέγεθος των χαρτών χαρακτηριστικών πριν από τη ανάλυση (convolution) 3x3 και να βελτιώσουν την απόδοση των υπολογιστών. Το B έρχεται μετά το όνομα Bottleneck layer με το οποίο είστε ήδη εξοικειωμένοι με την εργασία στο ResNets.

DenseNets-BC

Το DenseNets-C είναι ένα ακόμη μικρό βήμα στο DenseNets-B, για τις περιπτώσεις όπου θα θέλαμε να μειώσουμε τον αριθμό των χαρτών χαρακτηριστικών εξόδου. Ο συντελεστής συμπίεσης (θήτα) καθορίζει αυτή τη μείωση. Αντί να έχουμε χάρτες χαρακτηριστικών m σε ένα συγκεκριμένο επίπεδο, θα έχουμε theta * m. Φυσικά, βρίσκεται στο εύρος [0–1]. Έτσι, το DenseNets θα παραμείνει το ίδιο όταν το theta = 1 και θα είναι DenseNets-B διαφορετικά.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

Πίνακας 1: Sizes of outputs and convolutional kernels for different DenseNets [1] architectures on ImageNet.

Sizes of outputs and convolutional kernels for different DenseNets [1] architectures on ImageNet.

Ο παρακάτω πίνακας δείχνει τα αποτελέσματα του DenseNets σε σύνολα δεδομένων CIFAR. Το σύμβολο "+" στο τέλος υποδηλώνει την τυπική αύξηση δεδομένων (τυχαία περικοπή μετά το μηδέν, και οριζόντια αναστροφή). Για ένα μοντέλο DenseNet, το L δηλώνει το βάθος του και το k δηλώνει το ρυθμό ανάπτυξής του. Σε CIFAR-10 και

CIFAR-100 χωρίς αύξηση δεδομένων, ένα στρώμα Dropout με ρυθμό πτώσης 0,2 εισάγεται μετά από κάθε συνελικτικό στρώμα εκτός από το πρώτο.

Model	Parameters	CIFAR-10	CIFAR-10+	CIFAR-100	CIFAR-100+
DenseNet (L=40, k=12)	1.0M	7.00	5.24	27.55	24.42
DenseNet (L=100, k=12)	7.0M	5.77	4.10	23.79	20.20
DenseNet (L=100, k=24)	27.2M	5.83	3.74	23.42	19.25
DenseNet-BC (L=100, k=12)	0.8M	5.92	4.51	24.15	22.27
DenseNet-BC (L=250, k=24)	15.3M	5.19	3.62	19.64	17.60
DenseNet-BC (L=190, k=40)	25.6M	-	3.46	-	17.18

Πίνακας 2: Results of DenseNets on CIFAR datasets

3.3 Dense and Transition Blocks

Dense and Transition Blocks

Ωστόσο, λόγω του εξαιρετικά πυκνού αριθμού συνδέσεων στο DenseNet, η οπτικοποίηση γίνεται λίγο πιο περίπλοκη από ότι ήταν για VGG και ResNet. Το Σχήμα 17 δείχνει ένα πολύ απλό σχήμα στην αρχιτεκτονική του DenseNet-121, το οποίο θα είναι το DenseNet στο οποίο θα επικεντρωθούμε. Αυτό συμβαίνει επειδή είναι το DenseNet μεταξύ αυτών που έχουν σχεδιαστεί μέσω του συνόλου δεδομένων ImageNet.

Μπορούμε να συγκρίνουμε το Σχήμα 17 με το Σχήμα 18 στο DenseNet-121. Τα μέτρα κάτω από κάθε όγκο αντιπροσωπεύουν τα μεγέθη του πλάτους και του βάθους, ενώ οι αριθμοί στην κορυφή αντιπροσωπεύουν τη διάσταση χαρτών χαρακτηριστικών.

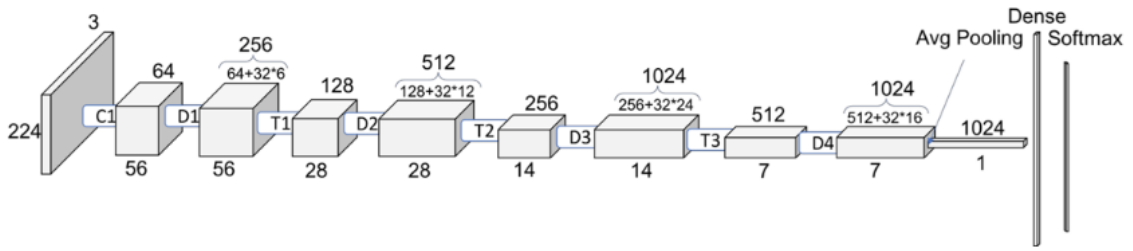
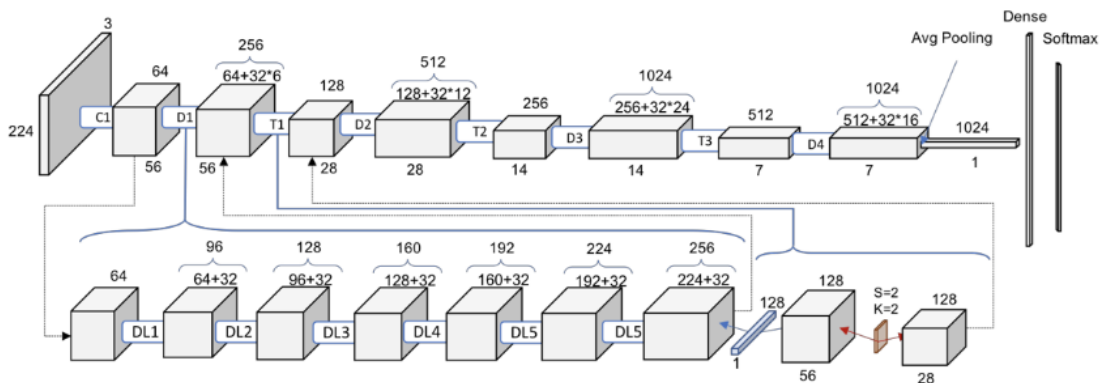


Figure 18. Another look at Dense-121. Dx: Dense Block x. Tx: Transition Block x.

Όπως μπορεί να παρατηρηθεί στο σχήμα δεν υπάρχουν οι συνδέσεις παράκαμψης στο DenseNet. Πρόκειται ουσιαστικά για το απλούστερο σχήμα πριν πάει βαθύτερα σε ολόκληρη την δομή του DenseNet. Ο πρώτος αριθμός της προσθήκης για τον υπολογισμό των χαρτών χαρακτηριστικών κάθε νέου τόμου ταιριάζει με τη διάσταση χαρτών χαρακτηριστικών του προηγούμενου τόμου. Αυτό εξηγεί την σύνδεση παράκαμψης διότι σημαίνει συννέωση.

Ο όγκος μετά από κάθε Dense Block αυξάνεται κατά το ρυθμό ανάπτυξης επί τον αριθμό των Dense Layers σε αυτό το Dense Block.

Στο Σχήμα 4 πηγαίνουμε βαθύτερα τώρα και καταλαβαίνουμε τι συμβαίνει πραγματικά σε κάθε μπλοκ.



Σχήμα 19: One level deeper look at DenseNet - 121

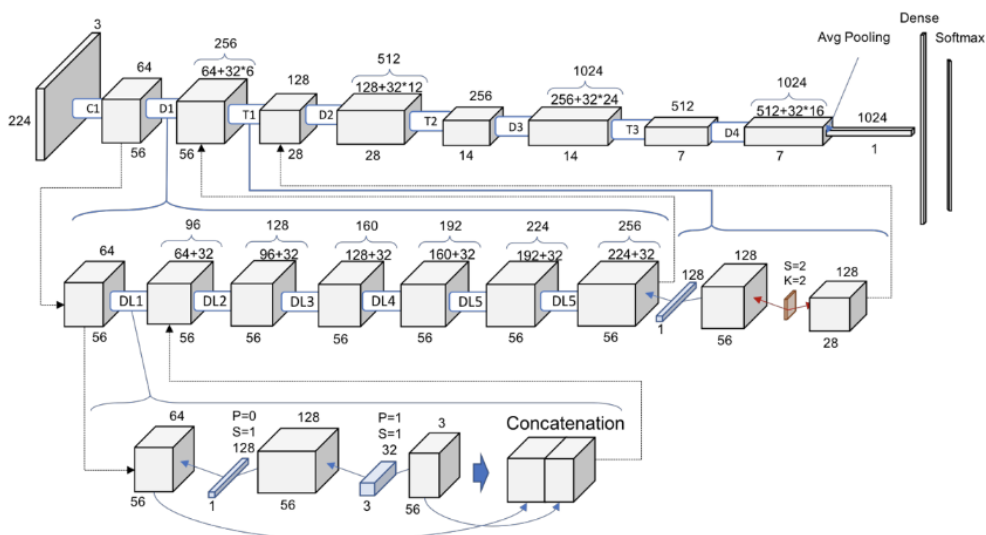
Σχήμα 4. One level deeper look at DenseNet-121. Dense Block and Transition Block. DLx: Dense Layer x

Dense Layers

Κάθε επίπεδο προσθέτει στον προηγούμενο τόμο αυτούς τους 32 νέους χάρτες χαρακτηριστικών. Γι' αυτό πηγαίνουμε από 64 σε 256 μετά από 6 επίπεδα. Επιπλέον, το Transition Block αποδίδει ως ανάλυση 1x1 με 128 φίλτρα. ακολουθούμενη από συγκέντρωση 2x2 με βήμα 2, με αποτέλεσμα το διαχωρισμό του μεγέθους του όγκου και του αριθμού χαρτών χαρακτηριστικών στο μισό.

Από αυτήν την παρατήρηση προτύπων βγαίνουν τα εξής συμπεράσματα :

- Η ένταση σε ένα Dense Block παραμένει σταθερή.
- Ο όγκος και οι χάρτες χαρακτηριστικών μειώνονται στο μισό μετά από κάθε μεταβατικό μπλοκ (Transition Block).



Σχήμα 20: Full Schematic representation of DenseNet - 121

Ωστόσο, μπορούμε ακόμα να πάμε 1 επίπεδο βαθύτερα. Για να κατανοηθεί πλήρως το τι συμβαίνει σε κάθε Dense Layer και σε κάθε Dense Block, καθώς δεν είναι ακόμη ασήμαντο. Η πλήρης εικόνα φαίνεται παρακάτω :

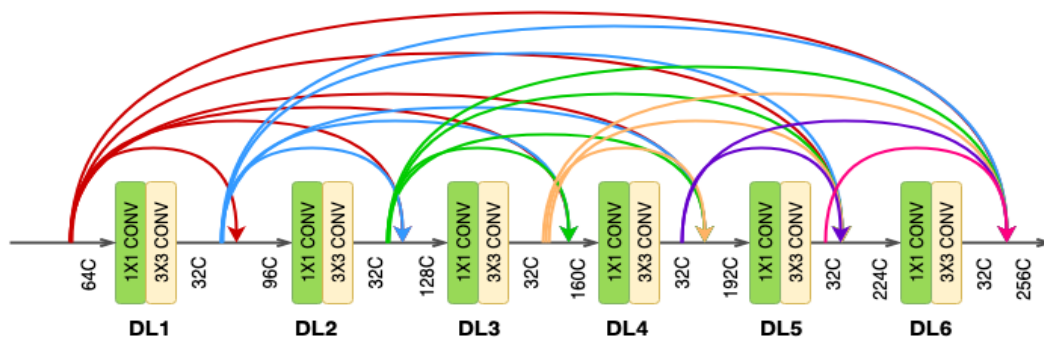
Στο νέο βαθύτερο επίπεδο που αντιπροσωπεύει το πρώτο Dense Layer στο πρώτο Dense Block, φαίνεται πόσο πραγματικά επιτυγχάνεται αυτή η συμπεριφορά της προσθήκης 32 φορές τον αριθμό των επιπέδων. Προτείνεται μια συνέλιξη 1x1 με 128 φίλτρα για να μειώσει * το μέγεθος χαρτών χαρακτηριστικών και να εκτελέσει μια πιο ακριβή ανάλυση 3x3 με αυτόν τον επιλεγμένο αριθμό 32 χαρτών ρυθμού ανάπτυξης.

Στη συνέχεια, ο όγκος εισόδου και το αποτέλεσμα των δύο λειτουργιών (οι οποίες είναι οι ίδιες για κάθε Dense Layer σε κάθε Dense Block) συνενώνονται, με την προσθήκη νέων πληροφοριών στην κοινή γνώση του δικτύου.

Growth – Rate (k) – Ποσοστό Ανάπτυξης

Είναι βασικά ο αριθμός των καναλιών που εξέρχονται από ένα πυκνό στρώμα (1x1 μετατρ. → 3x3 μετατρ.). Οι συγγραφείς έχουν χρησιμοποιήσει μια τιμή $k = 32$ για τα πειράματα. Αυτό σημαίνει ότι ο αριθμός των χαρακτηριστικών που λαμβάνονται από ένα πυκνό στρώμα (l) από το προηγούμενο πυκνό στρώμα (l-1) είναι 32. Αυτό αναφέρεται ως ο ρυθμός ανάπτυξης επειδή μετά από κάθε στρώση, 32 χαρακτηριστικά καναλιών συνδυάζονται και τροφοδοτούνται ως είσοδος στο επόμενο επίπεδο.

Σχήμα 21: Transition layer/block 1

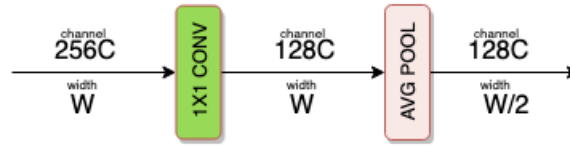


Transition layer

Στο τέλος κάθε πυκνού (dense) μπλοκ, ο αριθμός των χαρτών χαρακτηριστικών συσσωρεύεται σε μια τιμή - χαρακτηριστικά εισόδου + (αριθμός πυκνών στρωμάτων x ρυθμός ανάπτυξης). Έτσι, για 64 χαρακτηριστικά καναλιών που εισέρχονται σε ένα πυκνό μπλοκ 6 πυκνών επιπέδων ρυθμού ανάπτυξης 32, ο αριθμός των καναλιών που συσσωρεύονται στο τέλος του μπλοκ θα είναι $64 + (6 \times 32) = 256$. Για να μειώσετε αυτόν

τον αριθμό καναλιών, ένα στρώμα μετάβασης (ή μπλοκ) προστίθεται μεταξύ δύο πυκνών μπλοκ

Το στρώμα μετάβασης – Transition Layer αποτελείται από -



Σχήμα 22: Transition layer/block 2

- **1 X 1**
- **CONV** operation
- **2 X 2 AVG POOL** operation

Η λειτουργία 1 X 1 CONV μειώνει τον αριθμό καναλιών στο μισό.

Το στρώμα 2 X 2 AVG POOL είναι υπεύθυνο για τη χαμηλή δειγματοληψία των χαρακτηριστικών όσον αφορά το πλάτος και το ύψος.

ΕΠΙΛΟΓΟΣ

Σε αυτό το κεφάλαιο γίνεται παρουσίαση των Densely Connected Convolutional Networks ή αλλιώς τα DenseNets. Το DenseNet είναι ένα δίκτυο που απεικονίζει τη σημασία της ύπαρξης πυκνών συνδέσεων σε ένα δίκτυο χρησιμοποιώντας πυκνά μπλοκ. Αυτό βοηθά στην επαναχρησιμοποίηση χαρακτηριστικών, στην καλύτερη ροή διαβάθμισης, στον μειωμένο αριθμό παραμέτρων και στην καλύτερη μετάδοση των χαρακτηριστικών στο δίκτυο. Μια τέτοια εφαρμογή μπορεί να βοηθήσει στην εκπαίδευση βαθύτερων δικτύων χρησιμοποιώντας λιγότερους υπολογιστικούς πόρους και με καλύτερα αποτελέσματα. Αρχικά γίνεται παρουσίαση της δομής και της αρχιτεκτονικής των DenseNets. Εν συνεχεία υπάρχει μια ανάλυση του DenseNet Block και Transition Block, καθώς και τα Dense Layers που είναι απο τα δομικά στοιχεία των DenseNets.

ΚΕΦΑΛΑΙΟ 4 – ΣΥΓΚΡΙΣΗ RESNET - DENSENET

ΕΙΣΑΓΩΓΗ

Η αρχική σκέψη, ότι η ακρίβεια αυξάνεται καθώς το δίκτυο γίνεται βαθύτερο, δεν είναι απαραίτητα αλήθεια - Καθώς το δίκτυο γίνεται βαθύτερο, προκύπτουν ορισμένα ζητήματα απόδοσης και ακρίβειας. Ένα από τα μεγαλύτερα προβλήματα είναι το πρόβλημα της εξαφάνισης της κλίσης: το DCNN εκπαιδεύει το μοντέλο υπολογίζοντας τα παράγωγα των παραμέτρων σε σχέση με τη λειτουργία απώλειας εκπαίδευσης. Ωστόσο, οι κλίσεις της λειτουργίας απώλειας πλησιάζουν το μηδέν καθώς προστίθενται περισσότερα επίπεδα στα νευρωνικά δίκτυα. Επομένως, η επίδραση της λειτουργίας απώλειας στις λειτουργίες ενεργοποίησης μειώνεται, γεγονός που καθιστά το δίκτυο δύσκολο να εκπαιδευτεί. Ορισμένες τεχνικές ομαλοποίησης θα μπορούσαν να αντιμετωπίσουν αυτό το πρόβλημα, αλλά μόνο σε στρώματα μεσαίας κλίμακας, όχι σε στρώματα μεγάλης κλίμακας.

Οι αρχιτεκτονικές DenseNets και ResNets προσπαθούν να αντιμετωπίσουν αυτό το πρόβλημα και, συνεπώς είναι αρκετά καινοτόμες όσον αφορά την απόδοση των χρήσεων των αρχιτεκτονικών DCNN.

Σε αυτό το κεφάλαιο θα γίνει μια συγκριτική ανάλυση ανάμεσα στα ResNets και στα DenseNets. Θα παρουσιαστούν τα πλεονεκτήματα και τα μειονεκτήματα των δύο αυτών CNNs. Επίσης θα γίνει σύγκριση ανάμεσα στις αρχιτεκτονικές τους και τέλος θα γίνει εξαγωγή κάποιων σημαντικών συμπερασμάτων από αυτή την έρευνα.

4.1 Πλεονεκτήματα ResNet - DenseNet

Τα Υπολειμματικά Δίκτυα (ResNet) προχωρούν το CNN (Convolutional Neural Network) ένα βήμα παραπέρα με «μικρές» αλλά «σημαντικές» αλλαγές στην αρχιτεκτονική. Παρακάτω ακολουθεί μια μικρή ανάλυση στα πλεονεκτήματα του ResNet:

- 1) Επιτάχυνση της ταχύτητας της εκπαίδευσης των βαθιών δικτύων (deep networks) .

2) Αντί να διευρύνει το δίκτυο, η αύξηση του βάθους του δικτύου οδηγεί σε λιγότερες επιπλέον παραμέτρους .

3) Μείωση της επίδρασης του Εξαφανιστικού Προβλήματος Διαβάθμισης (Vanish Gradient Problem) .

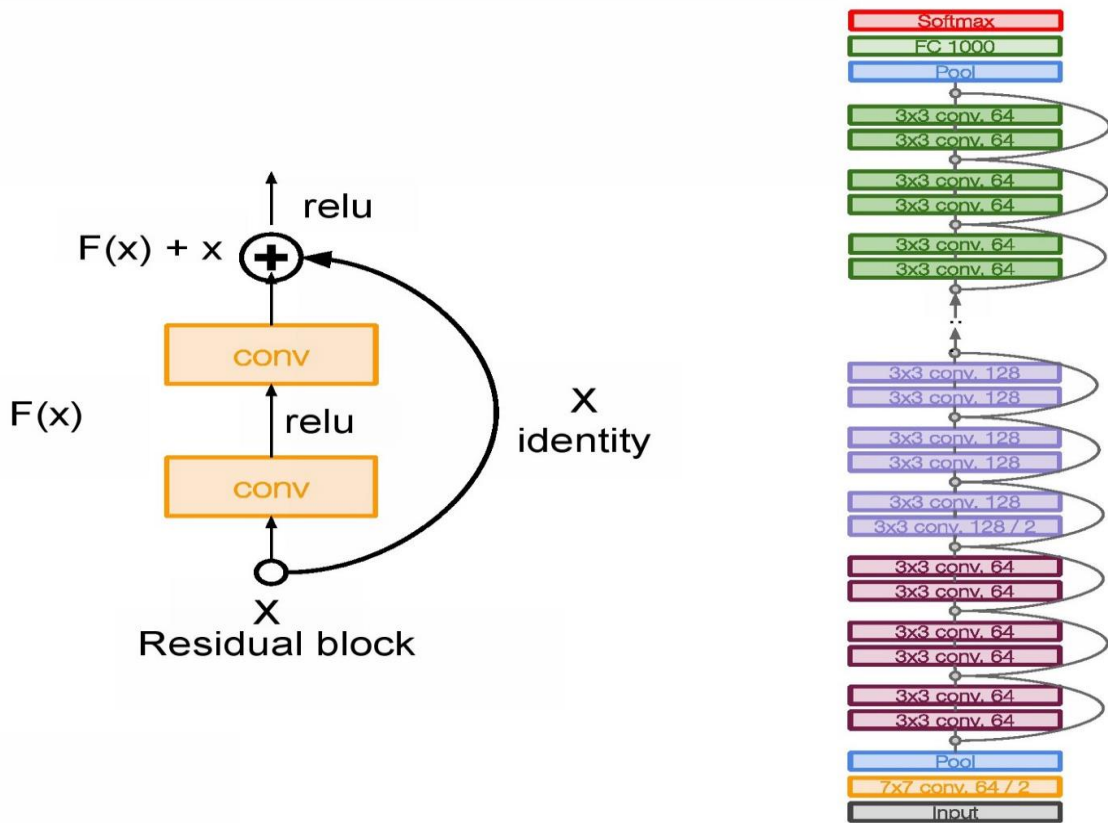
4) Απόκτηση μεγαλύτερης ακρίβειας στην απόδοση του δικτύου, ιδίως στην ταξινόμηση εικόνας.

Με την σειρά τους τώρα τα DenseNets έχουν και αυτά κάποια σημαντικά πλεονεκτήματα τα οποία είναι :

1. Ανακουφίζουν το πρόβλημα της εξαφάνισης (vanish gradient problem)
2. Ενιχύουν την διάδοση χαρακτηριστικών.
3. ενθαρρύνουν την επαναχρησιμοποίηση χαρακτηριστικών.
4. μειώνουν σημαντικά τον αριθμό των παραμέτρων.

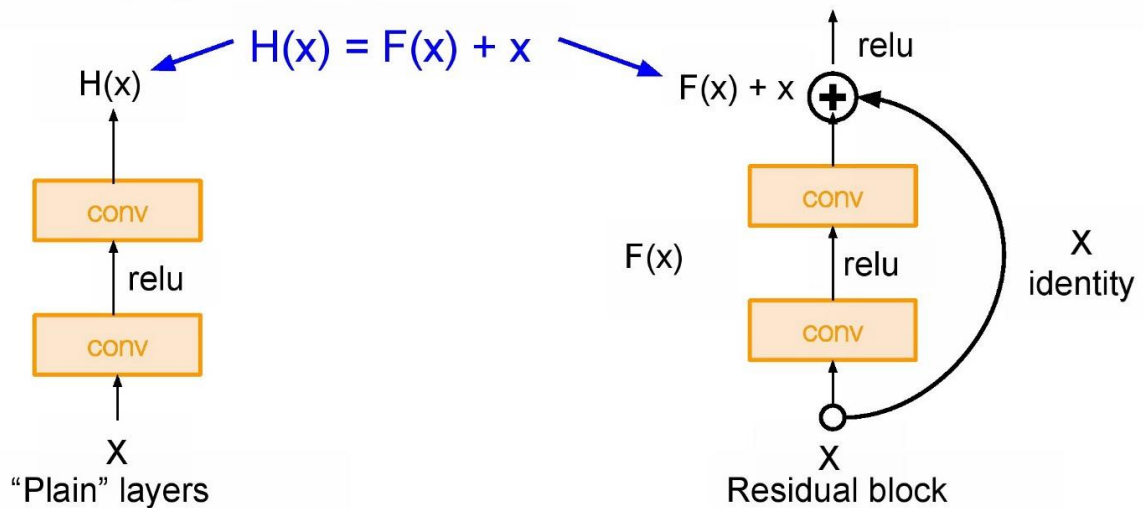
4.2 Σύγκριση αρχιτεκτονικής ResNet και DenseNet

Residual Network - ResNet



Σχήμα 23: ResNet Architecture and Layer Details

Το κύριο βασικό στοιχείο του ResNet είναι το υπόλοιπο μπλοκ. Καθώς το δίκτυο γίνεται βαθύτερο με μεγάλο αριθμό επιπέδων, ο υπολογισμός γίνεται πιο περίπλοκος. Αυτά τα επίπεδα τοποθετούνται το ένα πάνω στο άλλο και κάθε στρώμα προσπαθεί να μάθει κάποια υποκείμενη χαρτογράφηση της επιθυμητής λειτουργίας και αντί να έχουμε αυτά τα μπλοκ, προσπαθούμε να ταιριάξουμε μια υπολειμματική χαρτογράφηση.



Σχήμα 24: Plain Layers and Residual Block

Η είσοδος σε αυτά τα μπλοκ είναι ακριβώς η είσοδος, ενώ στην άλλη πλευρά, θα χρησιμοποιήσουμε τα στρώματά μας για να προσπαθήσουμε να τοποθετήσουμε κάποια υπολείμματα του $H(X) - X$ αντί της επιθυμητής συνάρτησης $H(X)$ άμεσα. Βασικά λοιπόν, στο τέλος αυτού του μπλοκ παίρνει τη σύνδεση παράλειψης (skip connection) εδώ ακριβώς, όπου παίρνει απλώς την είσοδο και την περνά ως ταυτότητα, και έτσι αν δεν είχε στρώματα βάρους μεταξύ του, απλώς επρόκειτο να είναι η ταυτότητα. Θα ήταν το ίδιο πράγμα με την έξοδο, αλλά τώρα χρησιμοποιούμε επιπλέον στρώματα βάρους για να μάθουμε κάποια δέλτα, για κάποια υπολείμματα από το X .

Με λίγα λόγια, καθώς μπαίνουμε βαθύτερα στο δίκτυο, είναι τόσο δύσκολο να μάθουμε την $H(X)$, καθώς έχουμε μεγάλο αριθμό επιπέδων. Εδώ χρησιμοποιήσαμε παράλειψη σύνδεσης και μάθηση $F(x)$ άμεση είσοδο του x ως τελικό αποτέλεσμα. Έτσι το $F(x)$ ονομάζεται υπολειπόμενο. Εδώ χρησιμοποιήσαμε παράλειψη σύνδεσης και μάθαμε την $F(x)$ με άμεση είσοδο του x ως τελικό αποτέλεσμα. Έτσι το $F(x)$ ονομάζεται υπολειπόμενο (residual).

Στο ResNet, συγκεντρώνονται όλα αυτά τα μπλοκ μαζί πολύ βαθιά. Ένα άλλο πράγμα με αυτήν την πολύ βαθιά αρχιτεκτονική είναι ότι επιτρέπει σε βάθος έως και 150 στρώματα και, στη συνέχεια, αυτό που κάνουμε είναι να στοιβάζουμε περιοδικά όλα αυτά τα επίπεδα. Διπλασιάζουμε επίσης τον αριθμό των φίλτρων και υποβαθμίζουμε χωρικά το δεύτερο βήμα.

Στο ResNet, χρησιμοποιεί Batch Normalization μετά από κάθε επίπεδο μετατροπής. Χρησιμοποιεί επίσης την αρχικοποίηση Xavier με SGD + Momentum. Το ποσοστό εκμάθησης είναι 0,1 και διαιρείται με 10 καθώς το σφάλμα επικύρωσης γίνεται σταθερό. Επιπλέον, το μέγεθος παρτίδας είναι 256 και η αποσύνθεση βάρους είναι $1e-5$. Το σημαντικό μέρος είναι ότι δεν υπάρχει εγκατάλειψη (dropout) που χρησιμοποιείται στο ResNet.

Densely Connected Convolutional Networks (DenseNet)

Το DenseNet αποτελείται από μπλοκ Dense όπως φαίνεται παρακάτω. Μέσα σε αυτά τα μπλοκ, τα στρώματα συνδέονται πυκνά μεταξύ τους: Κάθε επίπεδο λαμβάνει την είσοδο από χάρτες χαρακτηριστικών εξόδου προηγούμενων επιπέδων. Αυτή η ακραία επαναχρησιμοποίηση υπολειμμάτων δημιουργεί βαθιά επίβλεψη επειδή κάθε επίπεδο λαμβάνει περισσότερη επίβλεψη από το προηγούμενο στρώμα και έτσι η λειτουργία απώλειας θα αντιδρά αναλόγως και λόγω αυτής της μεθοδολογίας, το καθιστά ένα πιο ισχυρό δίκτυο.

Το DenseNet αποτελείται από 2 μπλοκ:

1. Πυκνό μπλοκ – Dense Block : Ένα μονό μπλοκ αποτελείται από αυτά τα επίπεδα:

- Ομαλοποίηση παρτίδας – Batch Normalization
- Ενεργοποίηση ReLU – ReLU Activation
- 3x3 Convolution

2. Επίπεδο μετάβασης – Transition Layer : Στο ResNet θα εκτελεστεί άθροισμα υπολειμμάτων, αντί να αθροιστεί το υπόλοιπο (residual) DenseNet που συνδυάζει όλους τους χάρτες χαρακτηριστικών. Αυτό το στρώμα είναι κατασκευασμένο από :

- Batch Normalization – Ομαδοποίηση παρτίδας
- 1x1 Convolution
- Average Pooling

Βασικά, η συνεκτικότητα του Densenet δημιουργεί μικρότερο αριθμό χαρτών χαρακτηριστικών. Το DenseNet έχει χαμηλότερη ανάγκη για ευρεία στρώματα, καθώς καθώς τα στρώματα είναι πυκνά συνδεδεμένα, υπάρχει μικρή πλεονασματικότητα στα γνωστά χαρακτηριστικά. Τα επίπεδα πυκνού μπλοκ μοιράζονται ένα κομμάτι συλλογικής γνώσης. Ο αριθμός χαρτών χαρακτηριστικών εξόδου ενός επιπέδου ορίζεται ως ο ρυθμός ανάπτυξης. Τελικά, ο ρυθμός ανάπτυξης ελέγχει πόσες νέες πληροφορίες συμβάλλει κάθε επίπεδο.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 × 56 28 × 28	1 × 1 conv 2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 × 28 14 × 14	1 × 1 conv 2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14 × 14 7 × 7	1 × 1 conv 2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1 × 1	7 × 7 global average pool 1000D fully-connected, softmax			

Πίνακας 3: DenseNet Architecture

Με λίγα λόγια, η αρχιτεκτονική DenseNet χρησιμοποιεί τον υπολειπόμενο μηχανισμό στο μέγιστο, καθιστώντας κάθε στρώμα πυκνά συνδεδεμένο με τα επόμενα στρώματά του. Η συμπαγής εφαρμογή του μοντέλου καθιστά τα γνωστά χαρακτηριστικά μη περιττά, καθώς όλα μοιράζονται μέσω συλλογικών γνώσεων.

4.3 Η αποτελεσματικότητα του Densenet

Από τα Restnets, τα δίκτυα αυτοκινητόδρομων (Highway Networks) και τα βαθιά και ευρεία νευρωνικά δίκτυα, οι άνθρωποι προσπαθούν να προσθέσουν περισσότερες συνδέσεις μεταξύ επιπέδων εκτός από την άμεση σύνδεση σε παρακείμενα στρώματα για να ενισχύσουν τη ροή πληροφοριών κατά μήκος των επιπέδων. Παρόμοια με το Resnet, το Densenet προσθέτει συντομεύσεις μεταξύ των επιπέδων. Διαφορετικό από το Resnet, ένα

επίπεδο σε πυκνό δέχεται όλα τα outs των προηγούμενων επιπέδων και τα συνδυάζει στην διάσταση βάθους. Στο Resnet, ένα επίπεδο δέχεται εξόδους μόνο από το προηγούμενο δεύτερο ή τρίτο επίπεδο και οι εξοδοι προστίθενται μαζί στο ίδιο βάθος, επομένως δεν θα αλλάξει το βάθος προσθέτοντας συντομεύσεις. Με άλλα λόγια, στο Resnet η έξοδος του επιπέδου k είναι $x[k] = f(w * x[k-1] + x[k-2])$, ενώ στο Densenet είναι $x[k] = f(w * H(x[k-1], x[k-2], \dots, x[1]))$ όπου H σημαίνει στοίβαγμα πάνω από τη διάσταση βάθους. Εκτός αυτού, το Resnet διευκολύνει τη λειτουργία εκμάθησης ταυτότητας, ενώ το Densenet προσθέτει άμεσα τη λειτουργία ταυτότητας.

Τα κυριότερα σημεία

Η βασική ιδέα του Densenet είναι η ίδια με το ResNet, αλλά δημιουργεί την πυκνή σύνδεση μεταξύ όλων των προηγούμενων επιπέδων και του πίσω στρώματος, και το όνομά του προέρχεται επίσης από αυτό. Ένα άλλο σημαντικό χαρακτηριστικό του DenseNet είναι η επαναχρησιμοποίηση δυνατοτήτων μέσω της σύνδεσης χαρακτηριστικών στο κανάλι. Αυτές οι δυνατότητες επιτρέπουν στο DenseNet να επιτύχει καλύτερη απόδοση από το ResNet με λιγότερες παραμέτρους και υπολογιστικό κόστος

Το Densenet είναι πιο αποτελεσματικό σε ορισμένα σημεία αναφοράς ταξινόμησης εικόνας. Από τα ακόλουθα γραφήματα, μπορούμε να δούμε ότι το densenet είναι πολύ πιο αποτελεσματικό όσον αφορά τις παραμέτρους και τον υπολογισμό για το ίδιο επίπεδο ακρίβειας, σε σύγκριση με το resnet.

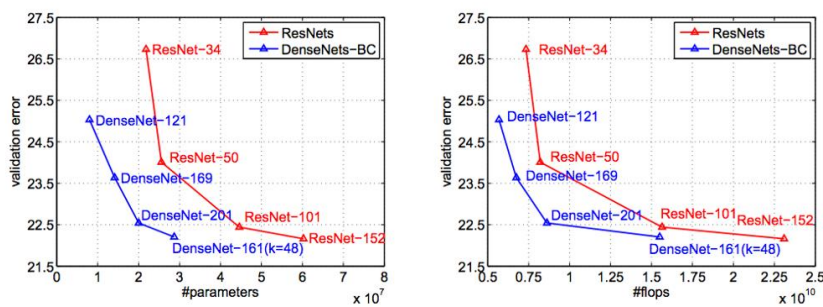


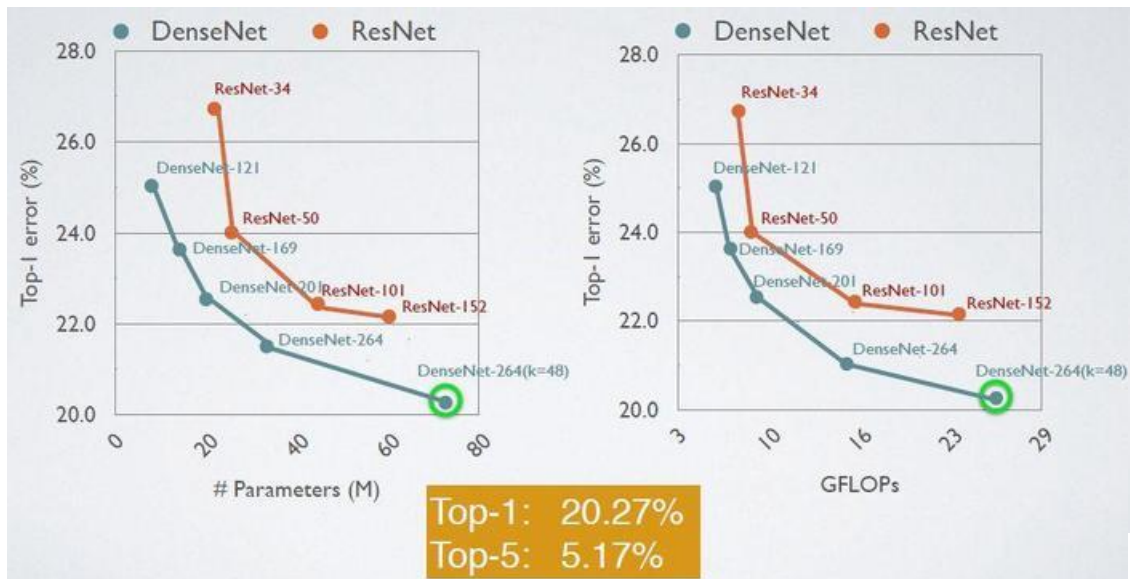
Figure 3. Comparison of the DenseNet and ResNet Top-1 (single model and single-crop) error rates on the ImageNet classification dataset as a function of learned parameters (left) and flops during test-time (right).

Σχήμα 25: Comparison of the DenseNet and ResNet

Σε σύγκριση με το ResNet, το DenseNet προτείνει έναν πιο ριζικό και πυκνό μηχανισμό σύνδεσης: διασύνδεοντας όλα τα επίπεδα, ειδικά κάθε επίπεδο δέχεται όλα τα προηγούμενα στρώματά του ως πρόσθετη είσοδο. Μπορεί να φανεί ότι το ResNet είναι βραχυκυκλωμένο με κάθε στρώμα μπροστά από ένα συγκεκριμένο επίπεδο (γενικά 2 ~ 3 επίπεδα) και η μέθοδος σύνδεσης προστίθεται κατά επίπεδο στοιχείου. Στο DenseNet, κάθε επίπεδο συνδυάζεται με όλα τα προηγούμενα επίπεδα στη διάσταση καναλιού (όπου οι χάρτες χαρακτηριστικών των επιπέδων είναι οι ίδιοι) και χρησιμοποιούνται ως είσοδο στο επόμενο επίπεδο. Για ένα δίκτυο L-layer, το DenseNet περιέχει συνολικά συνδέσεις $L * (L + 1) / 2$, που είναι μια πυκνή σύνδεση σε σύγκριση με το ResNet. Επιπλέον, το DenseNet είναι ένας άμεσος χάρτης δυνατοτήτων από διαφορετικά επίπεδα, ο οποίος μπορεί να επιτύχει την επαναχρησιμοποίηση χαρακτηριστικών και να βελτιώσει την αποδοτικότητα. Αυτή η δυνατότητα είναι η κύρια διαφορά μεταξύ DenseNet και ResNet.

Πειραματικά αποτελέσματα και συζήτηση

Ακολουθεί μια σύγκριση του DenseNet με το ResNet στα σύνολα δεδομένων CIFAR-100 και ImageNet, όπως φαίνεται στα Σχήματα 23 και 24. Όπως φαίνεται από το Σχήμα 23, μόνο η απόδοση 0,8M DenseNet-100 έχει ξεπεράσει το ResNet-1001 και το τελευταίο το μέγεθος της παραμέτρου είναι 10,2M. Όπως φαίνεται από το Σχήμα 24, το DenseNet είναι επίσης ανώτερο από το δίκτυο ResNet όταν το μέγεθος των παραμέτρων είναι το ίδιο. Άλλα πειραματικά αποτελέσματα βρίσκονται στο πρωτότυπο έγγραφο.



Σχήμα 26: ResNet vs DenseNet on the ImageNet dataset

Συνολικά, τα πλεονεκτήματα του DenseNet σε σχέση με το ResNet αντικατοπτρίζονται κυρίως στις ακόλουθες πτυχές :

- Λόγω της πυκνής σύνδεσης, το DenseNet βελτιώνει την οπίσθια διάδοση της κλίσης, κάνοντας το δίκτυο ευκολότερο να εκπαιδευτεί. Δεδομένου ότι κάθε επίπεδο μπορεί να φτάσει απευθείας στο τελικό σήμα σφάλματος, εφαρμόζεται μια σιωπηρή «βαθιά επίβλεψη».
- Οι μικρότερες παράμετροι και οι πιο αποτελεσματικοί υπολογισμοί είναι λίγο αντίθετοι. Επειδή το DenseNet εφαρμόζει συνδέσεις βραχυκυκλώματος μέσω δυνατότητας συμπίκνωσης, επαναχρησιμοποίησης χαρακτηριστικών και μικρότερου ρυθμού ανάπτυξης, ο χάρτης χαρακτηριστικών που είναι μοναδικός σε κάθε επίπεδο είναι σχετικά μικρός .
- Λόγω της επαναχρησιμοποίησης χαρακτηριστικών, ο τελικός ταξινομητής χρησιμοποιεί δυνατότητες χαμηλού επιπέδου.

Ένα σημαντικό σημείο που πρέπει να σημειωθεί είναι ότι το DenseNet μπορεί να καταναλώνει πολλή μνήμη GPU εάν δεν εφαρμόζεται σωστά.

ΕΠΙΛΟΓΟΣ

Στο τέταρτο και τελευταίο κεφάλαιο της παρούσας πτυχιακής γίνεται μια σύγκριση ανάμεσα στα ResNet και DenseNet. Αρχικά παρουσιάζονται τα πλεονεκτήματα και των δύο δικτύων. Στη συνέχεια γίνεται μια σύγκριση ανάμεσα στις αρχιτεκτονικές του ResNet και του DenseNet καθώς και των Residual Block και Dense Block. Παρακάτω αναλύεται η αποτελεσματικότητα του DenseNet , Και παρουσιάζονται οι λόγοι για τους οποίους το DenseNet είναι αποτελεσματικότερο απο το ResNet. Τέλος εξάγονται κάποια συμπεράσματα απο την σύγκριση καθώς και παρουσιάζονται κάποια πλεονεκτήματα και των δύο τεχνητών νευρωνικών δικτύων.

BIBΛΙΟΓΡΑΦΙΑ

- Artificial Neural Networks - Methodological Advances and Biomedical Applications Edited by Kenji Suzuki
- Artificial Neural Networks Advantages and Disadvantages, Maad M. Mijwil
- A deep residual network for recognizing cluster-based generic photo categories , Evi Sijben
- K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385,2015.
- K. He, X. Zhang, S. Ren, and J. Sun. Identity Mappings in Deep Residual Networks. arXiv preprint arXiv:1603.05027v3,2016.
- [1]. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1–9,2015.
- [2]. S. Xie, R. Girshick, P. Dollar, Z. Tu and K. He. Aggregated Residual Transformations for Deep Neural Networks. arXiv preprint arXiv:1611.05431v1,2016.
- [3]. G. Huang, Y. Sun, Z. Liu, D. Sedra and K. Q. Weinberger. Deep Networks with Stochastic Depth. arXiv:1603.09382v3,2016.
- [4]. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. The Journal of Machine Learning Research 15(1) (2014) 1929–1958.
- [5]. G. Huang, Z. Liu and L. van der Maaten, “Densely Connected Convolutional Networks,” 2018.
- [6]. G. Larsson, M. Maire and G. Shakhnarovich, “FractalNet: Ultra-Deep Neural Networks without Residuals,” 2017.
- [7]. K. He, X. Zhang, S. Ren and Jian Sun, “Deep Residual Learning for Image Recognition,” 2015.
- [8]. DenResNet: Ensembling Dense Networks and Residual Networks , Victor Cheung, Stanford Universit
- [9]. K. Obermayer, G. G. Blasdel, and K. Schulten, “A neural-network model for the formation and for the spatial structure of retinotopic maps, orientation, and ocular dominance columns,” in Artificial Neural Networks, T. Kohonen, K. Makisara, O. Simula, and J. Kangas, Eds. “ Amsterdam, Netherlands: Elsevier, 1991, pp. 505–511.

- [10]. K. Obermayer, H. Ritter, and K. Schulten, “Large-scale simulation of a self-organizing neural network: Formation of a somatotopic map,” in *Parallel Processing in Neural Systems and Computers*, R. Eckmiller, G. Hartmann, and G. Hauske, Eds. Amsterdam, Netherlands: North–Holland, 1990, pp. 71–74.
- [11]. M. Mandischer, “Evolving recurrent neural networks with nonbinary encoding,” in *Proc. 1995 IEEE Int. Conf. Evolutionary Computation, Part 2 (of 2)*, pp. 584–589.
- [12]. F. Heimes, G. Zaleski, W. L. Jr., and M. Oshima, “Traditional and evolved dynamic neural networks for aircraft simulation,” in *Proc. 1997 IEEE Int. Conf. Systems, Man, and Cybernetics, Part 3 (of 5)*, pp. 1995–2000.
- [13]. J. Sietsma and R. J. F. Dow, “Creating artificial neural networks that generalize,” *Neural Networks*, vol. 4, no. 1, pp. 67–79, 1991.
- [14]. N. Dodd, D. Macfarlane, and C. Marland, “Optimization of artificial neural network structure using genetic techniques implemented on multiple transputers,” in *Proc. Transputing’91*, P. Welch, D. Stiles, T. L. Kunii, and A. Bakkers, Eds. Amsterdam, The Netherlands: IOS, 1991, pp. 687–700.
- [15]. B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. Enhanced deep residual networks for single image super-resolution. In *CVPRW*, 2017.
- [16]. H. Zhang and V. M. Patel. Densely connected pyramid dehazing network. In *CVPR*, 2018.
- [17]. S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *ICML*, 2015.
- [18]. M. Jaderberg, W. M. Czarnecki, S. Osindero, O. Vinyals, A. Graves, and K. Kavukcuoglu. Decoupled neural interfaces using synthetic gradients. *arXiv preprint arXiv:1608.05343*, 2016.
- [19]. N. Kalchbrenner, L. Espeholt, K. Simonyan, A. v. d. Oord, A. Graves, and K. Kavukcuoglu. Neural machine translation in linear time. *arXiv preprint arXiv:1610.10099*, 2016.
- [20]. Nikolaus Mayer, Eddy Ilg, Philip Husser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox, “A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation,” in *CVPR*, 2016.
- [21]. Simon Jgou, Michal Drozdal, David Vazquez, Adriana Romero, and Yoshua Bengio, “The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation,” *arXiv preprint arXiv:1611.09326*, 2016.