



RSLab

Remote Sensing Laboratory
National Technical University of Athens



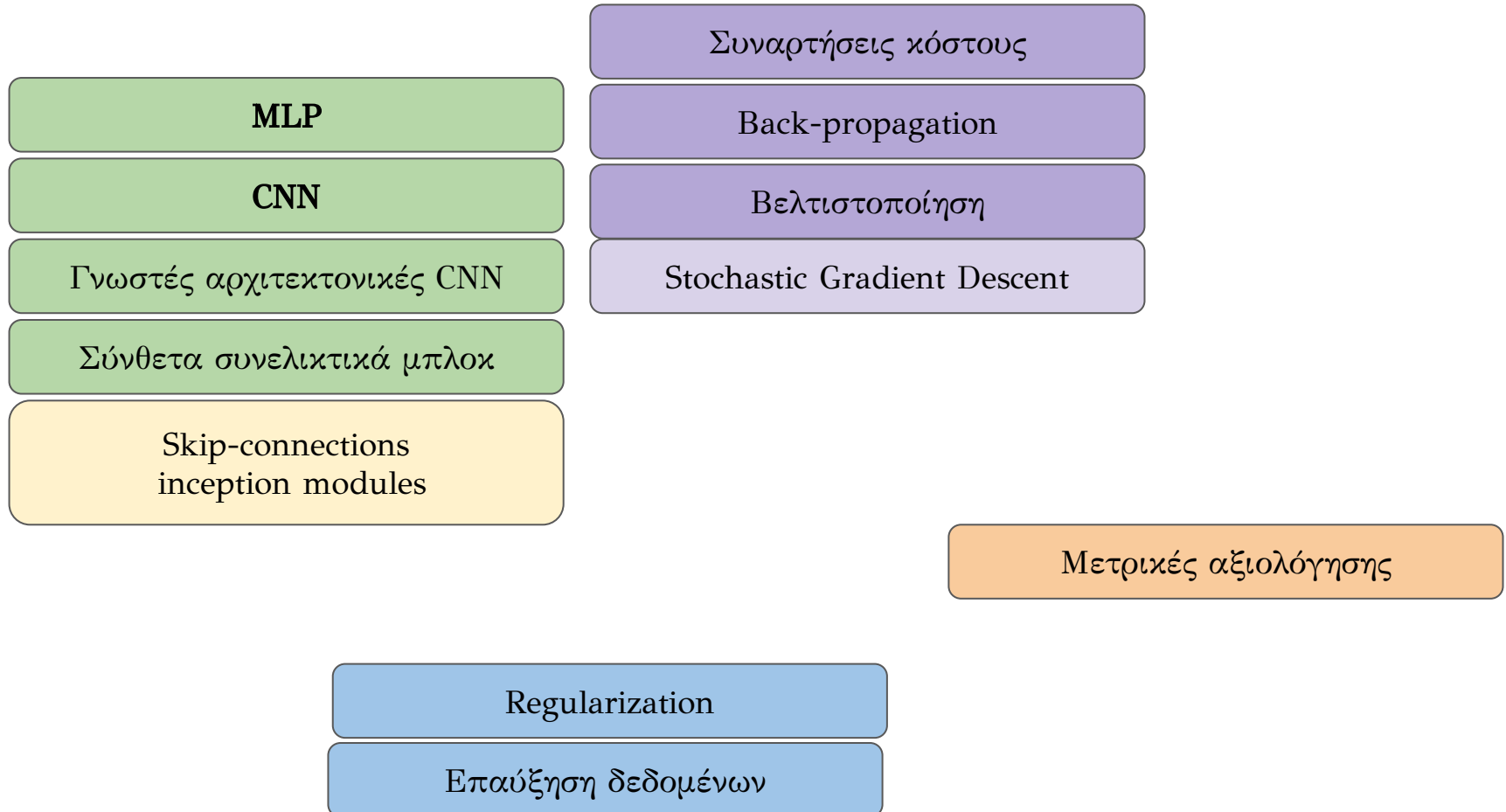
Διαχείριση και Επεξεργασία Μεγάλων
Δεδομένων Παρατήρησης Γης

Μεταφορά Μάθησης & Πλήρως Συνελικτικά Δίκτυα

Αθηνά Ψάλτα
Βασίλειος Τσιρώνης
Κωνσταντίνος Καράντζαλος

Εαρινό εξάμηνο 2022

Ανακεφαλαίωση



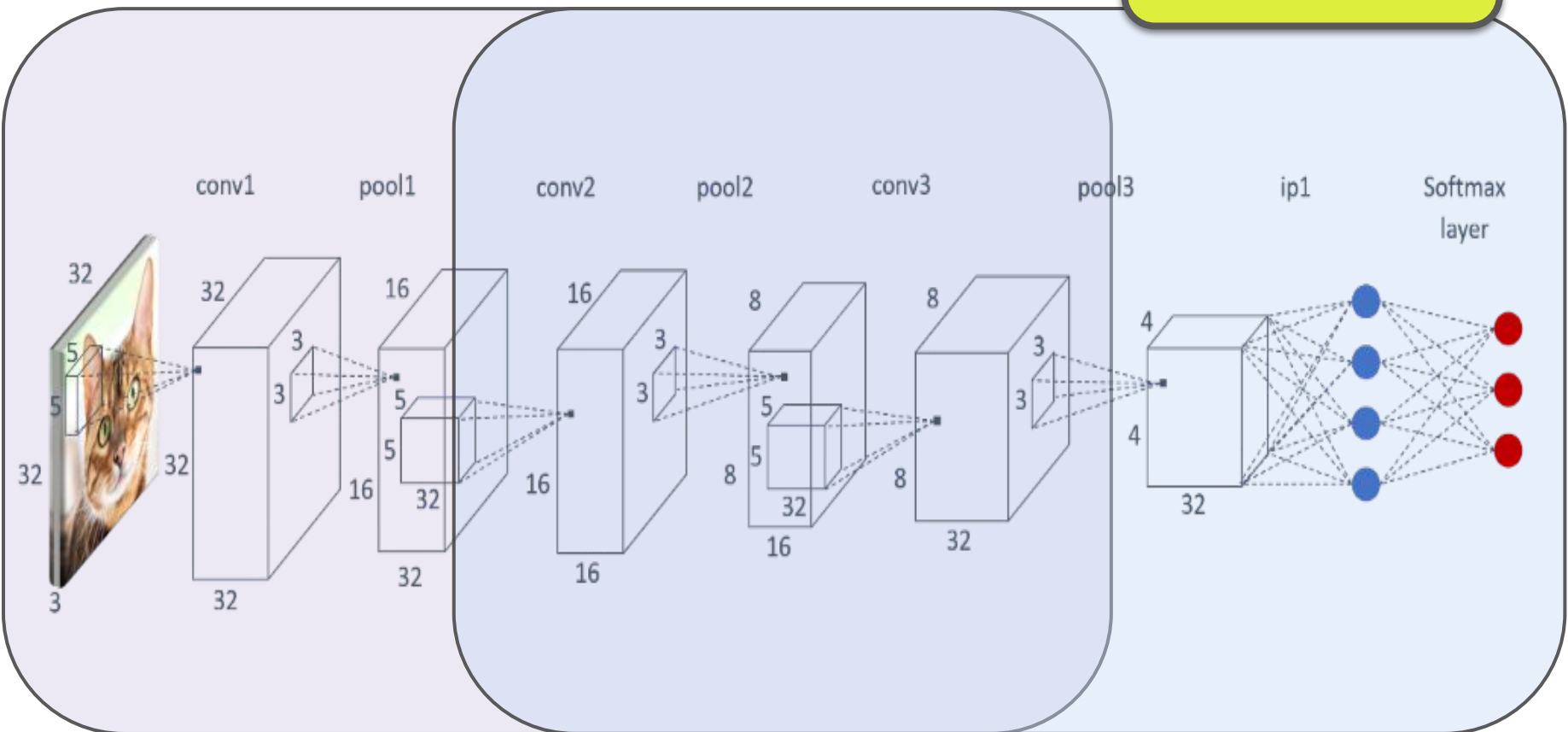
Περιεχόμενα

1. Μεταφορά Μάθησης
2. Πλήρως Συνελικτικά Δίκτυα / Fully-Convolutional Networks (FCN)
3. Δίκτυα τύπου Autoencoder
4. Το πρόβλημα της σημασιολογικής κατάτμησης (segmentation)
 - a. Συναρτήσεις κόστους και Μετρικές αξιολόγησης
 - b. FCN για κατάτμηση, δίκτυα τύπου UNet
 - c. State-of-the-art παραδείγματα

Μεταφορά Μάθησης

Χαρακτηριστικά και ταξινομητές

Τυπικό CNN

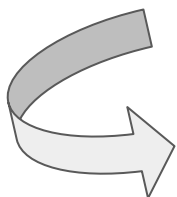


**ΜΑΘΗΣΗ
ΑΝΑΠΑΡΑΣΤΑΣΗΣ**

**ΜΗΧΑΝΙΚΗ
ΜΑΘΗΣΗ**

Βαθιές Αρχιτεκτονικές

- Χρειαζόμαστε **μεγάλα σετ δεδομένων** με διαφορετικά παραδείγματα για κάθε κατηγορία ώστε το μοντέλο που προκύπτει να έχει την ανάλογη γενίκευση και ευρωστία (robustness)
- Ανάγκη μεγάλων σετ δεδομένων → Διαφορετικές προσεγγίσεις από τους ερευνητές
 - Dense Captioning (Andrej Karpathy): 94.000 εικόνες και 4.100.00 περιγραφές (captions) → πολλά δεδομένα, απλοί αλγόριθμοι, “αδύναμα” labels
 - DeepFace by Facebook : 4.000.000 πρόσωπα και 4000 identities → πολλά δεδομένα για αποφυγή overfitting
 - Μετακίνηση ρομποτικού βραχίονα με συντονισμό ματιού & χεριού (Google Brain, Alex Krizhevsky) : 800.000 δοκιμές “πιασίματος” (grasp) για την εκμάθηση πολλών στρατηγικών



Δεν μπορούμε να χρησιμοποιήσουμε βαθιές αρχιτεκτονικές αν δεν έχουμε πολύ μεγάλα σετ δεδομένων

Βαθιές Αρχιτεκτονικές

- Χρειαζόμαστε **μεγάλα σετ δεδομένων** με διαφορετικά παραδείγματα για κάθε κατηγορία ώστε το μοντέλο που προκύπτει να έχει την ανάλογη γενίκευση και ευρωστία (robustness)
- Ανάγκη μεγάλων σετ δεδομένων → Διαφορετικές προσεγγίσεις από τους ερευνητές
 - Dense Captioning (Andrej Karpathy): 94.000 εικόνες και 4.100.00 περιγραφές (captions) → πολλά δεδομένα, απλοί αλγόριθμοι, “αδύναμα” labels
 - DeepFace by Facebook : 4.000.000 πρόσωπα και 4000 identities → πολλά δεδομένα για αποφυγή overfitting
 - Μετακίνηση ρομποτικού βραχίονα με συντονισμό ματιού & χεριού (Google Brain, Alex Krizhevsky) : 800.000 δοκιμές “πιασίματος” (grasp) για την εκμάθηση πολλών στρατηγικών



Δεν μπορούμε να χρησιμοποιήσουμε βαθιές αρχιτεκτονικές αν δεν έχουμε πολύ μεγάλα σετ δεδομένων

Βαθιές Αρχιτεκτονικές

Στην πραγματικότητα μπορούμε να :

- Μάθουμε **νέες αναπαραστάσεις** από δεδομένα που δεν διαθέτουν δεδομένα αληθείας (unlabeled data)
- Εκπαιδεύσουμε σε μια **κοντινή “θεματικά”** εφαρμογή για την οποία είναι εύκολη η παραγωγή δεδομένων αληθείας
- **Μεταφέρουμε** ήδη γνωστές αναπαραστάσεις/χαρακτηριστικά για μια παρεμφερή θεματικά εφαρμογή
 - π.χ. Ταξινόμηση εικόνας (image classification) & ανίχνευση αντικειμένου (object detection)

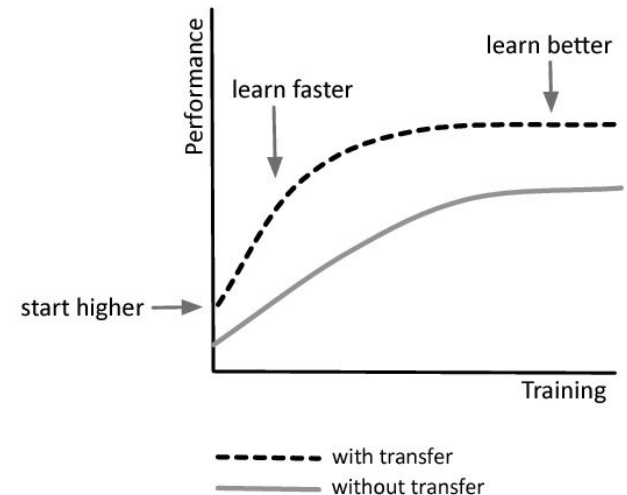
Μεταφορά Μάθησης : Βασική Ιδέα

Τι είναι;

- Η χρήση “τμήματος” από ένα προεκπαιδευμένο μοντέλο για τη δημιουργία ενός διαφορετικού μοντέλου που επιλύει ένα διαφορετικό πρόβλημα

Σε ποιες περιπτώσεις είναι χρήσιμη;

- Ελλιπής όγκος δεδομένων
- Περιορισμένοι υπολογιστικοί πόροι



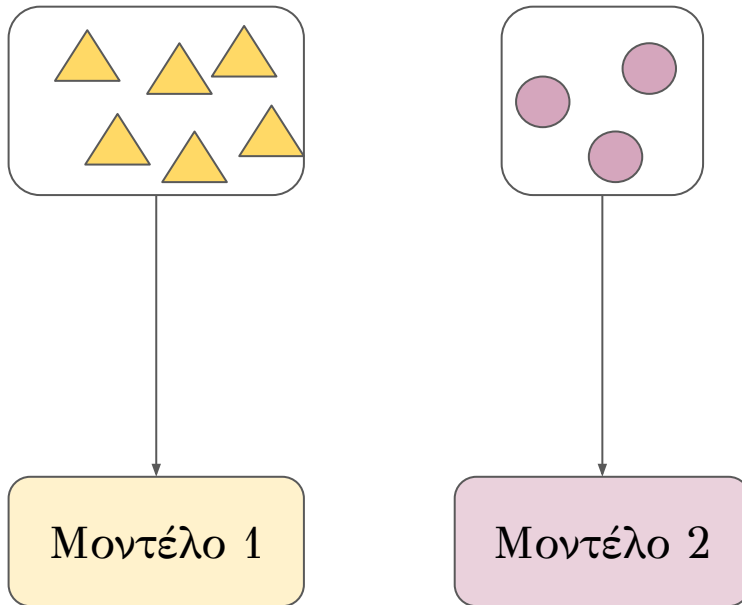
Μεταφορά Μάθησης : Βασική Ιδέα

Αντί να εκπαιδεύσουμε ένα δίκτυο από την αρχή μπορούμε να :

- A. Χρησιμοποιήσουμε ένα δίκτυο που έχει εκπαιδευτεί σε έναν διαφορετικό τομέα (domain) για να αποκτήσουμε **γνώση** για ένα παρεμφερές θεματικά πρόβλημα
- B. Προσαρμόσουμε το δίκτυο για τον τομέα του προβλήματός μας (domain) και την εφαρμογή μας (task) με την αξιοποίηση αυτής της γνώσης

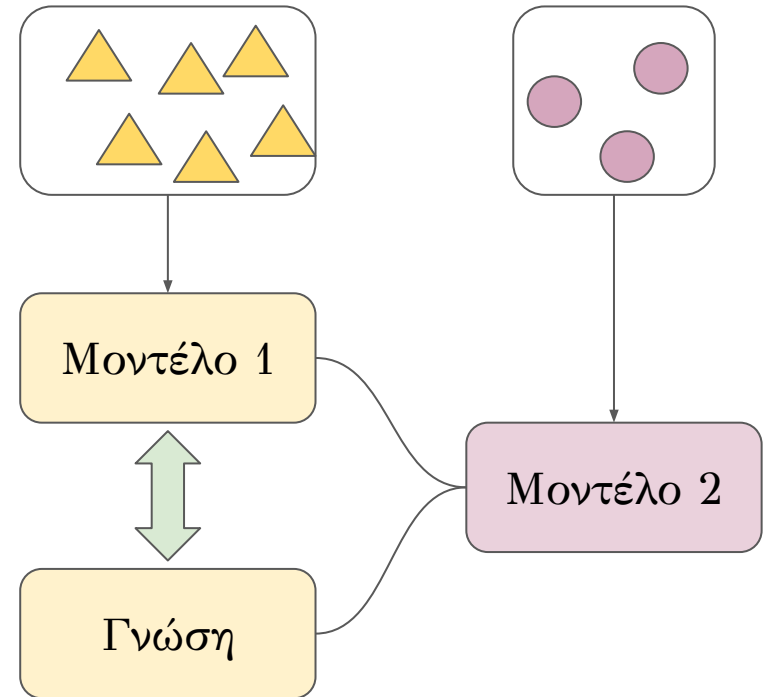
Η **μεταφορά μάθησης (transfer learning)** βασίζεται στην ιδέα ότι για να ξεπεράσουμε ένα μεμονωμένο learning paradigm πρέπει να μπορούμε να αξιοποιήσουμε γνώση που έχουμε αποκτήσει από παρόμοια paradigms.

Κλασική MM



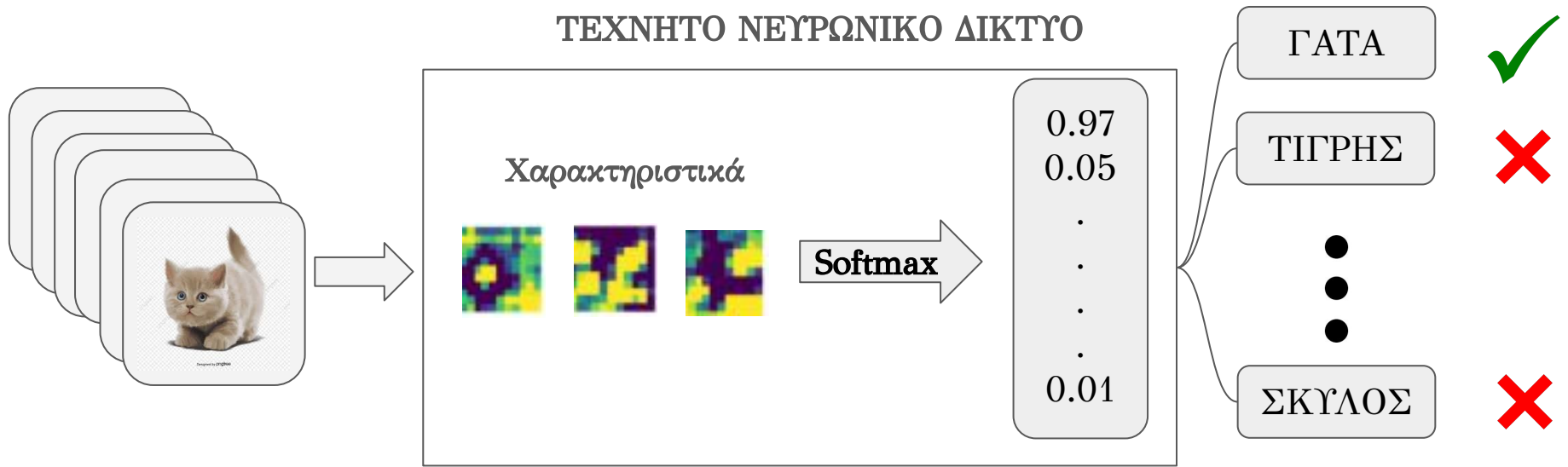
Χωρίς να λαμβάνουμε υπόψη πρότερη γνώση

Μεταφορά Μάθησης

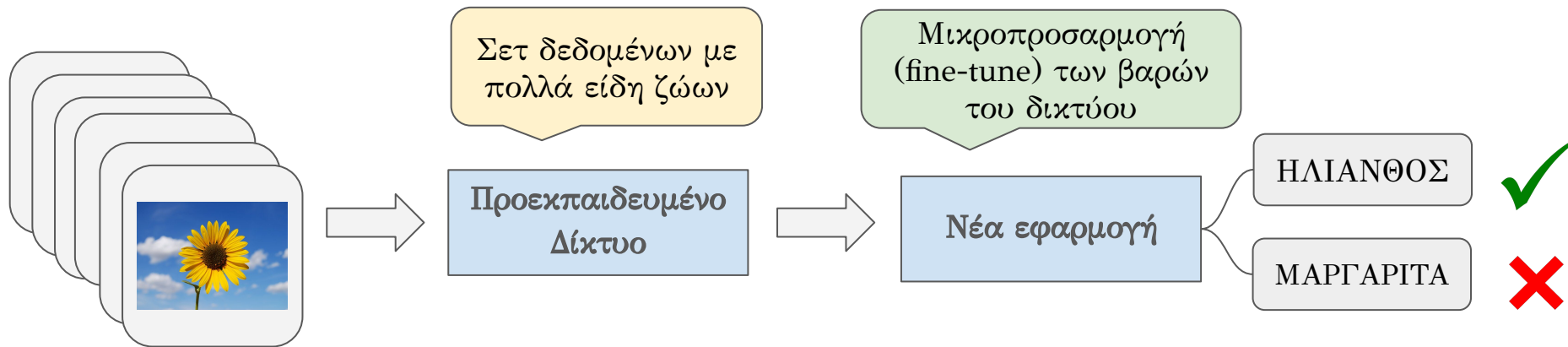


Πρότερη γνώση από εκπαίδευση σε “όμορη” εφαρμογή

ΚΛΑΣΙΚΗ ΔΙΑΔΙΚΑΣΙΑ ΕΚΠΑΙΔΕΥΣΗΣ



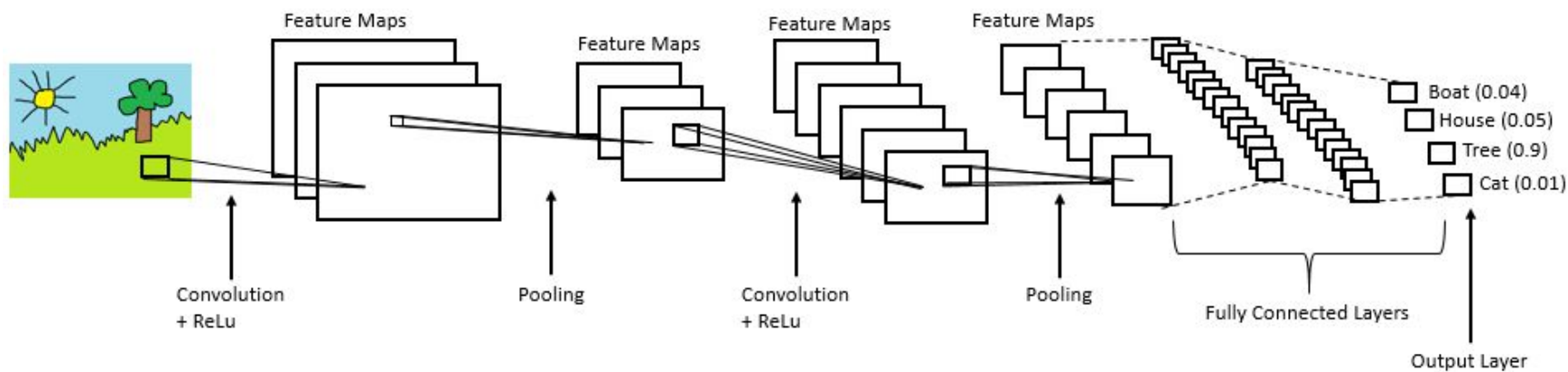
ΕΚΠΑΙΔΕΥΣΗ ΜΕ ΜΕΤΑΦΟΡΑ ΜΑΘΗΣΗΣ



Νέα δεδομένα

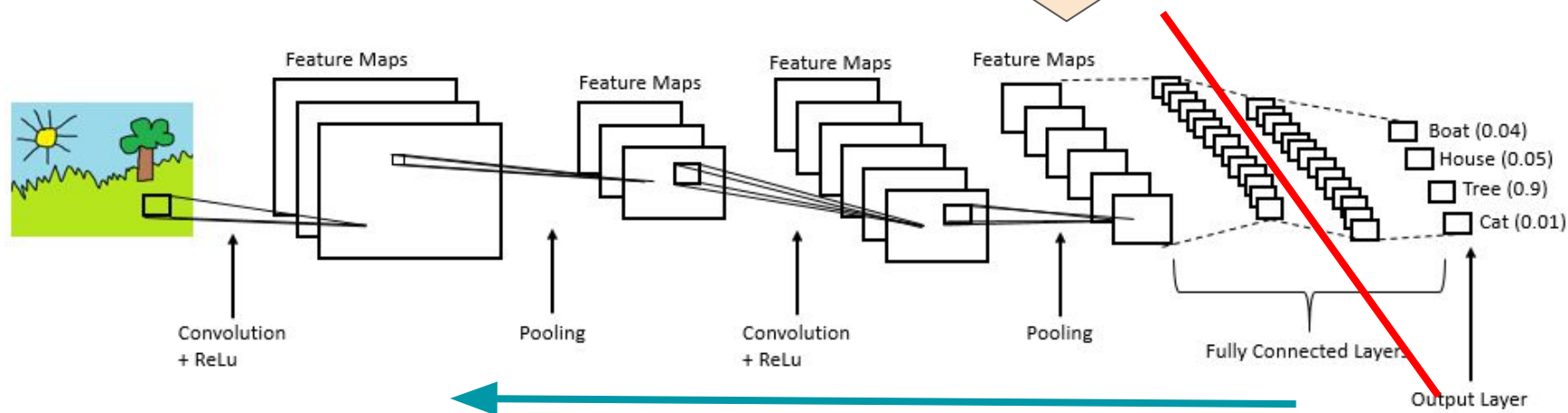
Μεταφορά Μάθησης : Πως;

Σε κάθε
περίπτωση
ξεκινάμε από ένα
**προεκπαιδευμένο
μοντέλο**



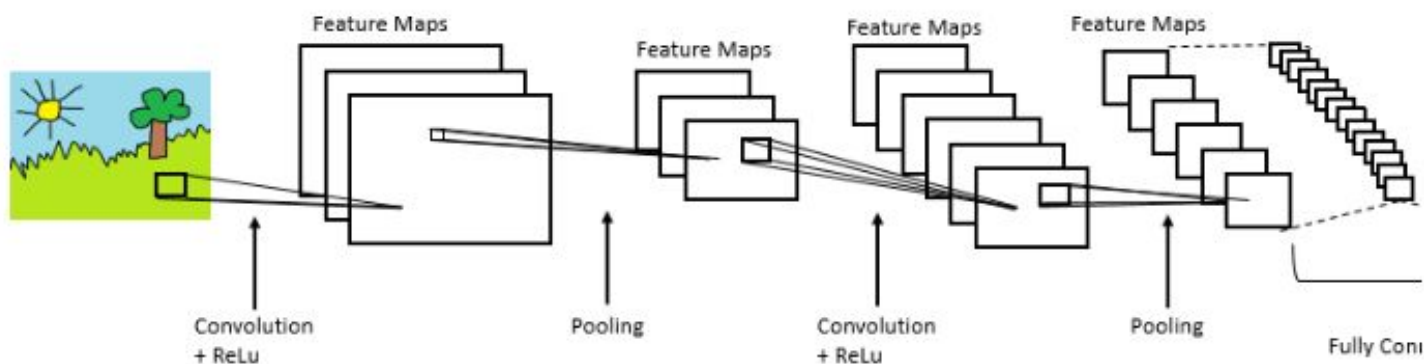
Μεταφορά Μάθησης : Πως;

Από το αρχικό
μοντέλο επιλέγουμε
να διατηρήσουμε τις
πρώτες η στρώσεις



Μεταφορά Μάθησης : Πως;

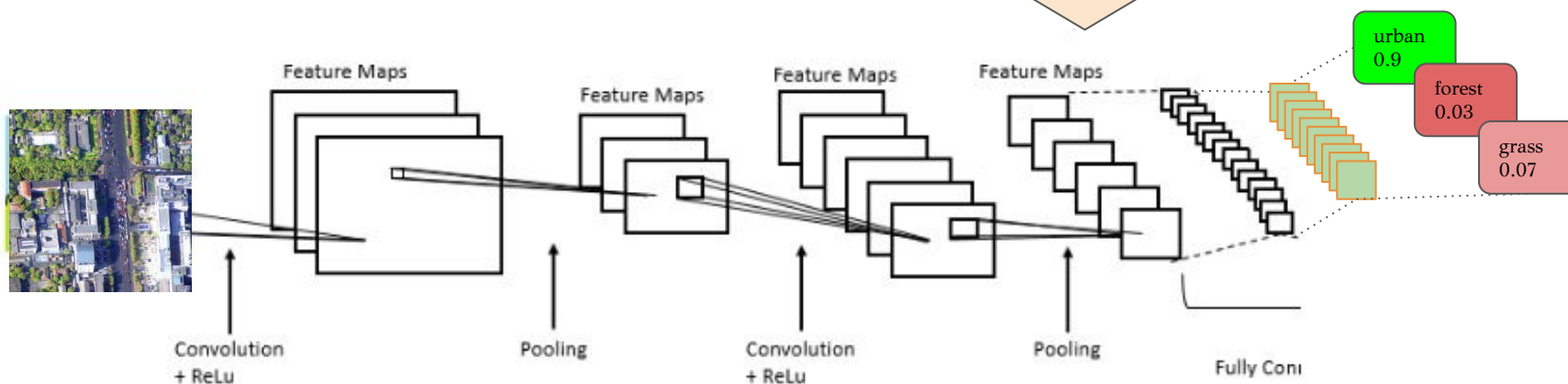
Το δίκτυο που απομένει μπορούμε να το θεωρήσουμε ως ένα **Feature Extractor**



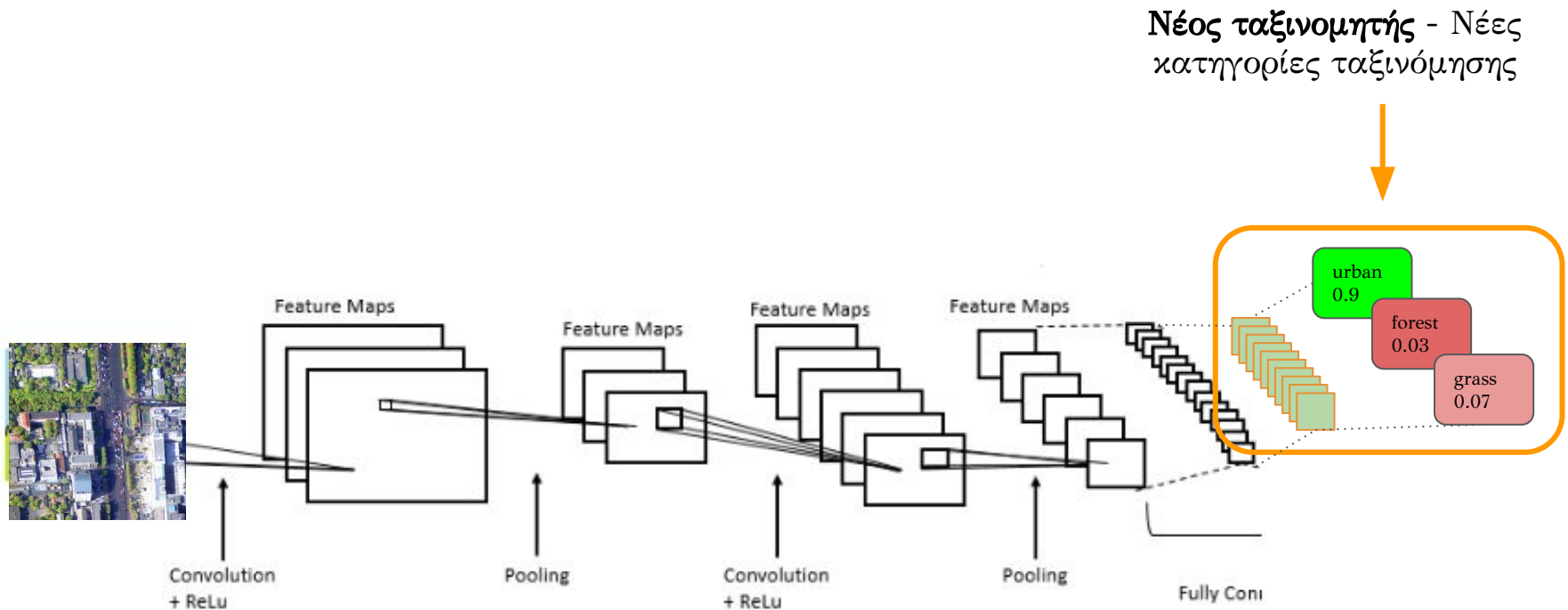
Πρακτικά, χρησιμοποιούμε το τμήμα του δικτύου που απομονώσαμε ως ένα μέσο “**προβολής**” των δεδομένων σε ένα **χώρο προτύπων** κατάλληλο για ταξινόμηση.

Μεταφορά Μάθησης : Πως;

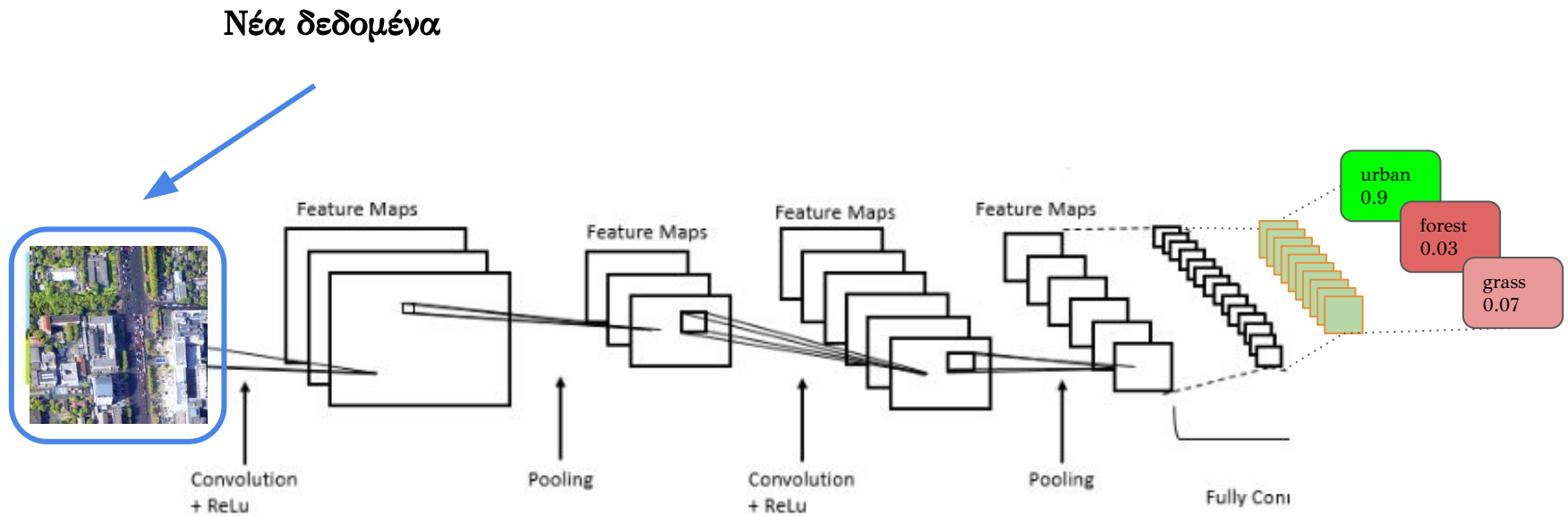
Για να ταξινομήσουμε
νέα δεδομένα, απλά
προσαρτούμε στο
τέλος του δικτύου έναν
νέο ταξινομητή



Μεταφορά Μάθησης : Πως;



Μεταφορά Μάθησης : Πως;



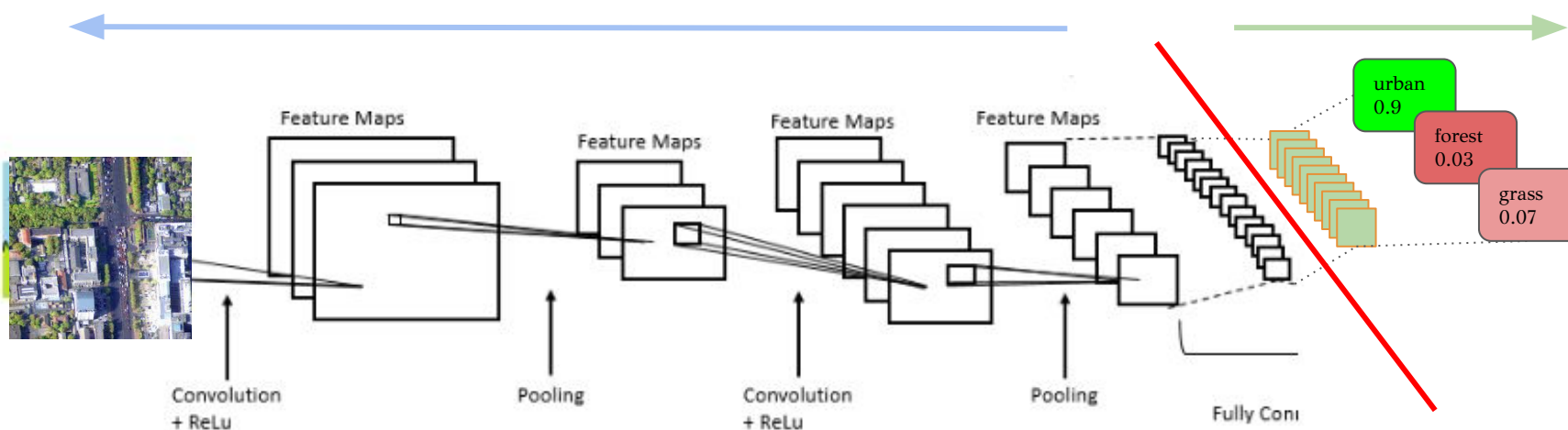
Μεταφορά Μάθησης : Εκπαίδευση

Γενικά υπάρχουν 2 τρόποι να εκπαιδευτεί ένα δίκτυο που έχει προκύψει από “μεταφορά μάθησης”

1ος Τρόπος: Εκπαίδευση **μόνο** του νέου ταξινομητή

“Πάγωμα” βαρών

Εκπαίδευση τελικών στρώσεων

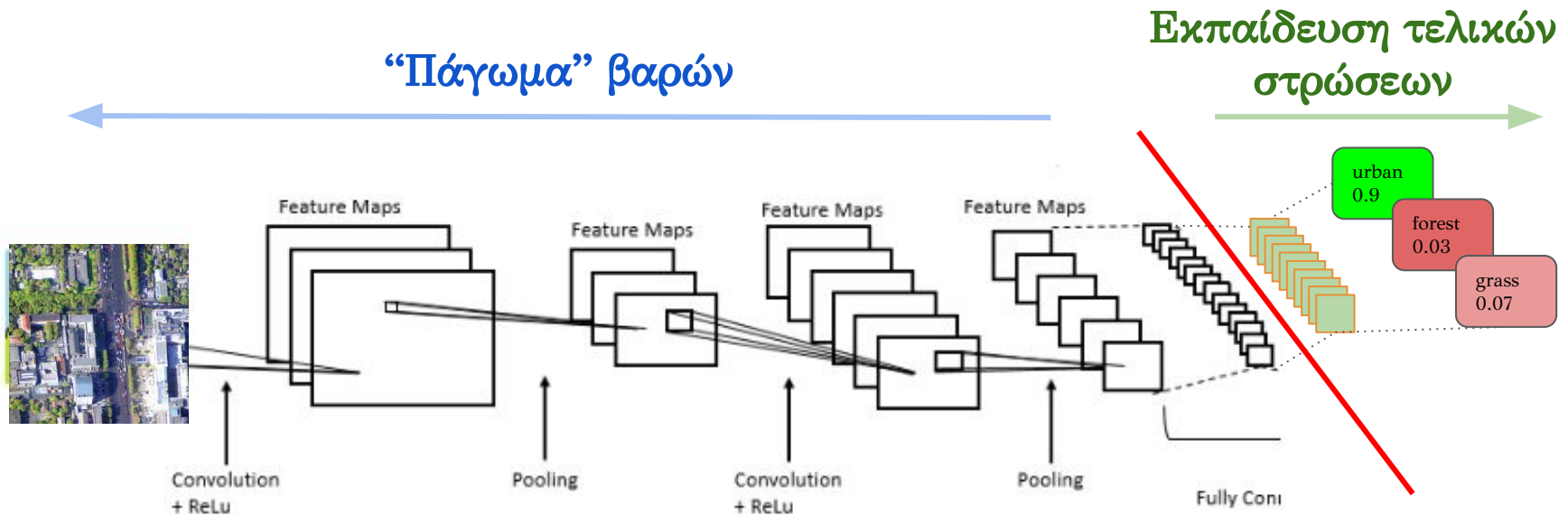


Μεταφορά Μάθησης : Εκπαίδευση

1ος Τρόπος: Εκπαίδευση **μόνο** του νέου ταξινομητή

Πλεονεκτήματα: Λίγοι άγνωστοι παράμετροι → Μειωμένες απαιτήσεις σε annotated δεδομένα και υπολογιστικούς πόρους

Μειονεκτήματα: Εμπιστευόμαστε “τυφλά” τα χαρακτηριστικά του προεκπαιδευμένου δικτύου

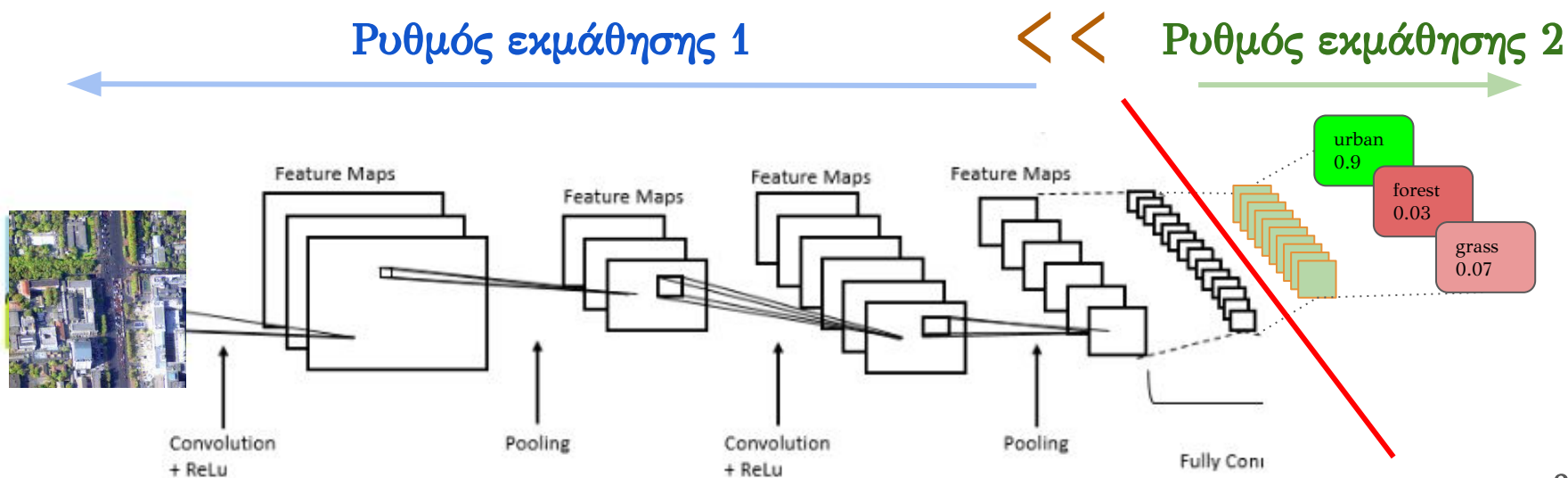


Μεταφορά Μάθησης : Εκπαίδευση

➤ Εκπαίδευση

Γενικά υπάρχουν 2 τρόποι να εκπαιδευτεί ένα δίκτυο που έχει προκύψει από “μεταφορά μάθησης”

2ος Τρόπος: Εκπαίδευση **όλου του δικτύου** με διαφορετικούς ρυθμούς εκμάθησης

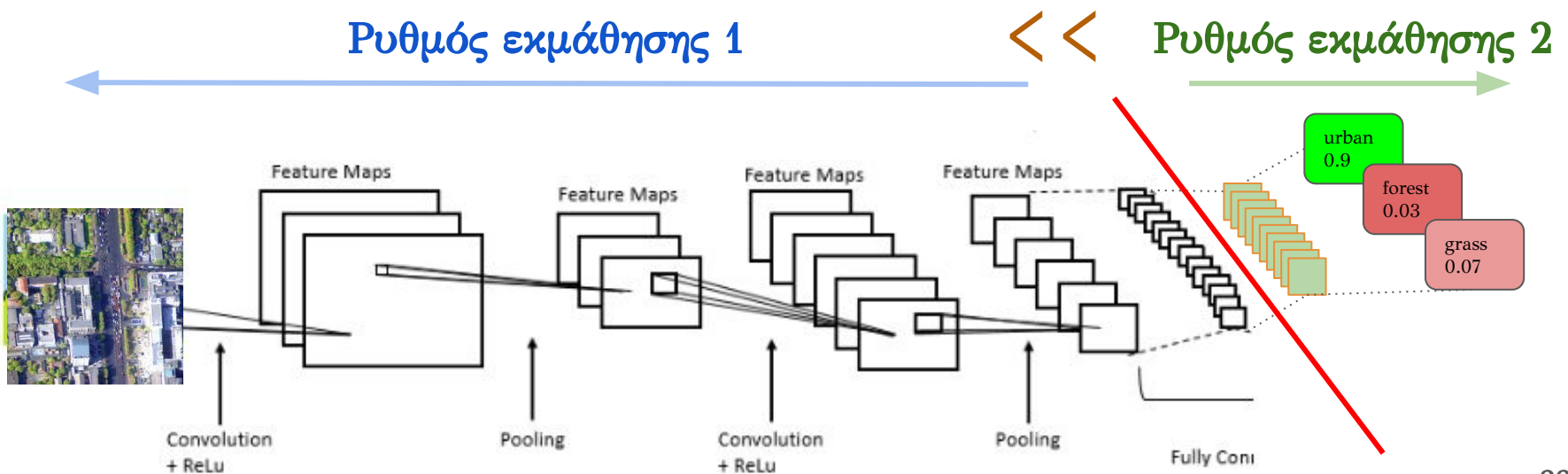


Μεταφορά Μάθησης : Εκπαίδευση

2ος Τρόπος: Εκπαίδευση με διαφορετικούς ρυθμούς εκμάθησης

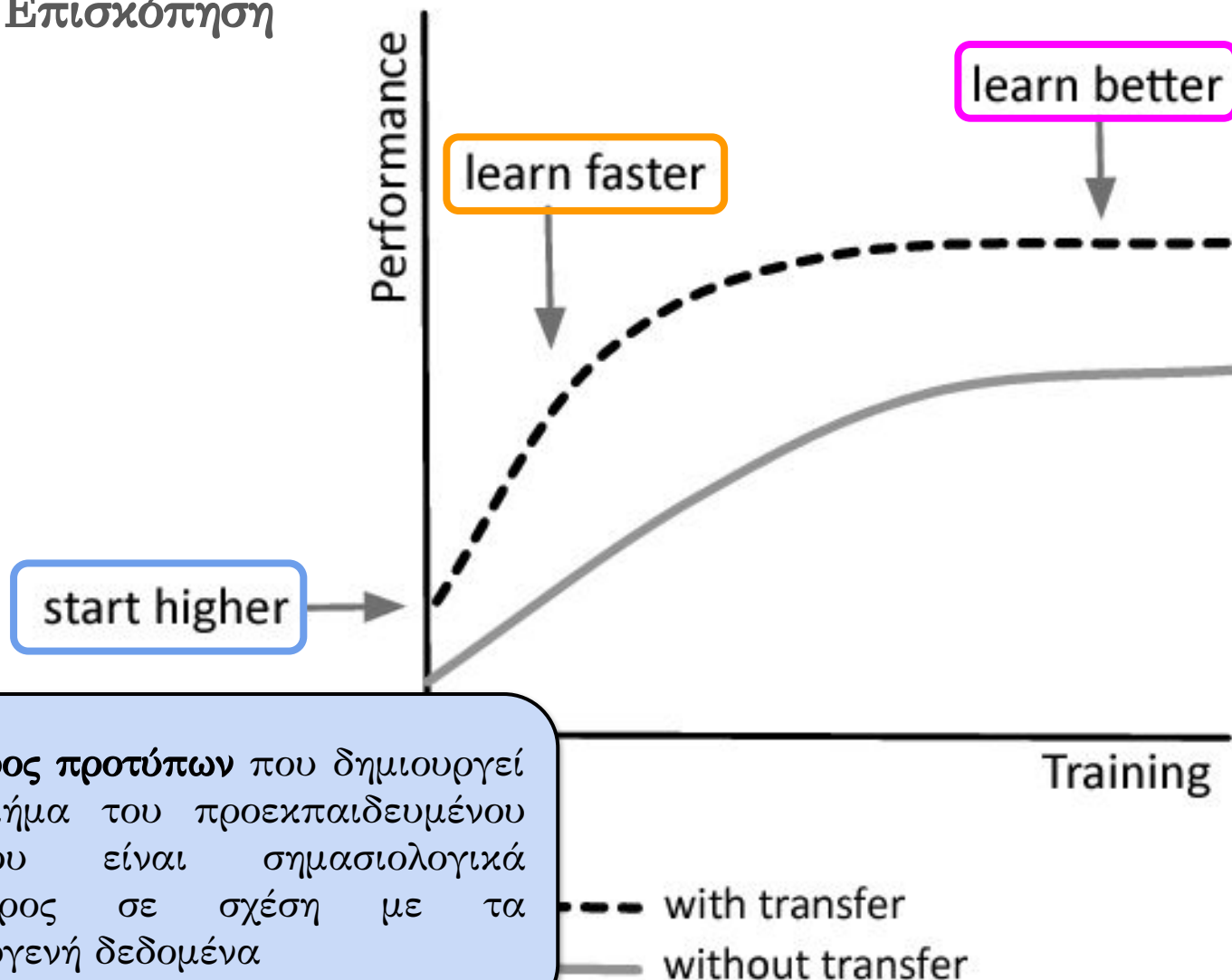
Πλεονεκτήματα: Εκπαίδευση νέου ταξινομητή και παράλληλα “μικρο-προσαρμογή” των βαρών του προεκπαιδευμένου δικτύου (**fine-tuning**)

Μειονεκτήματα: Μεγαλύτερη ανάγκη για υπολογιστικούς πόρους (κατά πολύ περισσότεροι παράμετροι προς εκπαίδευση), όχι όμως τεράστιες απαιτήσεις για annotated δεδομένα



Μεταφορά Μάθησης

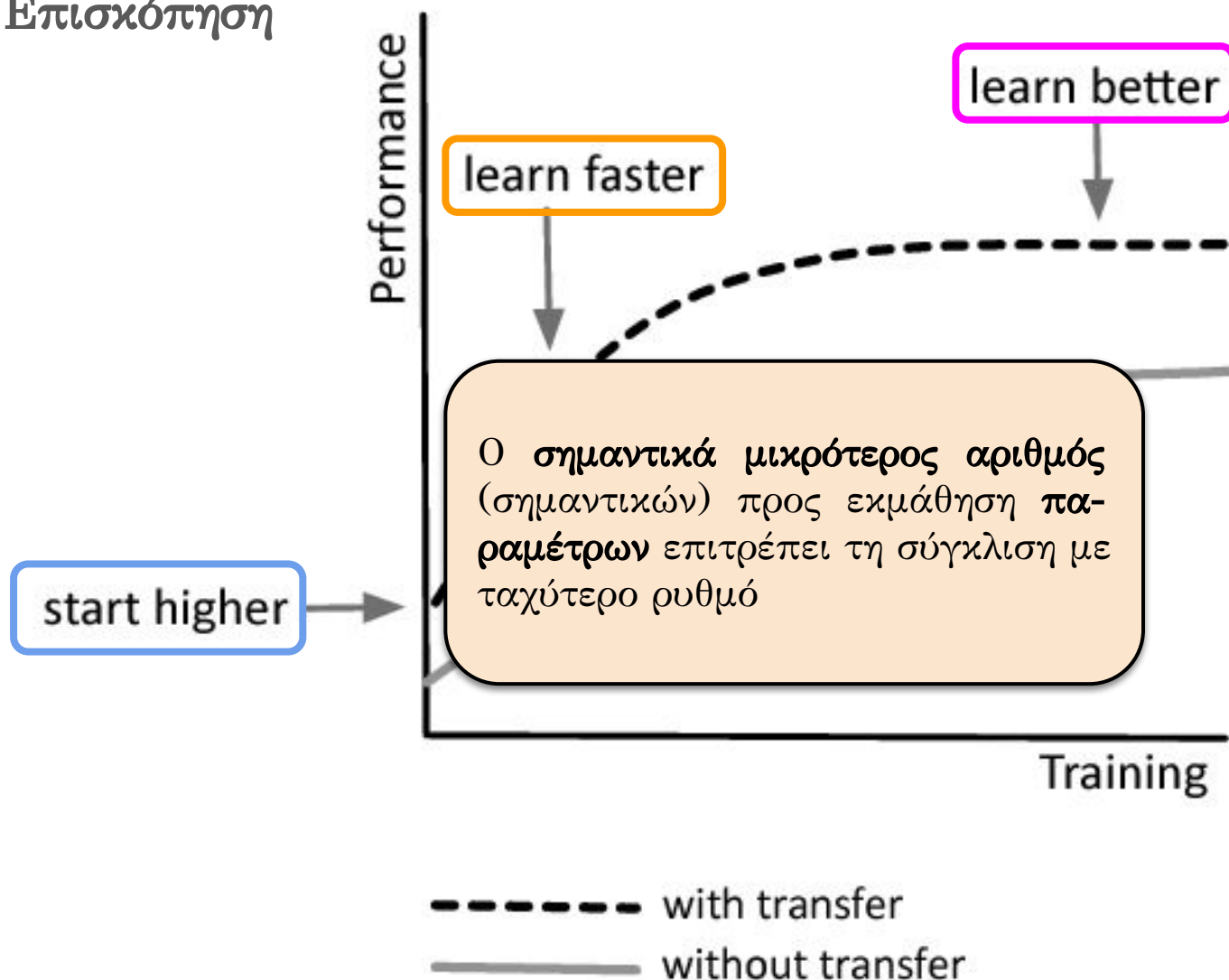
➤ Επισκόπηση



Ο χώρος προτύπων που δημιουργεί το τμήμα του προεκπαιδευμένου δικτύου είναι σημασιολογικά ανώτερος σε σχέση με τα πρωτογενή δεδομένα

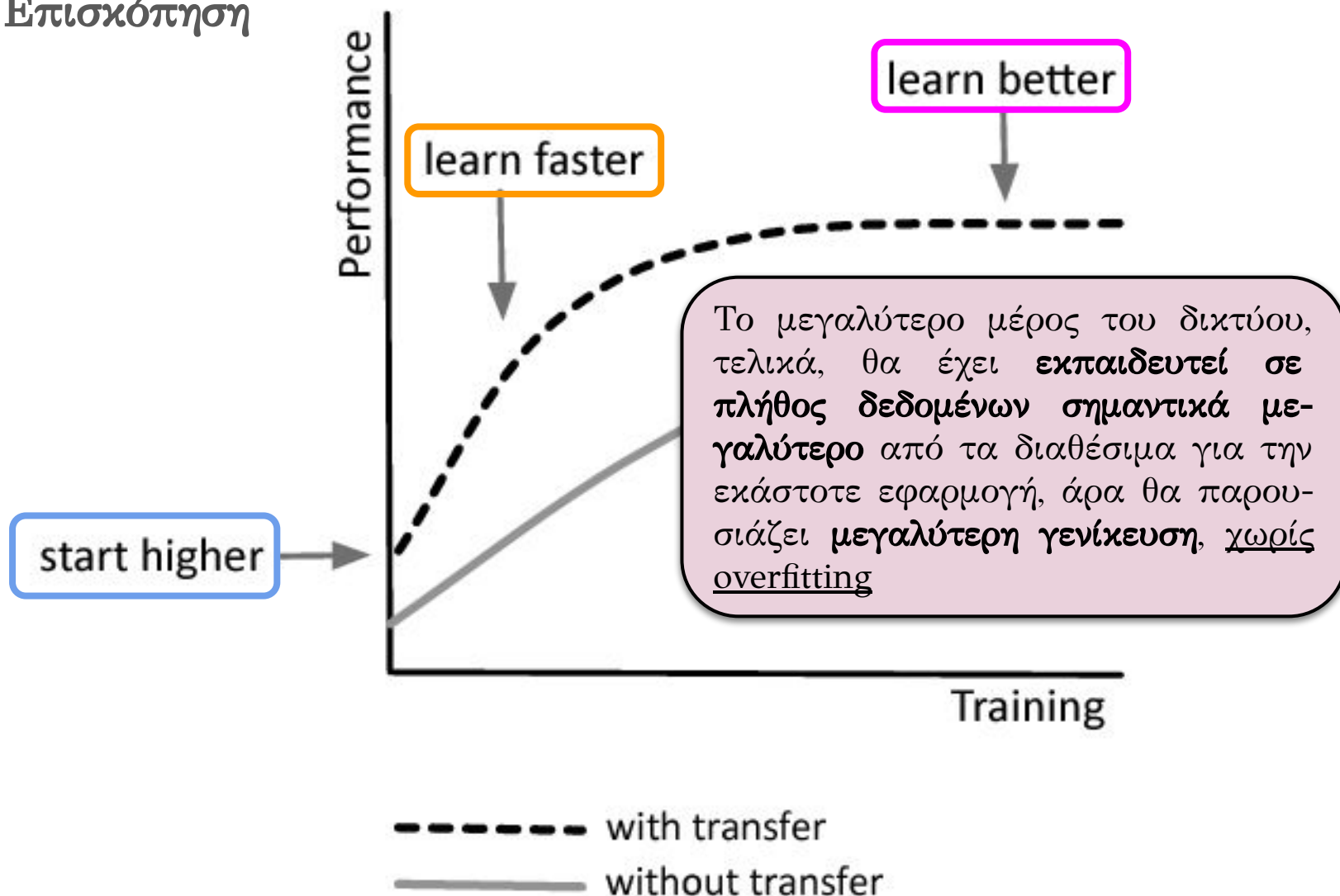
Μεταφορά Μάθησης

➤ Επισκόπηση



Μεταφορά Μάθησης

➤ Επισκόπηση



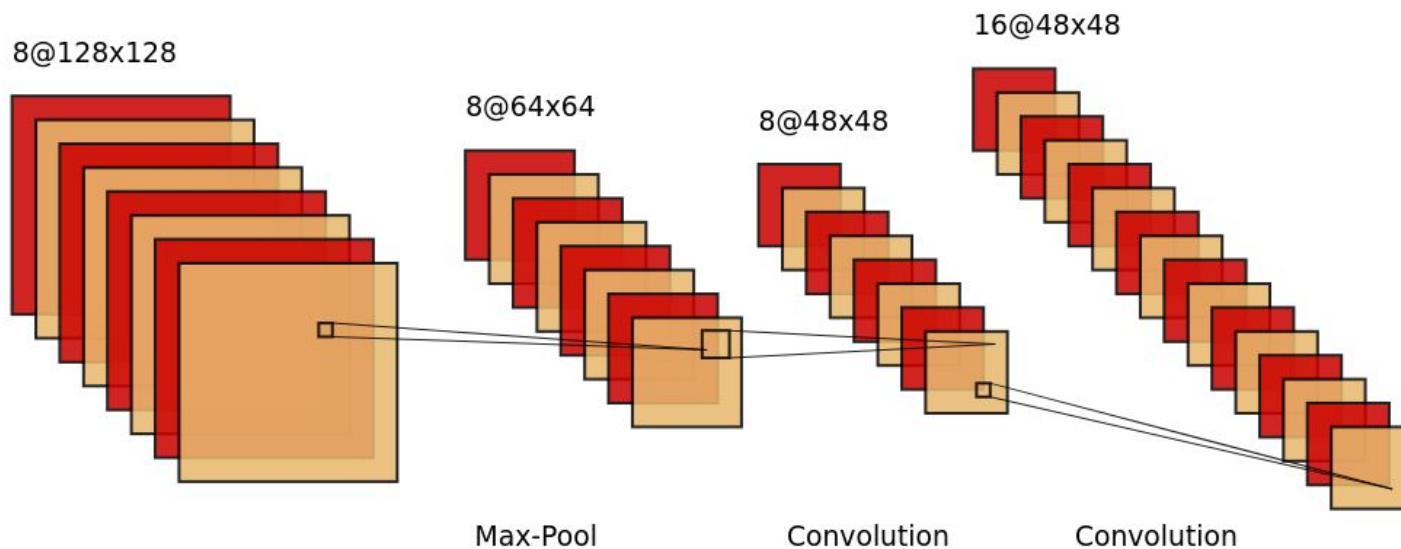
Πλήρως Συνελικτικά Δίκτυα

Δομή FCN, Auto-Encoders,
Σημασιολογική Κατάτμηση

Πλήρως Συνελικτικά δίκτυα

Ως πλήρως συνελικτικά δίκτυα (**Fully Convolutional Networks - FCN**) αναφέρονται τα ΤΝΔ των οποίων οι ενδιάμεσες στρώσεις αποτελούνται από συνελικτικές στρώσεις και στρώσεις μεταβολής χωρικών διαστάσεων (πχ pooling).

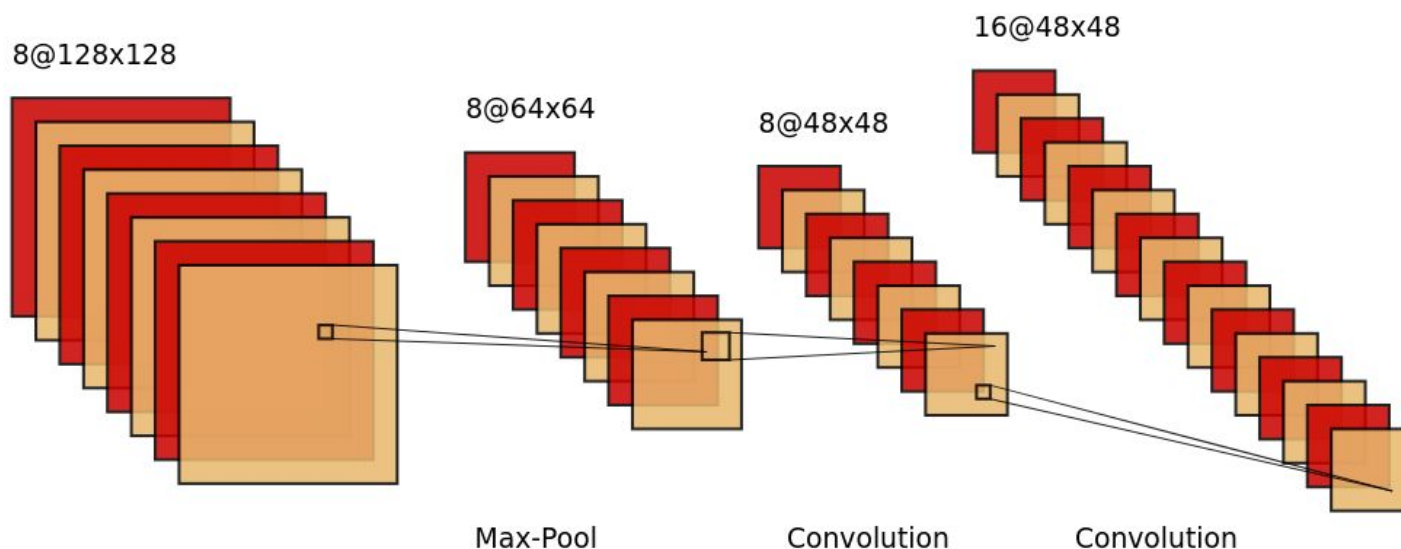
- Σε κάθε ενδιάμεση στρώση τα χαρακτηριστικά είναι ταχυστές τουλάχιστον 3ης τάξης
- Είσοδος εικόνα → έξοδος εικόνα



Πλήρως Συνελικτικά δίκτυα

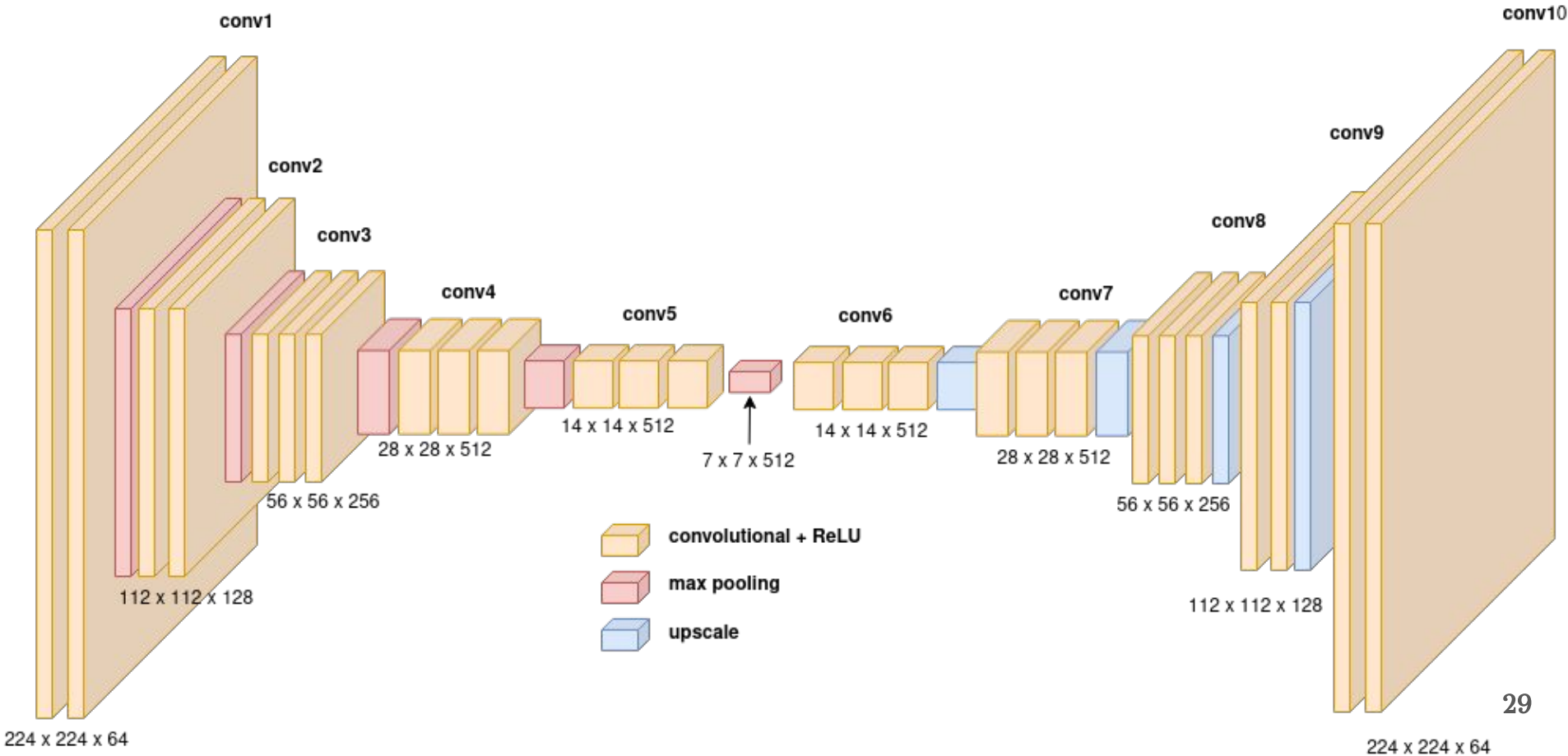
Ως πλήρως συνελικτικά δίκτυα (**Fully Convolutional Networks - FCN**) αναφέρονται τα ΤΝΔ των οποίων οι ενδιάμεσες στρώσεις αποτελούνται από συνελικτικές στρώσεις και στρώσεις μεταβολής χωρικών διαστάσεων (πχ pooling).

- **Δεν** εμπεριέχονται πλήρως συνδεδεμένες στρώσεις
- **Δεν** εμπεριέχονται τελεστές ισοπέδωσης (“flatten”)



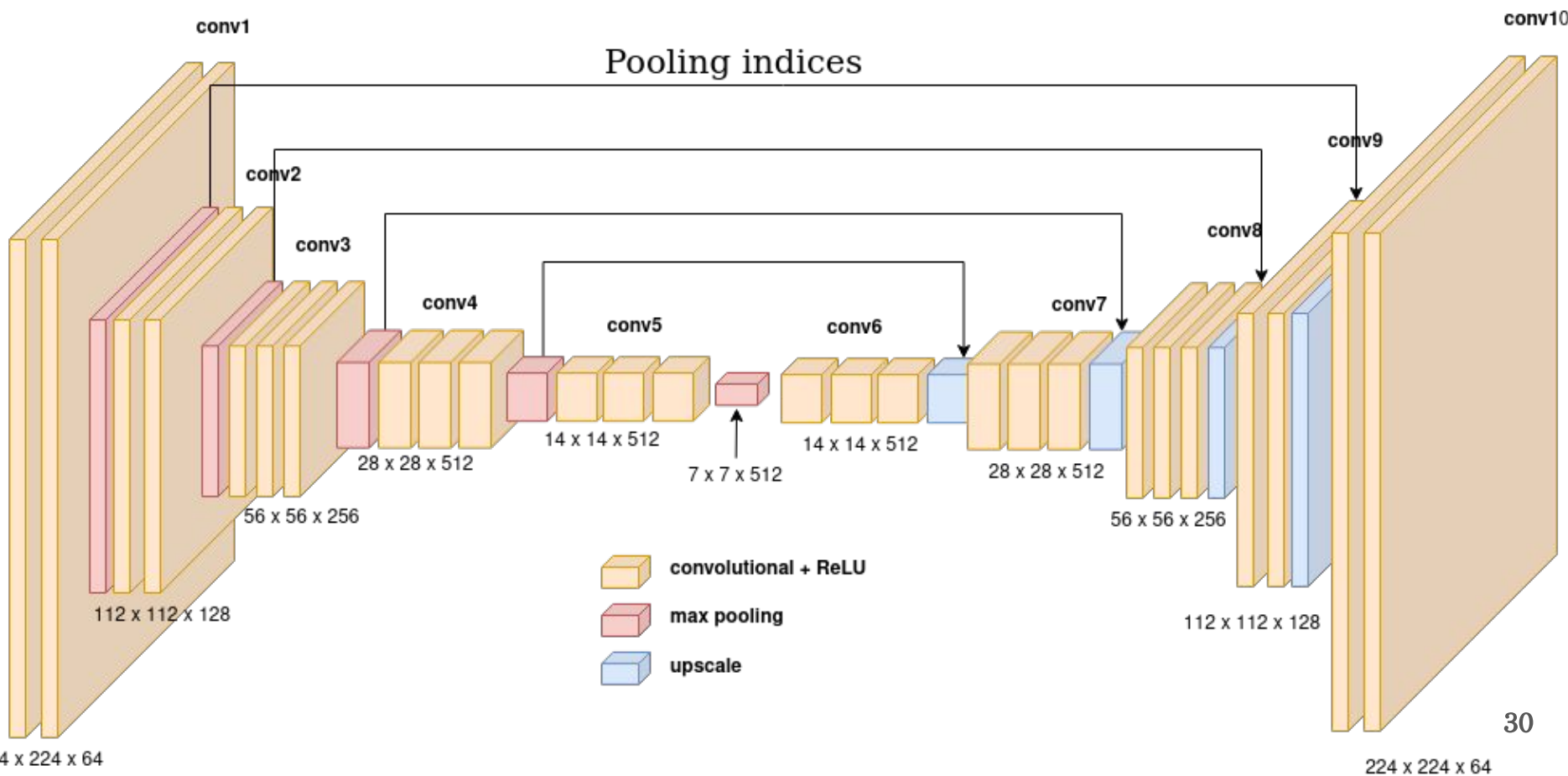
Autoencoders

Τα δίκτυα τύπου Autoencoder μετατρέπουν ένα τανυστή εισόδου σε ένα τανυστή εξόδου ίδιας τάξης και ίδιου μεγέθους



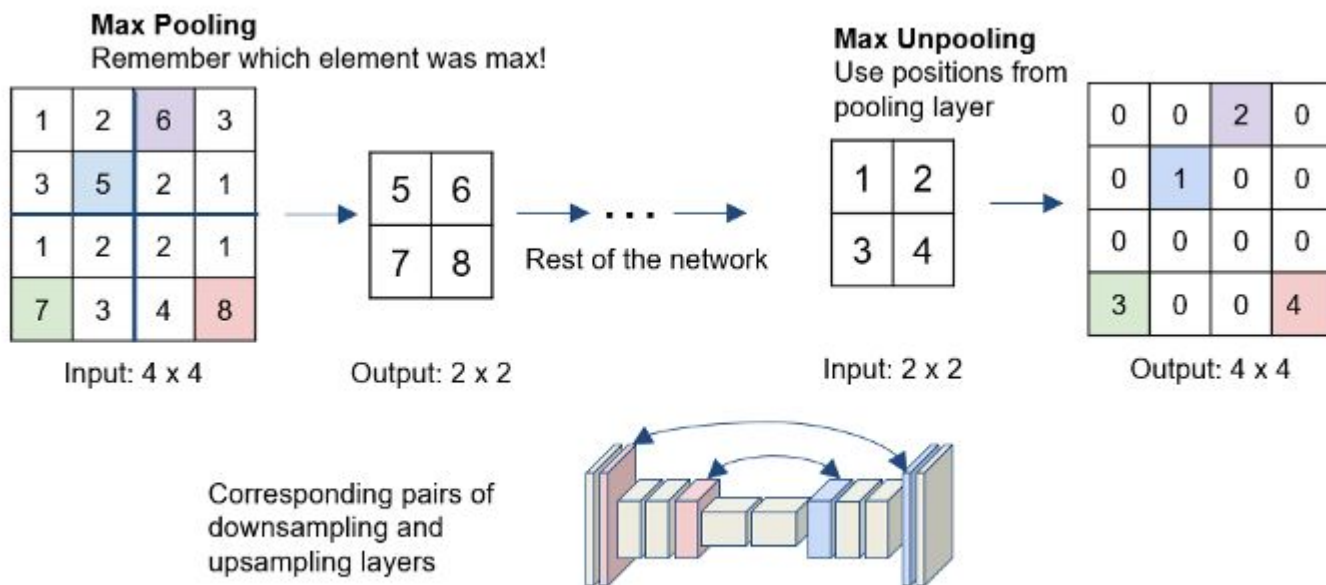
Autoencoders

Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι στρώσεις upscale

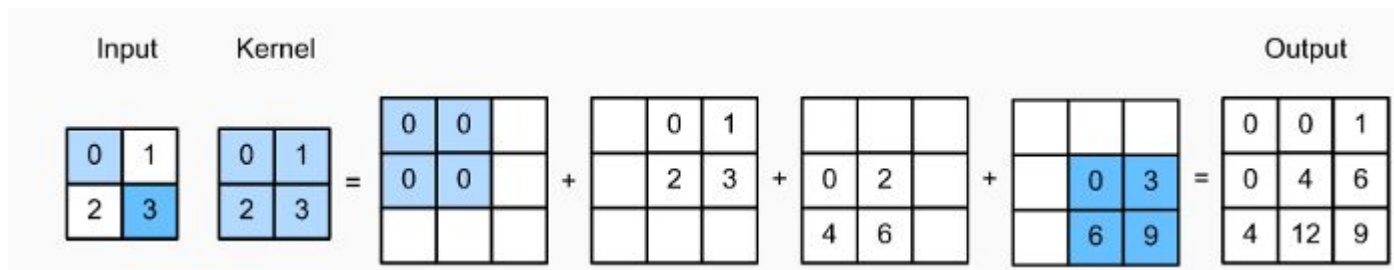


Upscale τεχνικές

Max-Unpooling: Χρησιμοποιώντας την πληροφορία από την “ομόλογη” max-pooling στρώση, η θέση που κατείχε την μέγιστη τιμή είναι αυτή που λαμβάνει την τιμή στο νέο τανυστή. Οι υπόλοιπες θέσεις λαμβάνουν την τιμή “0”.



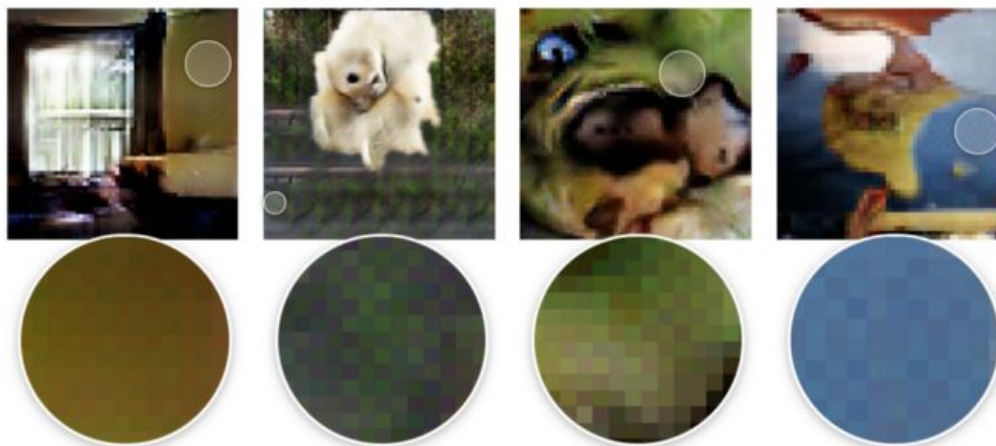
Upscale



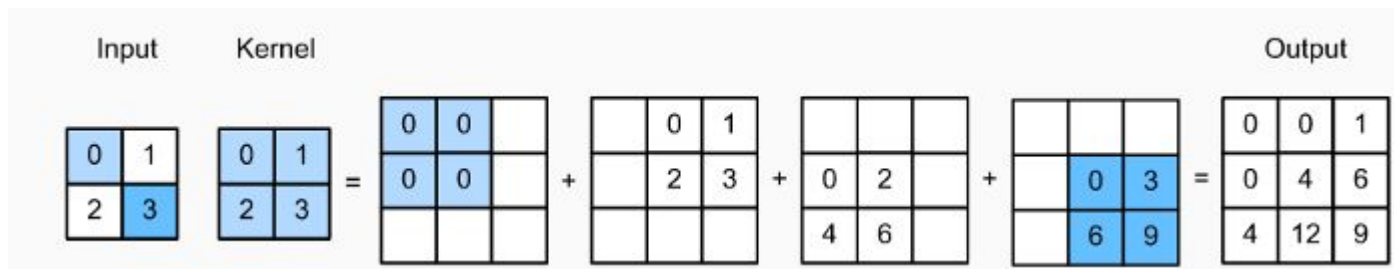
Ανάστροφη συνέλιξη (Transposed Convolution)

Πράξη η οποία εκτελεί upsampling της εικόνας αξιοποιώντας ένα σύνολο παραμέτρων προς εκμάθηση (βάρη πυρήνα).

Πολλές φορές μπορεί να προκύψει το πρόβλημα του checkerboarding ως κατάλοιπο στην εικόνα. Αυτό μπορεί να διορθωθεί μερικώς αν το μέγεθος πυρήνα διαιρείται από το stride



Upscale



Ανάστροφη συνέλιξη (Transposed Convolution)

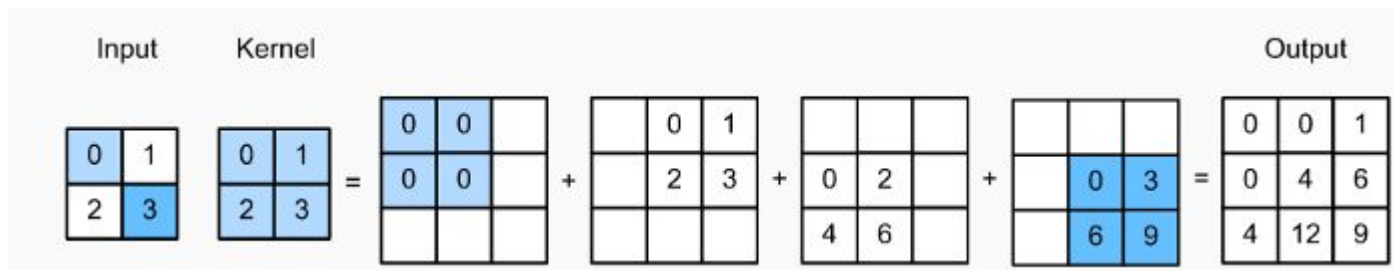
Πράξη η οποία εμπεριέχει ένα σύνολο παραμέτρων προς εκμάθηση (βάρος πυρήνα). Το stride αποτελεί σημαντική υπερπαραμέτρο.

Πολλές φορές μπορεί να προκύψει το πρόβλημα του checkerboarding ως κατάλοιπο στην εικόνα. Αυτό μπορεί να διορθωθεί μερικώς αν το μέγεθος πυρήνα διαιρείται από το stride

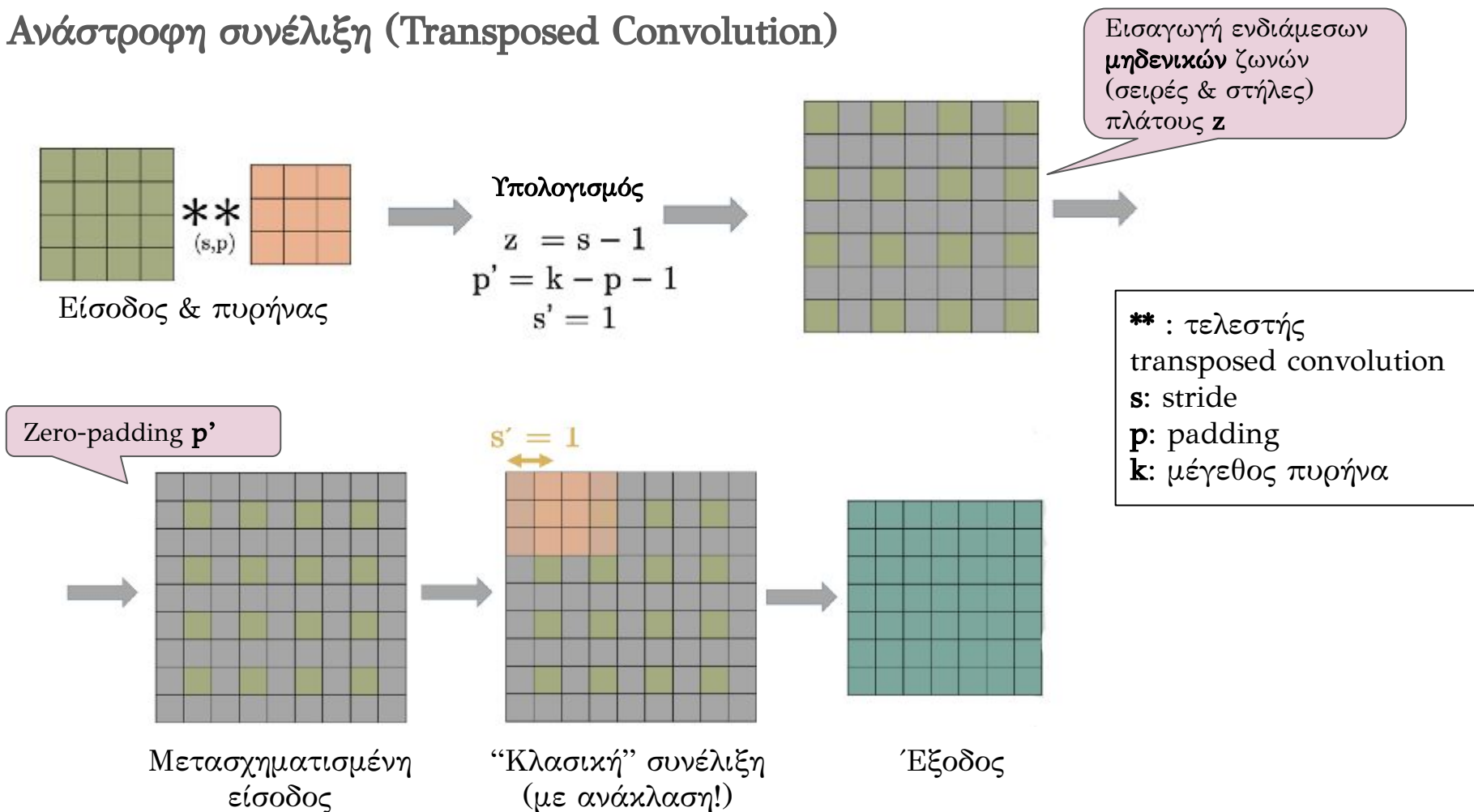


Η ανάστροφη συνέλιξη δεν αποτελεί “αντίστροφη” συνέλιξη με την μαθηματική έννοια. Πρακτικά το μόνο που “αντιστρέφεται” είναι οι (χωρικές) διαστάσεις των τανυστών εισόδου/εξόδου!

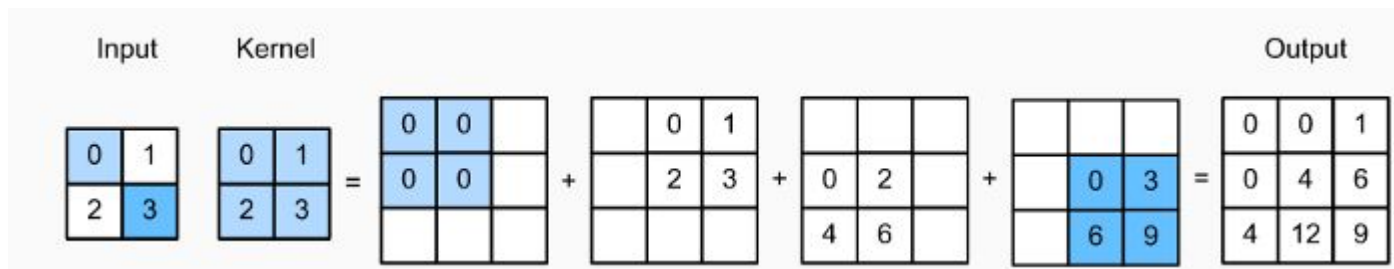
Upscale



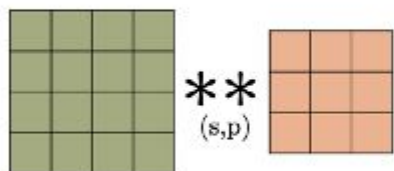
Ανάστροφη συνέλιξη (Transposed Convolution)



Upscale



Ανάστροφη συνέλιξη (Transposed Convolution)



Είσοδος & πυρήνας

$**$
(s,p)

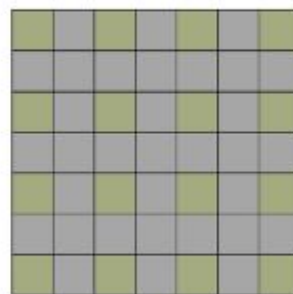


Υπολογισμός

$$z = s - 1$$

$$p' = k - p - 1$$

$$s' = 1$$

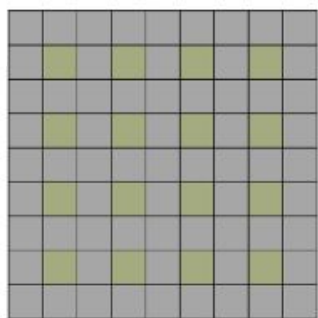


Εισαγωγή ενδιάμεσων
μηδενικών ζωνών
(σειρές & στήλες)
πλάτους z

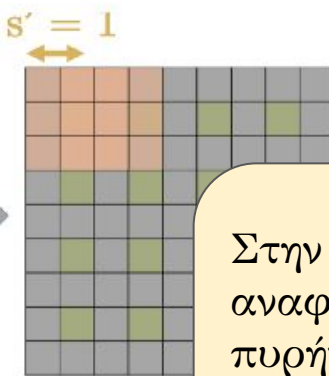


$**$: τελεστής
transposed convolution
 s : stride
 p : padding
 k : μέγεθος πυρήνα

Zero-padding p'



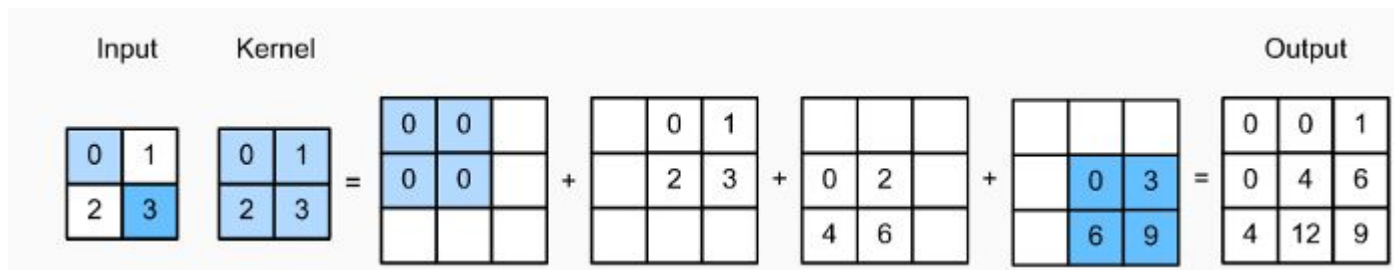
Μετασχηματισμένη
είσοδος



“Κλασική”
(με ανάκ

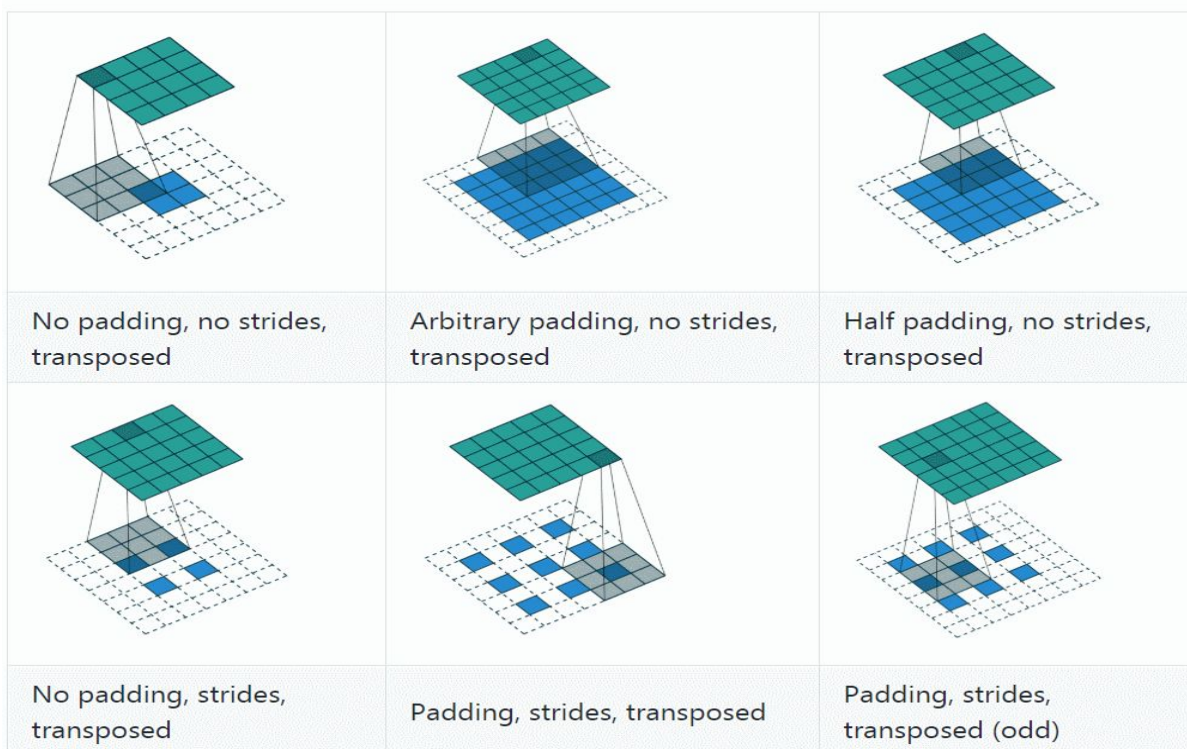
Στην **ανάστροφη συνέλιξη** το **stride** δεν
αναφέρεται σε βήματα μετάθεσης του
πυρήνα και το **padding** δεν αναφέρεται σε
προσθήκη στοιχείων “γύρω” από τα
δεδομένα!

Upscale



Ανάστροφη συνέλιξη (Transposed Convolution)

N.B.: Blue maps are inputs, and cyan maps are outputs.



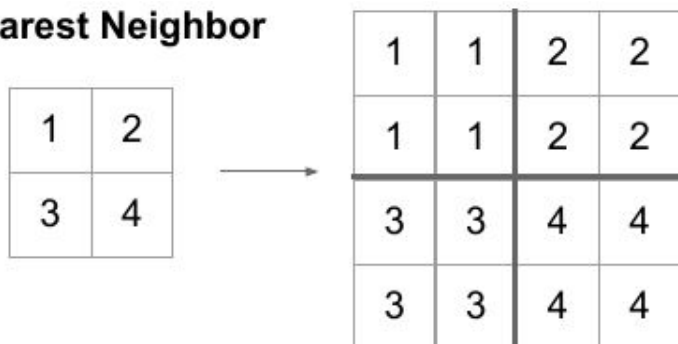
Comparison					
Conv Type	Operation	Zero Insertions	Padding	Stride	Output Size
Standard	Downsampling	0	p	s	$(i+2p-k)/s + 1$
Transposed	Upsampling	$(s - 1)$	$(k-p-1)$	1	$(i-1)*s+k-2p$

Upscale τεχνικές

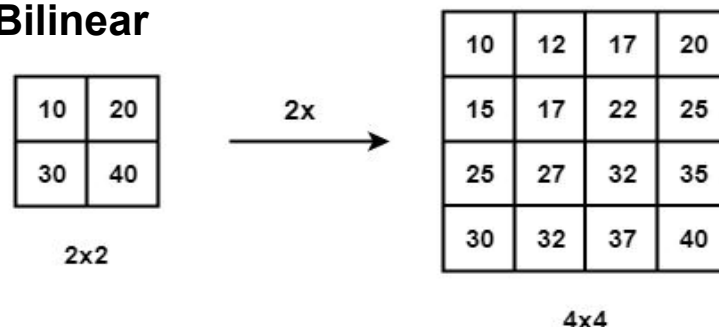
Μέθοδοι με παρεμβολή (interpolation):

1. **Εγγύτατου γείτονα** (Nearest Neighbor): Δίνεται η τιμή του εικονοστοιχείου της αρχικής εικόνας που βρίσκεται **πιο κοντά** έπειτα από τη “μεταφορά” του κάθε νέου εικονοστοιχείου στην αρχική εικόνα
2. **Διγραμμική παρεμβολή** (bilinear): Ως τιμή δίνεται ο **σταθμισμένος βάσει αποστάσεων μεσος όρος των 4 εγγυτέρων εικονοστοιχείων** της αρχικής εικόνας έπειτα από “μεταφορά” του εκάστοτε υπό εξέταση σημείου σε αυτήν.

Nearest Neighbor



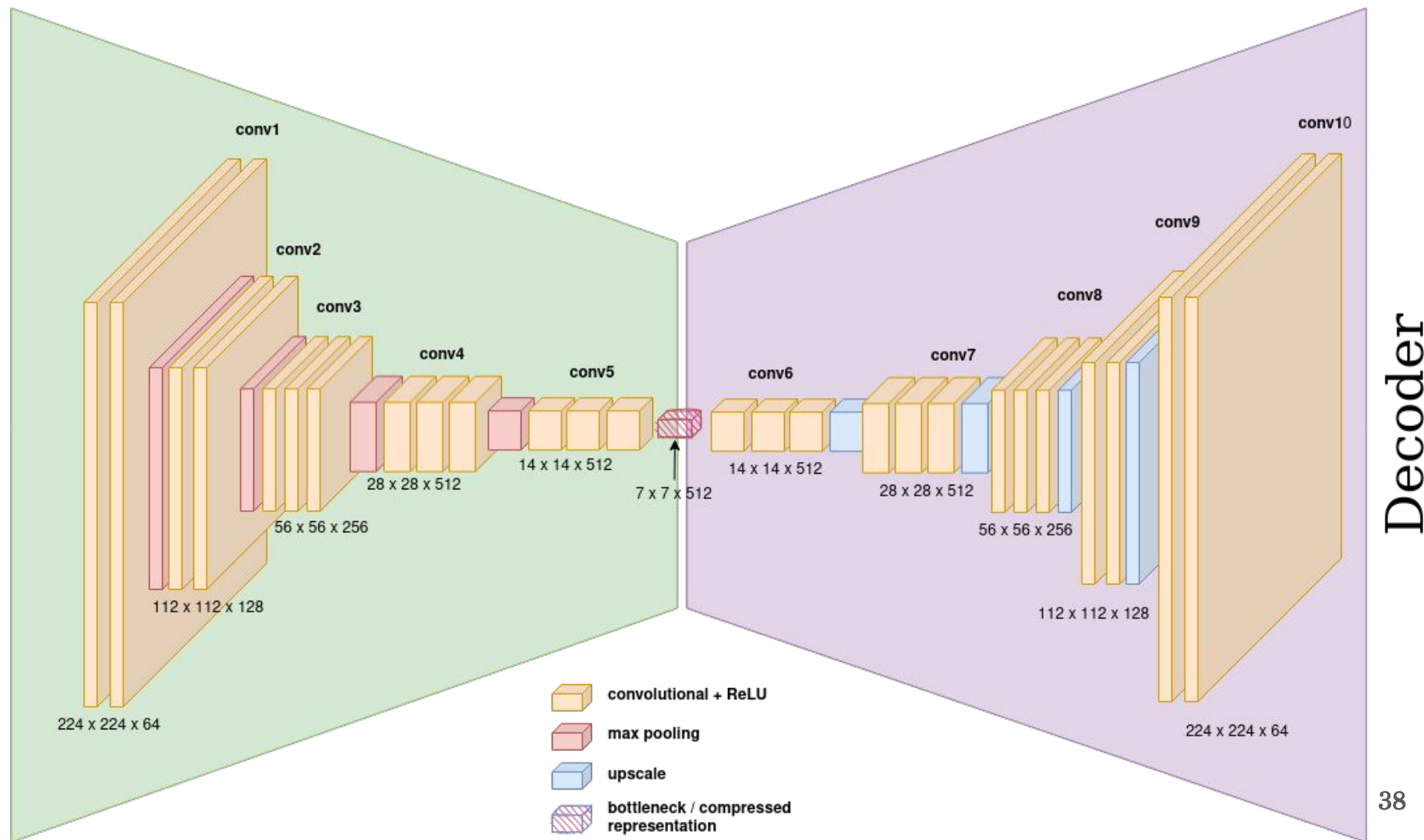
Bilinear



Autoencoders

Στόχος των AC είναι να δημιουργήσουν μία πυκνή αναπαράσταση η οποία θα εμπεριέχει όσο το δυνατόν περισσότερη πληροφορία “συμπιεσμένη” σε έναν τανυστή μικρότερου μεγέθους.

Encoder



Decoder

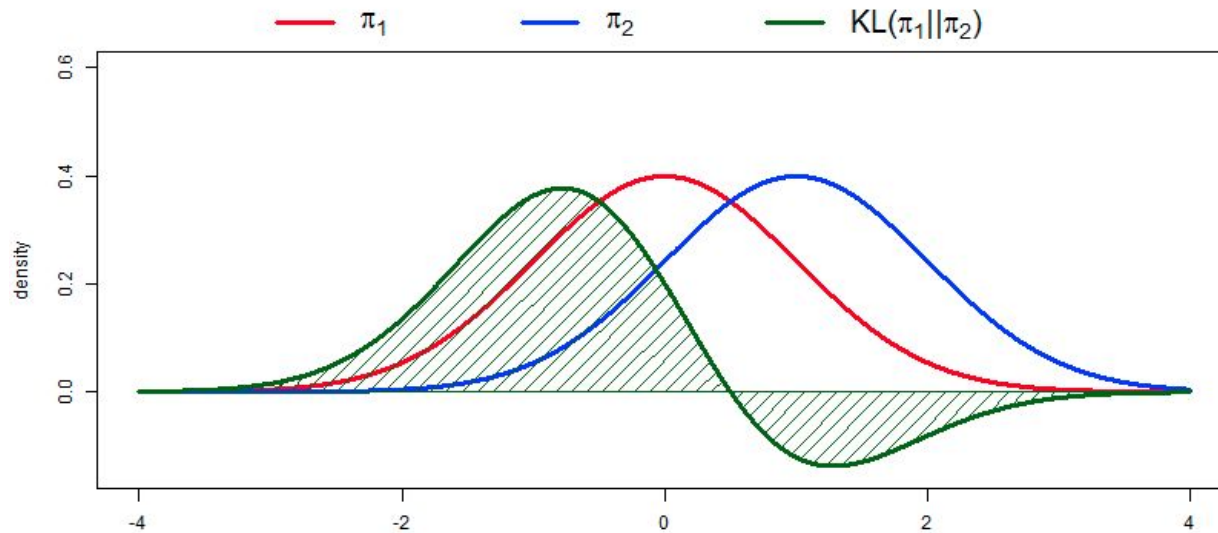
Autoencoders

Κατάλληλες συναρτήσεις κόστους:

- MSE loss
- L1 loss
- KL-divergence

Η **απόκλιση Kullback-Leibler** (KL-div)
“μετράει” τη διαφορά μεταξύ δύο κατανομών!

$$\mathcal{D}_{KL}(\mathcal{P} \parallel \mathcal{Q}) = \sum_i (\mathcal{P}_i \log \mathcal{P}_i - \mathcal{P}_i \log \mathcal{Q}_i)$$

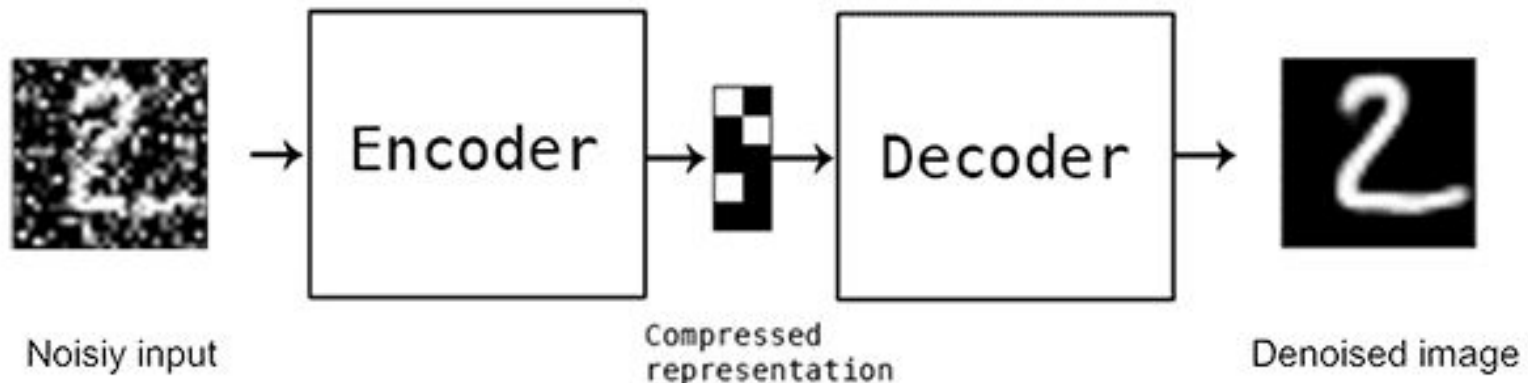


Autoencoders

Τύποι από autoencoders:

- **Denoising**
- Variational (VAE)
- Masked (MAE)
- Sparse
- Contractive

Τα denoising ACs χρησιμεύουν για την εκμάθηση **πιο εύρωστων αναπαραστάσεων** οι οποίες δεν είναι ευαίσθητες στην ύπαρξη θορύβου στα δεδομένα



Σημασιολογική Κατάτμηση (Semantic Segmentation)

- Η διαδικασία ανάθεσης μιας κατηγορίας/κλάσης σε κάθε **εικονοστοιχείο** μιας εικόνας
- Διαφορετικό πρόβλημα από αυτό της ταξινόμησης!
 - Ταξινόμηση : μία κατηγορία σε ολόκληρη την εικόνα
- Τα πολλαπλά αντικείμενα (σημασιολογικά) που ανήκουν στην ίδια κατηγορία αντιμετωπίζονται ως μία ξεχωριστή οντότητα/κατηγορία



predict →

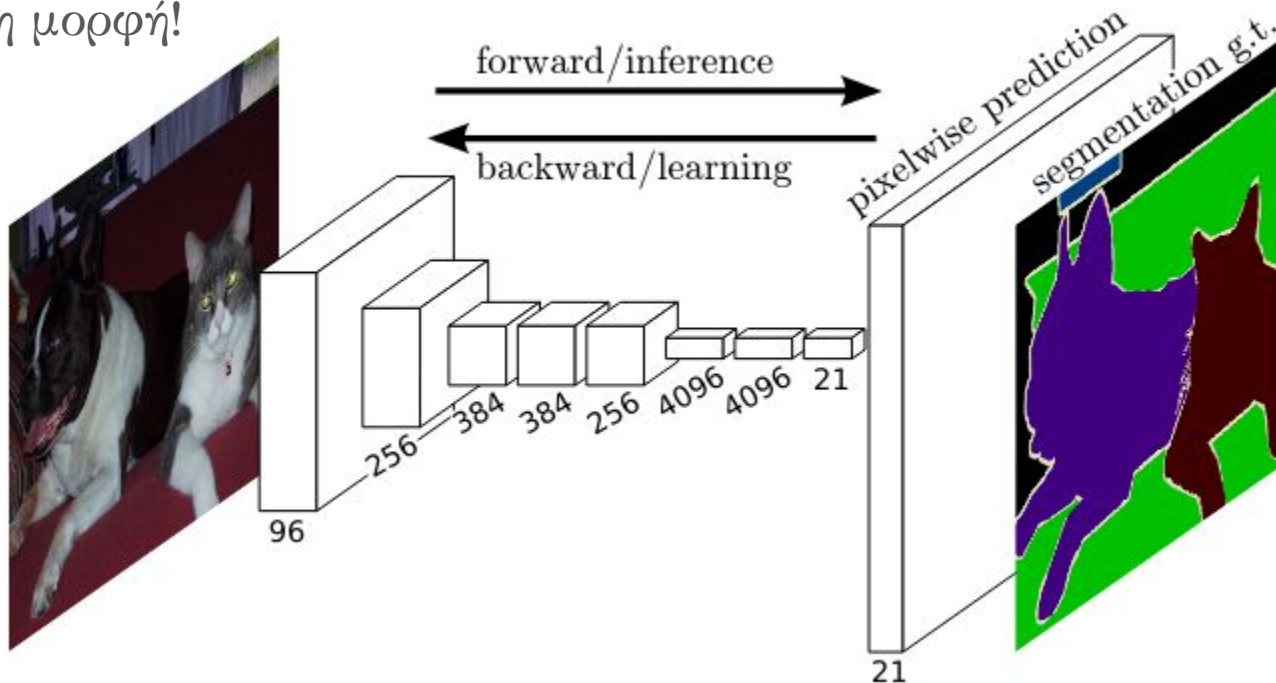


Person
Bicycle
Background

Σημασιολογική Κατάτμηση (Semantic Segmentation)

Encoder-Decoder FCN αρχιτεκτονικές μπορούν να χρησιμοποιηθούν και για τη σημασιολογική κατάτμηση εικόνων, καθώς η έξοδος για το συγκεκριμένο πρόβλημα είναι τανυστής 3ης τάξης (height x width x num_classes)!

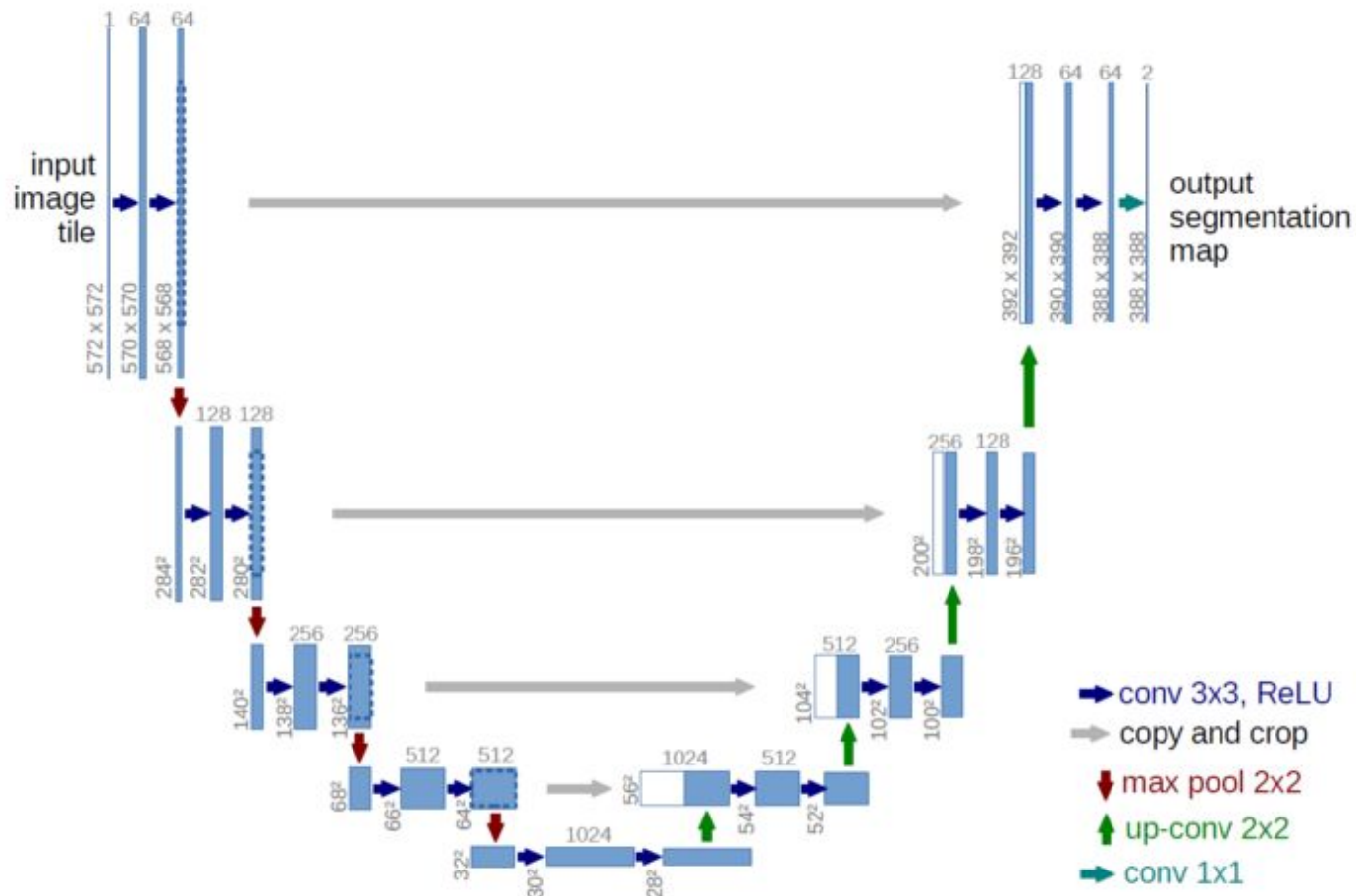
Τα απλά “σειριακά” FCN έχουν το πρόβλημα ότι τα χαρακτηριστικά κάθε κλίμακας ενσωματώνονται σειριακά με αποτέλεσμα να βρίσκονται σε πεπλεγμένη μορφή!

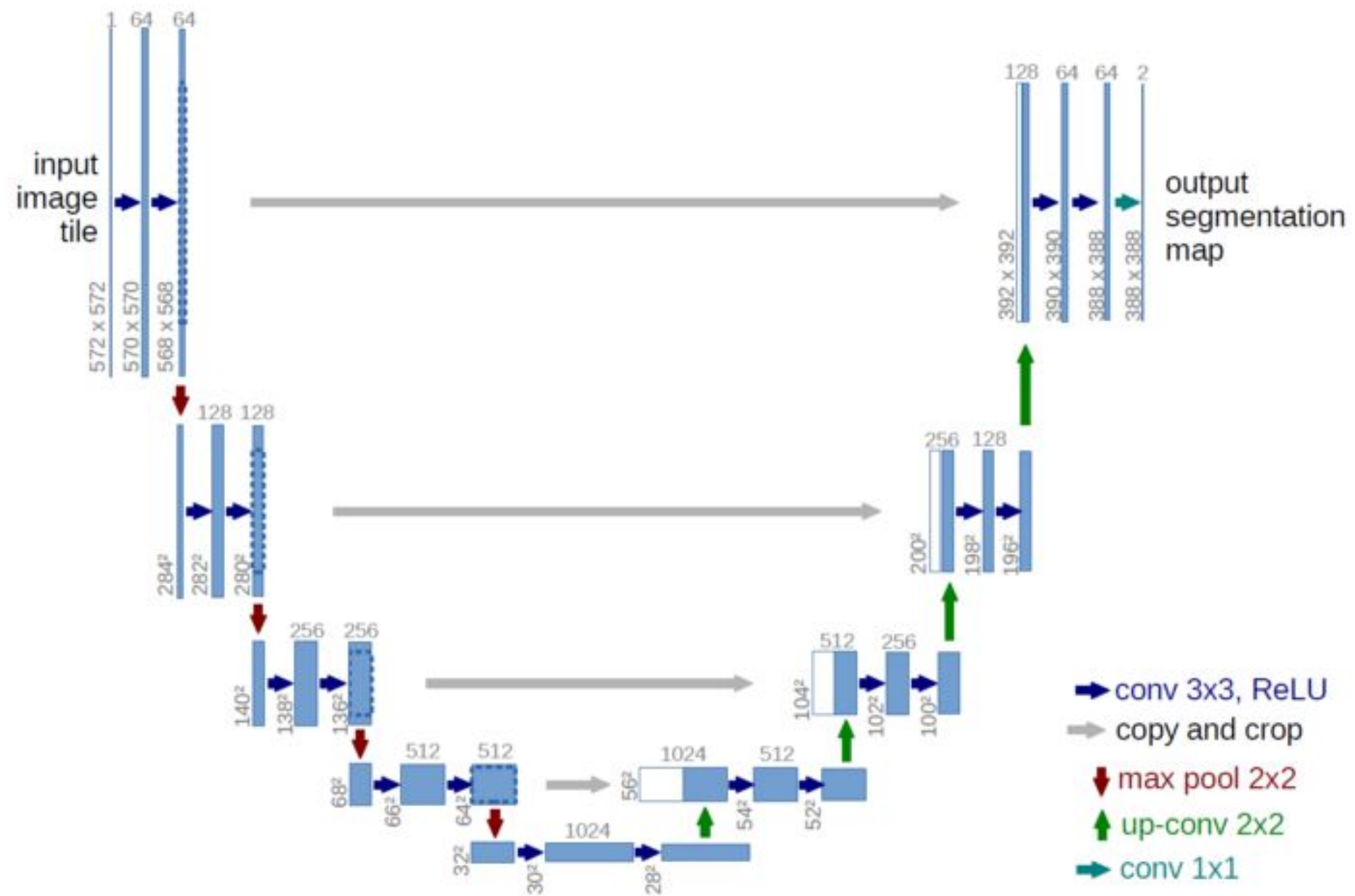


UNet αρχιτεκτονική

Το πρόβλημα της πιο άμεσης αναπαράστασης των κλιμακών καλείται να λύσουν τα δίκτυα τύπου U-Net.

- Μεταφορά με χρήση skip connections των χαρακτηριστικών προηγούμενων στρώσεων σε επόμενες





UNet αρχιτεκτονική

- Σε κάθε βήμα downsampling του decoder χρησιμοποιούμε **skip connections** συνενώνοντας τις εξόδους των στρώσεων upsampling (u) με τα **ομόλογα** feature maps του encoder στο ίδιο “επίπεδο (c), δηλαδή :

$$u6' = u6 + c4$$

$$u7' = u7 + c3$$

$$u8' = u8 + c2$$

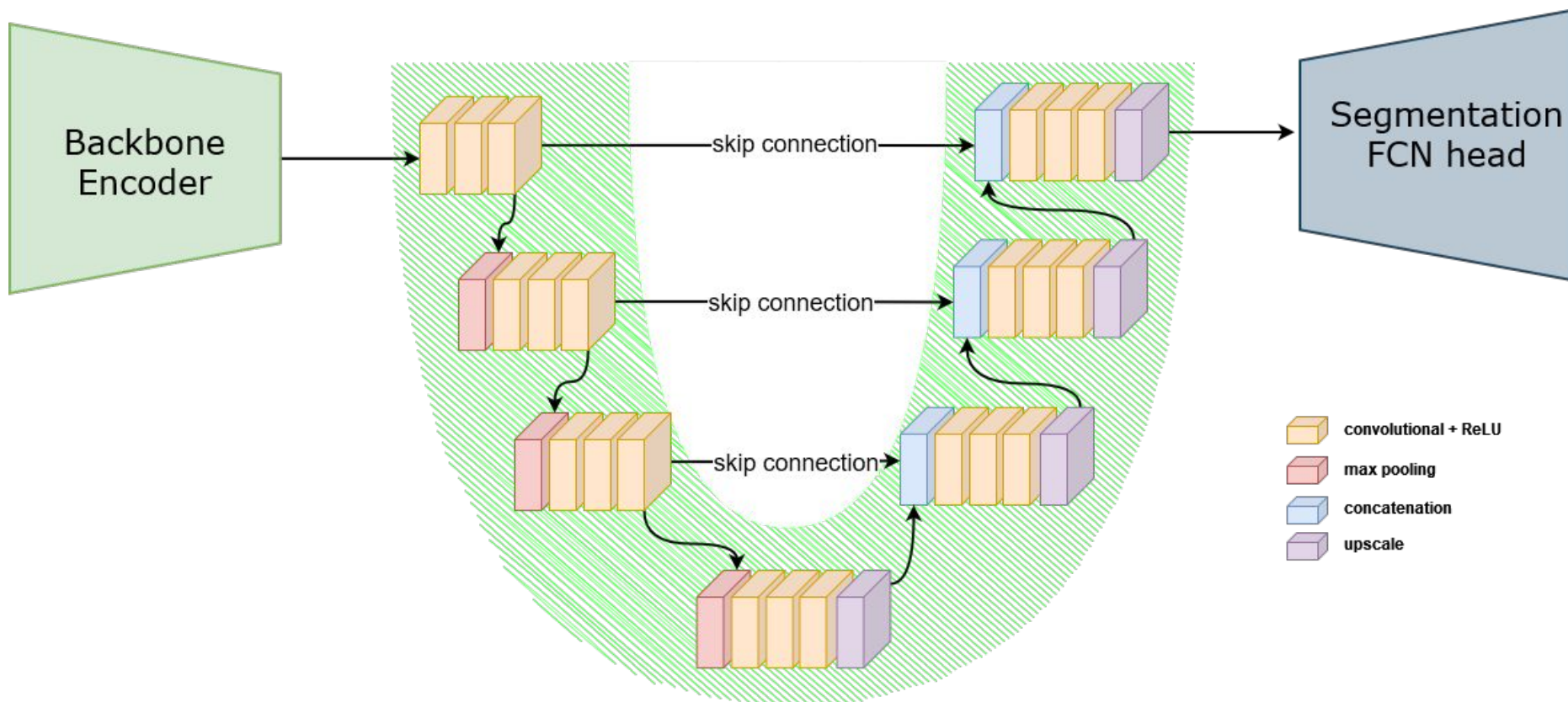
$$u9' = u9 + c1$$

Στη συνέχεια εφαρμόζονται διαδοχικά 2 συνελκτικές στρώσεις ώστε το μοντέλο να έχει τη δυνατότητα να μάθει μια “καλύτερη” αναπαράσταση

- Έτσι προκύπτει μια συμμετρική αρχιτεκτονική σχήματος U, προκύπτοντας έτσι το όνομα UNet από τους συγγραφείς

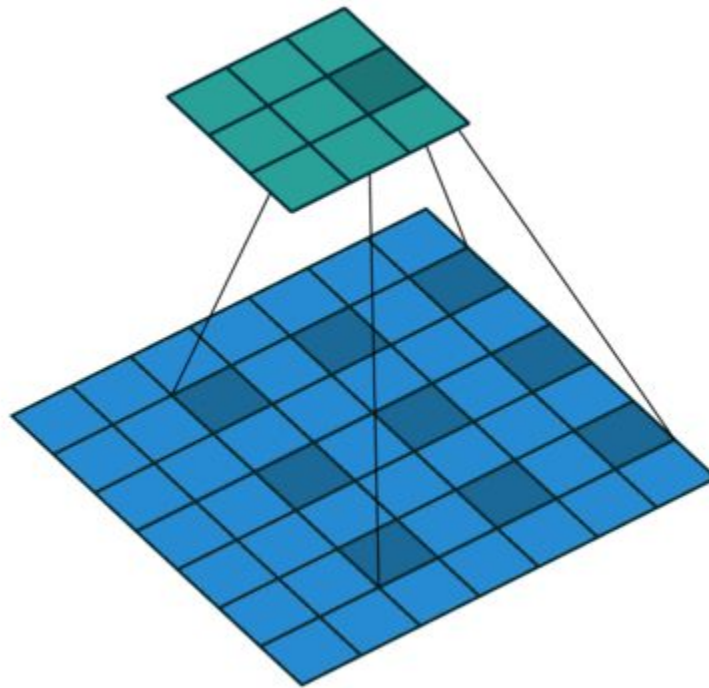
UNet αρχιτεκτονική

Γενικευμένη μορφή:



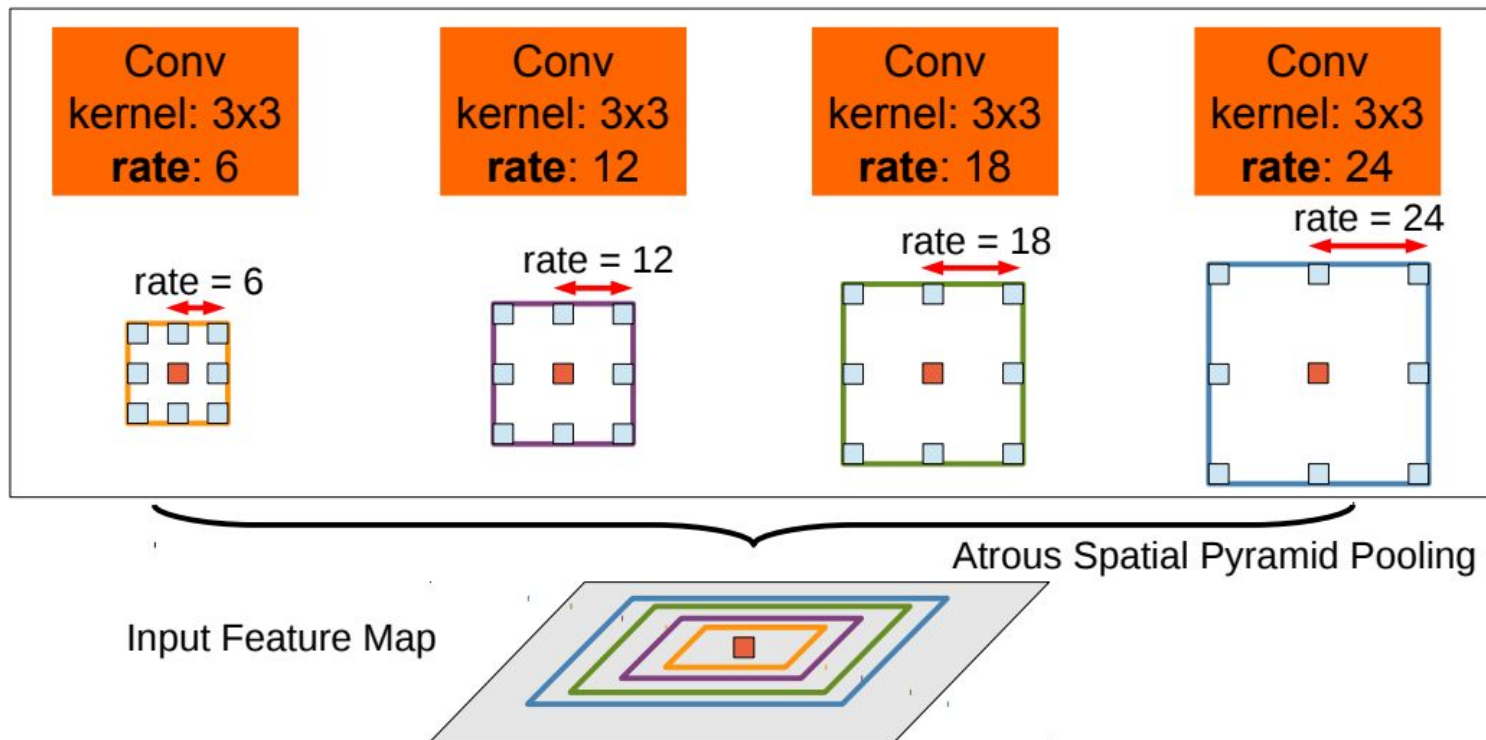
DeepLab αρχιτεκτονικές

- **Atrous convolutions** (deeplab v1)
- Spatial Pyramid pooling (deeplab v2)
- Depthwise separable atrous convolutions (deeplab v3)
- Specialized decoder (deeplab v3+)



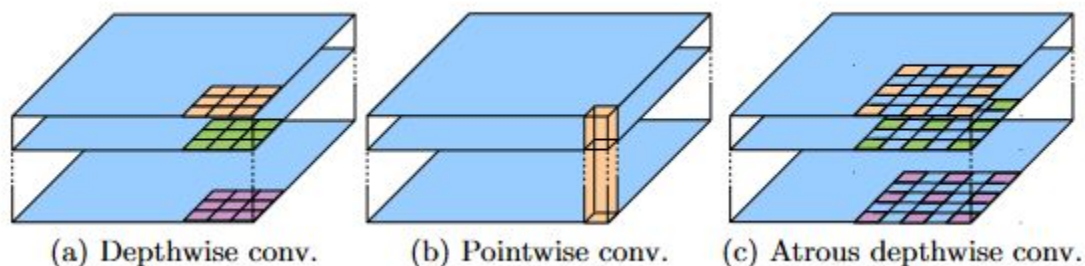
DeepLab αρχιτεκτονικές

- Atrous convolutions (deeplab v1)
- **Spatial Pyramid pooling** (deeplab v2)
- Depthwise separable atrous convolutions (deeplab v3)
- Specialized decoder (deeplab v3+)



DeepLab αρχιτεκτονικές

- Atrous convolutions (deeplab v1)
- Spatial Pyramid pooling (deeplab v2)
- **Depthwise separable atrous convolutions** (deeplab v3)
- Specialized decoder (deeplab v3+)



*DSCConv – parameters : $(kernelsize)^2 * channels_{in} + numfilters * channels_{in}$*

*Conv – parameters : $((kernelsize)^2 + 1) * channels_{in} * numfilters$*

Στις *depthwise separable convolutions* η απλή συνέλιξη αποσυντίθεται σε δύο σειριακά βήματα:

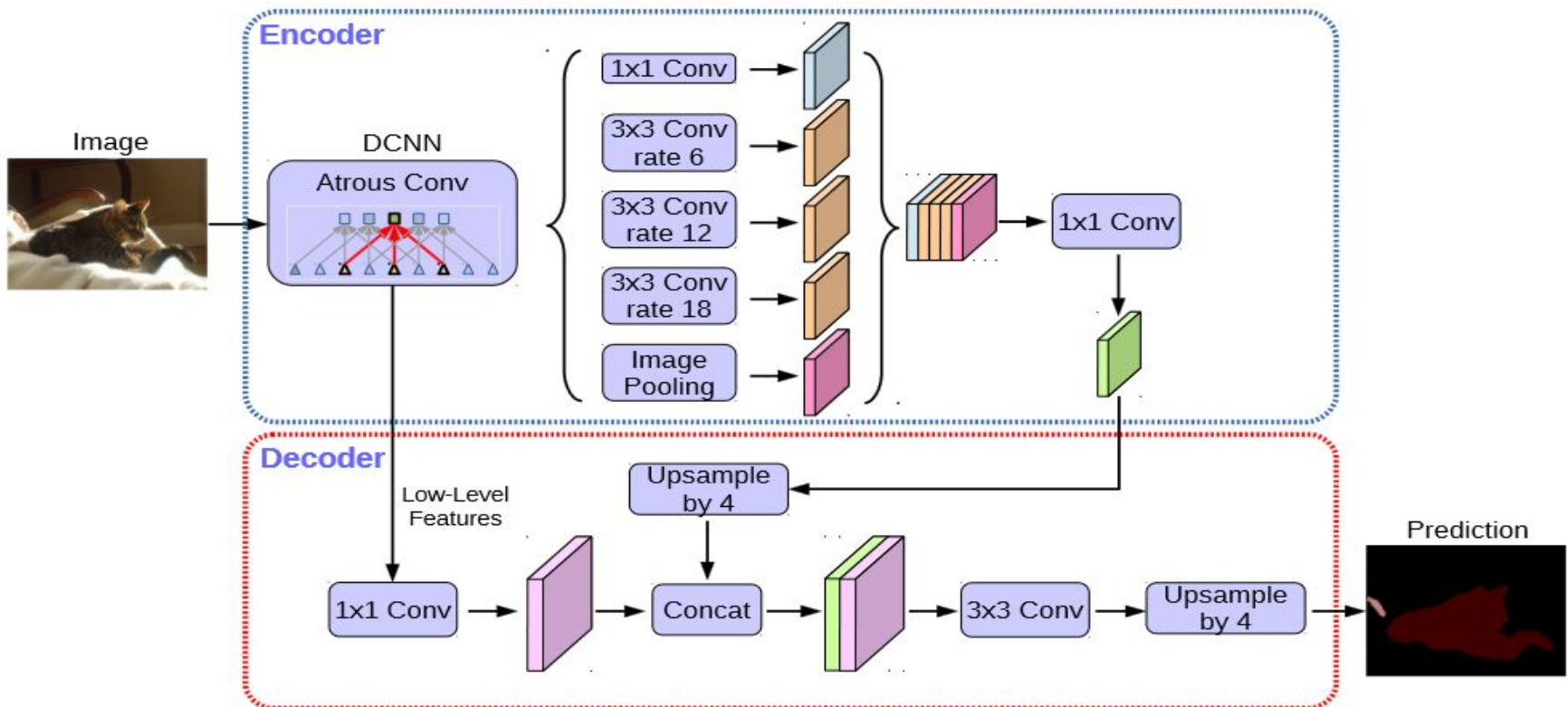
α) Συνέλιξη **ανά μεμονωμένο κανάλι**

β) **Συνέλιξη 1x1** για συνδυασμό των ανα κανάλι
συνελιγμένων χαρτών σε έναν τελικό χάρτη

Χωρίς
όρο bias

DeepLab αρχιτεκτονικές

- Atrous convolutions (deeplab v1)
- Spatial Pyramid pooling (deeplab v2)
- Depthwise separable atrous convolutions (deeplab v3)
- **Specialized decoder** (deeplab v3+)



Συντελεστής Dice & Soft-dice loss

Ο **συντελεστής Dice** μετράει την επικάλυψη μεταξύ δύο δειγμάτων / συνόλων

$$Dice = \frac{2|A \cap B|}{|A| + |B|}$$

για τη σημασιολογική κατάτμηση, θεωρώντας ένα κατώφλι T για τον τανυστή προβλέψεων \mathbf{B} , ο όρος $|A \cap B|$ αναφέρεται στο κατά σημείο γινόμενο μεταξύ των “δυαδικών” τελεστών \mathbf{A} , \mathbf{B} . Αντίστοιχα οι όροι $|A|$, $|B|$ αναφέρονται στο άθροισμα των στοιχείων των τανυστών.

Η συνάρτηση κόστους Dice loss προκύπτει ως $1-Dice$.

Στην πράξη, το βήμα κατωφλίωσης “βλάπτει” τη συνέχεια του Dice ως προς τον τανυστή προβλέψεων \mathbf{B} , συνεπώς και προς τα βάρη του ίδιου του δικτύου. Παραλείποντας την πράξη κατωφλίωσης προκύπτει η **soft Dice loss**

Συντελεστής Dice & Soft-dice loss

$$|A \cap B| = \begin{bmatrix} 0.01 & 0.03 & 0.02 & 0.02 \\ 0.05 & 0.12 & 0.09 & 0.07 \\ 0.89 & 0.85 & 0.88 & 0.91 \\ 0.99 & 0.97 & 0.95 & 0.97 \end{bmatrix} * \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{\text{element-wise multiply}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.89 & 0.85 & 0.88 & 0.91 \\ 0.99 & 0.97 & 0.95 & 0.97 \end{bmatrix} \xrightarrow{\text{sum}} 7.41$$

prediction target

$$|A| = \begin{bmatrix} 0.01 & 0.03 & 0.02 & 0.02 \\ 0.05 & 0.12 & 0.09 & 0.07 \\ 0.89 & 0.85 & 0.88 & 0.91 \\ 0.99 & 0.97 & 0.95 & 0.97 \end{bmatrix} \xrightarrow{\text{sum}} 7.82 \quad |B| = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{\text{sum}} 8$$

$^{2 \text{ (optional)}}$ $^{2 \text{ (optional)}}$

$$\text{soft Dice loss} = 1 - \frac{2 \sum_{\text{pixels}} y_{\text{true}} y_{\text{pred}}}{\sum_{\text{pixels}} y_{\text{true}}^2 + \sum_{\text{pixels}} y_{\text{pred}}^2}$$

Για τη σημασιολογική κατάτμηση η soft Dice loss υπολογίζεται ανά κατηγορία και στη συνέχεια λαμβάνεται ο (σταθμισμένος?) μέσος όρος

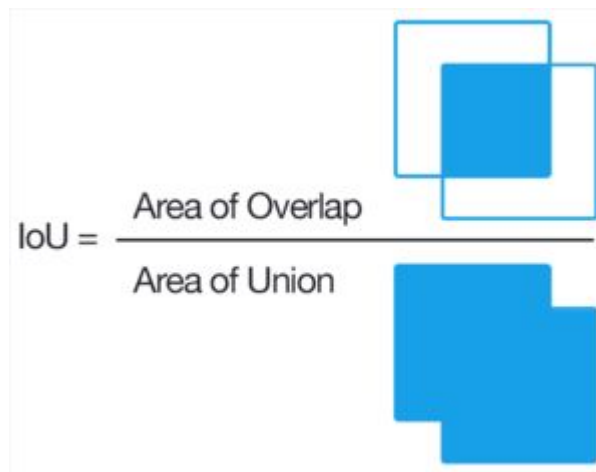
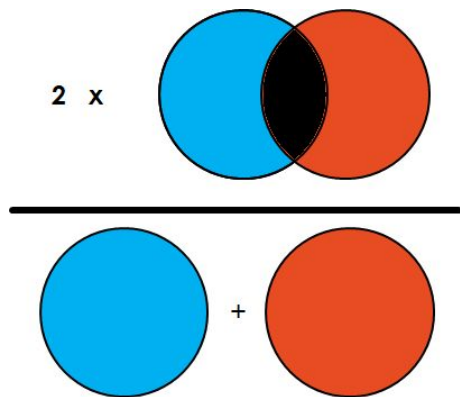
Σημασιολογική Κατάτμηση (Semantic Segmentation)

Συναρτήσεις κόστους:

- soft Dice loss
- Binary/Categorical cross entropy ανά εικονοστοιχείο

Μετρικές αξιολόγησης:

- Pixel accuracy (% των σωστά ταξινομημένων εικονοστοιχείων)
- Dice coefficient (~ F1 score)
- Jaccard-index / IoU metric





RSLab

Remote Sensing Laboratory
National Technical University of Athens



Διαχείριση και Επεξεργασία Μεγάλων Δεδομένων Παρατήρησης Γης



GitHub

<https://github.com/rslab-ntua>