



10 Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

## Μάθηση μεταφοράς (Transfer Learning) με συνελκτικά νευρωνικά δίκτυα (1D–CNN και 1D–Resnet) σε προηγμένη φασματοσκοπία Raman για εξατομικευμένη διάγνωση ορθοκολικού καρκίνου

Δημήτριος Καλατζής<sup>1</sup>, Ελλάς Σπυράτου<sup>1,2</sup>, Μαρία Καρναχωρίτη<sup>2,3</sup>, Μαρία Α. Κουρή<sup>1,2,4</sup>, Σπύρος Ορφανουδάκης<sup>5</sup>, Νεκτάριος Κουφόπουλος<sup>6</sup>, Αβραάμ Πουλιάκης<sup>6</sup>, Νικόλαος Δανιάς<sup>7</sup>, Ιωάννης Σεϊμένης<sup>8</sup>, Αθανάσιος Γ. Κόντος<sup>3</sup>, Ευστάθιος Π. Ευσταθόπουλος<sup>1,\*</sup>

<sup>1</sup> Β Εργαστήριο Ακτινολογίας, Ιατρική Σχολή, Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών, Ελλάδα

<sup>2</sup> Τμήμα Ιατρικής, Δημοκρίτειο Πανεπιστήμιο Θράκης, Αλεξανδρούπολη, Ελλάδα

<sup>3</sup> Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών, Εθνικό Μετσόβιο Πολυτεχνείο, Ελλάδα

<sup>4</sup> Πρόγραμμα Ιατρικής Φυσικής, Πανεπιστήμιο Μασαχουσέτης Λόουελ, Ηνωμένες Πολιτείες

<sup>5</sup> Αιρha ΠΛΗΡΟΦΟΡΙΚΗ Α.Ε., Αλεξανδρούπολη, Ελλάδα

<sup>6</sup> Β Παθολογική Κλινική, Ιατρική Σχολή, Πανεπιστημιακό Νοσοκομείο Αττικών, Ελλάδα

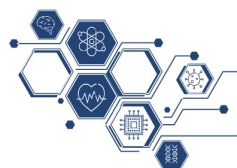
<sup>7</sup> Δ Χειρουργική Κλινική, Ιατρική Σχολή, Πανεπιστημιακό Νοσοκομείο Αττικών, Ελλάδα

<sup>8</sup> Ιατρική Σχολή, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Ελλάδα

Διοργανωτές



# 1. Εισαγωγή-Σκοπός



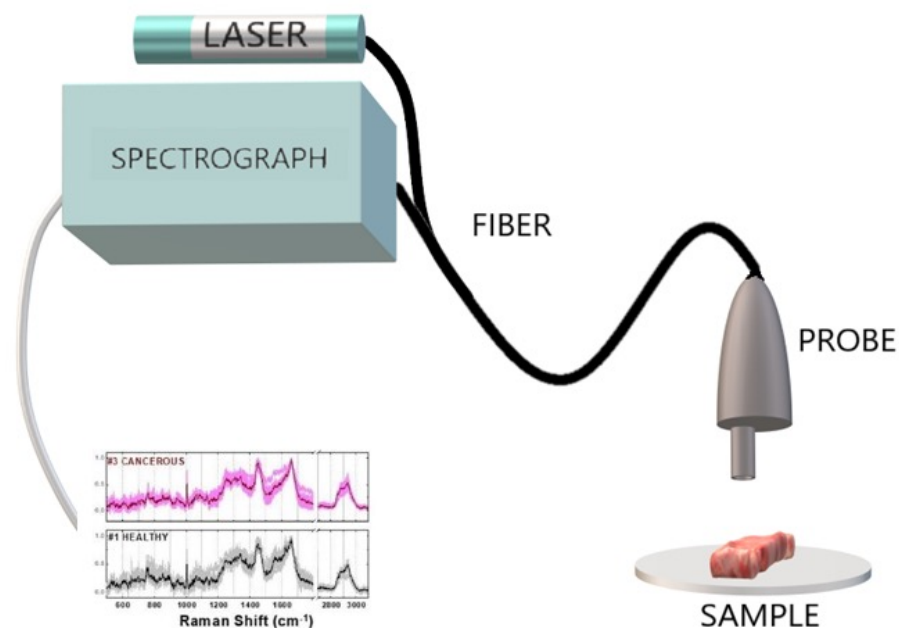
1ο Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

Η φασματοσκοπία Raman (Raman Spectroscopy, RS) είναι μια βιοφωτονική τεχνική με υψηλή ευαισθησία. Οποιαδήποτε βιοχημική αλλαγή στη σύσταση του ιστού αντικατοπτρίζεται στο φασματικό δακτυλικό του αποτύπωμα [1,2].

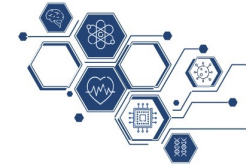
Η χρήση της τεχνητής νοημοσύνης (AI) για την ανάλυση φασματικών δεδομένων καθώς και η εξέλιξη των αισθητήρων επέτρεψαν την πρόσβαση της RS στην κλινική πράξη [3].

Περιορισμοί για την ανάπτυξη μοντέλων βαθιάς μάθησης:

- Απαιτούν μεγάλης κλίμακας βάση δεδομένων [4]
- Συλλογή μεγάλου πλήθους κλινικών δεδομένων [5]

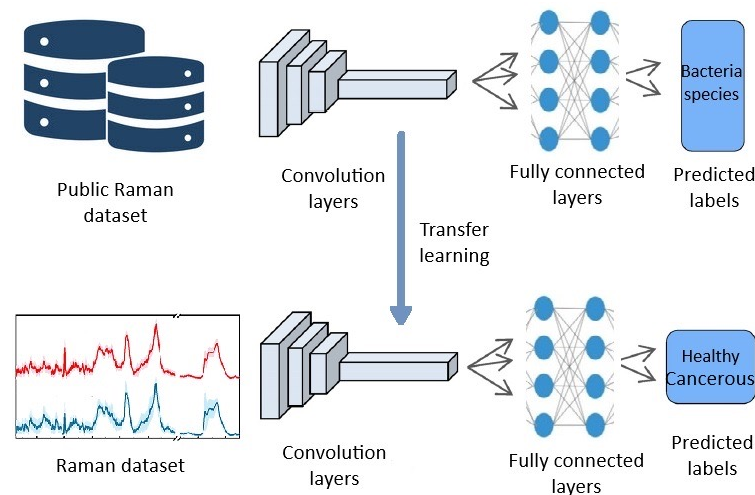


# 1. Εισαγωγή-Σκοπός

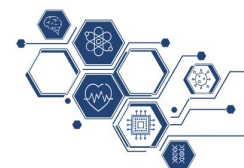


Στη παρούσα μελέτη:

- Χρησιμοποιήθηκε ανοικτή βάση δεδομένων Raman [6] για τη δημιουργία των προεκπαιδευμένων μοντέλων [7]
- Η γνώση που αποκτήθηκε εφαρμόστηκε στα δεδομένα Raman από τον καρκίνο του παχέος εντέρου για ταξινόμηση (υγιής/καρκινικός ιστός)
- Συγκρίθηκαν τα δύο μοντέλα, 1D-CNN και 1D-ResNet, με και χωρίς μεταφορά μάθησης (transfer learning, TL)



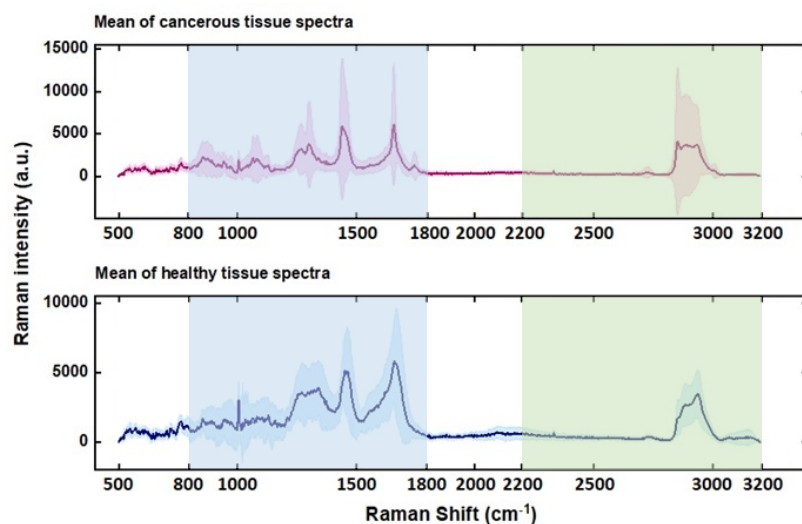
## 2. Μέθοδοι και Υλικά



1ο Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

Οι μετρήσεις Micro-Raman διεξήχθησαν με φασματόμετρο Renishaw Invia με λέιζερ 785 nm. Τα φάσματα καταγράφηκαν στο εύρος 500 έως 3300  $\text{cm}^{-1}$ .

Για τη μελέτη αυτή χρησιμοποιήθηκαν συνολικά 248 δείγματα 12 ασθενών/σετ με 134 δείγματα από υγιή ιστό και 114 από καρκινικό.

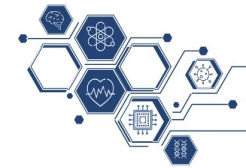


Προεπεξεργασία των δεδομένων:

- Διόρθωση φόντου χρησιμοποιώντας sensitive non-linear iterative peak-clipping (SNIP)
- Κανονικοποίηση min-max
- Η περιοχή των 800–1800  $\text{cm}^{-1}$  επιλέχθηκε ως είσοδος στα μοντέλα βαθιάς μάθησης

ασθενείς/σετ	Υγιής	Καρκινικός	Στάδιο	Βαθμός
set1	3	5	γpT3N2a	G2
set2	18	10	pT1N0	G1
set3	8	10	pT4bN1	G2
set4	13	11	pT3N0	G2
set5	9	10	pT3pN0	G1
set6	7	5	pT3N0	G1
set7	13	10	pT3N1M1	G2
set8	13	10	pT3N1c	G2
set9	10	9	pT2N0	G2
set10	14	10	pT3N0	G2
set11	16	15	pT2N0	G2
set12	10	9	pT3N0Mx	G2
<b>Σύνολο</b>	<b>134</b>	<b>114</b>		

## 2. Μέθοδοι και Υλικά



1ο Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

Χρησιμοποιήθηκε leave-p-out cross-validation για να αξιολογηθεί η απόδοση των μοντέλων. Η διαδικασία πραγματοποιήθηκε 6 φορές ενώ σε κάθε επανάληψη το p ήταν 2 ασθενείς για το τεστ και το n-p ήταν τα υπόλοιπα δεδομένα για την εκπαίδευση.

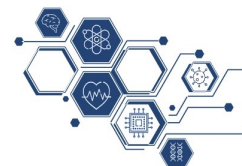


3 βήματα για την ανάπτυξη των μοντέλων ταξινόμησης.

- Τα μοντέλα εκπαιδεύτηκαν χρησιμοποιώντας το δημόσιο σύνολο δεδομένων Raman
- Αποθηκεύτηκαν τα στρώματα με τα εξαγόμενα χαρακτηριστικά
- Πάγωμα και εκπαίδευση (freeze then train) με το σύνολο των κλινικών δεδομένων

Τέλος, έγινε σύγκριση μεταξύ δύο μοντέλων (1D-CNN και 1D-ResNet) με και χωρίς μεταφορά μάθησης στα 248 φάσματα προκειμένου να αξιολογηθεί η ταξινόμηση (υγιής ή καρκινικός ιστός).

### 3. Αποτελέσματα



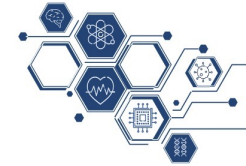
1ο Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

Με βάση τα εξαγόμενα αποτελέσματα από την παρακάτω διαδικασία επιλέχθηκε μόνο το εύρος 800–1800  $\text{cm}^{-1}$  για τη δημιουργία δεδομένων εκπαίδευσης.

Αρχικά, δημιουργήθηκαν δύο διαφορετικά σύνολα δεδομένων με εύρος 800–1800  $\text{cm}^{-1}$  και 2200–3200  $\text{cm}^{-1}$ . Στη συνέχεια, κάθε σύνολο δεδομένων χρησιμοποιήθηκε για την εκπαίδευση του 1D-CNN και του 1D-ResNet προκειμένου να αξιολογηθεί ποιο από τα δύο σύνολα δεδομένων δίνει τα καλύτερα αποτελέσματα.

Model	Wavenumbers ( $\text{cm}^{-1}$ )	Accuracy	Recall	Precision	f1_Score
1D-CNN	800–1800	0.834	0.859	0.796	<u>0.827</u>
1D-CNN	2200–3200	0.762	0.754	0.735	0.744
1D-ResNet	800–1800	0.850	0.859	0.823	<u>0.841</u>
1D-ResNet	2200–3200	0.814	0.850	0.769	0.808

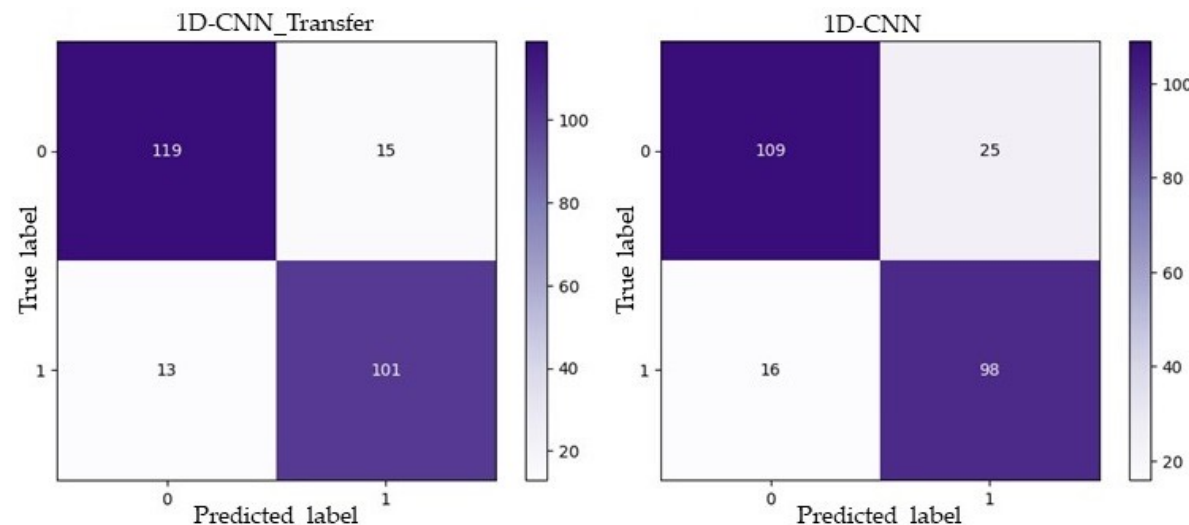
### 3. Αποτελέσματα



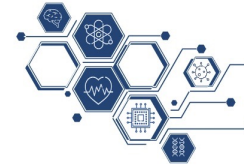
Στην τελική φάση των πειραμάτων, δύο μοντέλα βαθιάς μάθησης (1D-CNN και 1D-ResNet) με και χωρίς μάθηση μεταφοράς εκπαιδεύτηκαν με δεδομένα στην περιοχή  $800\text{--}1800\text{ cm}^{-1}$ .

- Η ακρίβεια του 1D-CNN με μάθηση μεταφοράς ήταν 5,3% υψηλότερη από αυτή του 1D-CNN χωρίς μάθηση μεταφοράς

Model	Accuracy	Recall	Precision	f1_Score
1D-CNN	0.834	0.859	0.796	0.827
1D-CNN transfer	0.887	0.885	0.870	<u>0.878</u>
1D-ResNet	0.850	0.859	0.823	0.841
1D-ResNet transfer	0.870	0.833	0.879	0.855



## 4. Συμπεράσματα



1ο Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

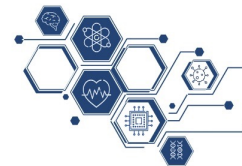
Τα μοντέλα μεταφοράς μάθησης 1D-CNN και 1D-ResNet επιτυγχάνουν:

- 5,3% και 2% υψηλότερη ακρίβεια, αντίστοιχα, σε αντίθεση με τα μοντέλα χωρίς μεταφοράς μάθησης
- Η μέθοδος μεταφοράς μάθησης βοήθησε να λυθούν οι σοβαροί περιορισμοί και οι δυσκολίες που αντιμετωπίστηκαν στη συλλογή μεγάλου όγκου κλινικών δεδομένων
- Υψηλές τιμές στην ευαισθησία που έφτασαν το 88,5%, αποδεικνύουν ότι το μοντέλο εντοπίζει σωστά την πλειοψηφία των καρκινικών ιστών, ελαχιστοποιώντας τις ψευδώς αρνητικές προβλέψεις

Συμπερασματικά, τα μοντέλα μπορούν να αποτελέσουν χρήσιμο εργαλείο στα χέρια των χειρουργών, βοηθώντας τους να προσδιορίσουν σε πραγματικό χρόνο, κατά τη διάρκεια της επέμβασης, τα όρια του καρκινικού όγκου.



## 5. Βιβλιογραφία



1ο Πανελλήνιο Συνέδριο Φυσικών Επιστημών στην Υγεία:  
Καινοτομίες και Προοπτικές  
22-23 Σεπτεμβρίου 2023

1. Krafft, C.; Schie, I.W.; Meyer, T.; Schmitt, M.; Popp, J. Developments in spontaneous and coherent Raman scattering microscopic imaging for biomedical applications. *Chem. Soc. Rev.* 2016, 45, 1819–1849.
2. Santos, I.P.; Caspers, P.J.; Schut, T.B.; et al. Implementation of a novel low-noise InGaAs detector enabling rapid near-infrared multichannel Raman spectroscopy of pigmented biological samples. *J. Raman Spectrosc.* 2015, 46, 652–660.
3. Krishna, H.; Majumder, S.K.; Chaturvedi, P.; Sidramesh, M.; Gupta, P.K. In vivo Raman spectroscopy for detection of oral neoplasia: A pilot clinical study. *J. Biophotonics* 2014, 7, 690–702.
4. Ngiam, K.Y. et al. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019, 20, 262–273.
5. Jain, S.H. et al. Is big data the new frontier for academic-industry collaboration? *JAMA* 2014, 311, 2171–2172.
6. Ho, C.-S.; Jean, N.; Hogan, C.A.; Blackmon, L.; Jeffrey, S.S.; et al. Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning. *Nat. Commun.* 2019, 10, 4927.
7. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* 2016, 3, 9.