ePublishing
ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ
NATIONAL DOCUMENTATION CENTRE

Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση

Τόμ. 1 (2021)

12ο Πανελλήνιο και Διεθνές Συνέδριο «Οι ΤΠΕ στην Εκπαίδευση»

Ανάπτυξη συστήματος αναγνώρισης συναισθημάτων για την εκπαίδευση

Κωνσταντίνος Παπαδόπουλος, Ιωάννης Λεύκος, Άννα-Μαρία Βελέντζα, Νικόλαος Φαχαντίδης

https://epublishing.ekt.gr | **e-Εκδότης:** ΕΚΤ | **Πρόσβαση:** 24/04/2024 12:12:20

Ανάπτυξη συστήματος αναγνώρισης συναισθημάτων για την εκπαίδευση

Κωνσταντίνος Παπαδόπουλος[1], Ιωάννης Λεύκος[2], Άννα-Μαρία Βελέντζα[3], Νικόλαος Φαχαντίδης[4]
kostaspap@uom.edu.gr, lefkos@uom.edu.gr, annamariavel@uom.edu.gr, nfachantidis@uom.edu.gr
[1,2,3,4] Τμήμα Εκπαιδευτικής και Κοινωνικής Πολιτικής, Πανεπιστήμιο Μακεδονίας, Θεσσαλονίκη

Περίληψη

Η αναγνώριση των συναισθημάτων των μαθητών είναι ουσιαστικό στοιχείο για μια επιτυχημένη εκπαιδευτική διαδικασία και ιδιαίτερα σε συνθήκες εξ αποστάσεως εκπαίδευσης είναι πολύ δύσκολο να το επιτύχει ένας εκπαιδευτικός. Η τεχνολογία αναγνώρισης συναισθημάτων μέσα από εικόνα ή βίντεο, μπορεί να συμβάλλει προς την κατεύθυνση αυτή. Η εργασία αυτή αφορά την παρουσίαση και την μεθοδολογία εγκυροποίησης μιας πλατφόρμας, με την οποία παρέχεται η δυνατότητα δοκιμής διαφόρων αλγορίθμων αναγνώρισης συναισθημάτων και επιλογής των καλύτερων μοντέλων τεχνητής ευφυΐας. Η πλατφόρμα εγκυροποιήθηκε επιτυχώς, μετά από χρήση κατάλληλων σετ εικόνων από προηγούμενες έρευνες και σε επόμενο στάδιο θα χρησιμοποιηθεί σε πραγματικές συνθήκες εξ αποστάσεως εκπαίδευσης.

Λέξεις κλειδιά: Τεχνητή Νοημοσύνη, Μηχανική μάθηση, Αλγόριθμοι αναγνώρισης συναισθημάτων, Εκπαιδευτική τεχνολογία

Εισαγωγή

Οποιαδήποτε εκπαιδευτική διαδικασία, εκτός από τη γνωστική, περιλαμβάνει και τη συναισθηματική διάσταση. Η αναγνώριση και κατανόησή της, από τους εκπαιδευτικούς, μπορεί να συνεισφέρει στον καλύτερο σχεδιασμό και υλοποίηση παρεμβάσεων, προάγοντας έτσι τη μάθηση. Το συναίσθημα φαίνεται πως επιδρά σημαντικά στις νοητικές διεργασίες των ανθρώπων, όπως η αντίληψη, η προσοχή, η μάθηση, η μνήμη και η επίλυση προβλημάτων, κρίσιμες για κάθε εκπαιδευτική διαδικασία, αλλά και στην κωδικοποίηση και ανάκτηση των πληροφοριών (Tyng et al., 2017).

Τα θετικά συναισθήματα, φαίνεται πως διευκολύνουν τη μάθηση και αυτό δεν αποδίδεται μόνο στο ευχάριστο κλίμα και τη θετική προδιάθεση για συμμετοχή, αλλά και στην ικανοποίηση που λαμβάνουν οι μαθητές και η οποία δρα σαν εσωτερικό κίνητρο (Li et al., 2020). Δηλαδή, όταν μια μαθησιακή διαδικασία εξελίσσεται θετικά για έναν μαθητή, την απεικονίζει στο μυαλό του με θετική προοπτική και έτσι πράγματι συνεχίζεται με τον τρόπο αυτό. Επίσης, τα θετικά συναισθήματα ενισχύουν την αυτο-ρύθμιση και μαθησιακή αυτονομία των μαθητών, ενώ αντίθετα, οι μαθητές που βιώνουν αρνητικά συναισθήματα έχουν την τάση να βασίζονται στην καθοδήγηση κάποιου άλλου (π.χ. του εκπαιδευτικού) (Villavicencio & Bernardo, 2013). Παρόμοια αποτελέσματα αναφέρονται και από μελέτες της Θετικής Ψυχολογίας, όπου τα θετικά συναισθήματα των μαθητών, φαίνεται να ενισχύουν τις επιδόσεις τους (Carmona-Halty et al., 2019).

Επίσης σημαντική σημασία για τη μάθηση φαίνεται να έχουν και τα συναισθήματα έκπληξης, καθώς εμφανίζουν ισχυρή θετική συσχέτιση με την ενίσχυση κινήτρων και την ικανότητα αποθήκευσης πληροφοριών (Reisenzein, Horstmann & Schützwohl, 2019). Η έκπληξη μπορεί να είναι η παραβίαση της προσδοκίας των μαθητών, μια ασυμφωνία μεταξύ προηγούμενων προσδοκιών και νέων πληροφοριών που με τον τρόπο αυτό τραβάει την προσοχή και ερεθίζει τη περιέργειά τους, δημιουργώντας μια τάση για εξερεύνηση και διευρύνοντας την ικανότητα για μάθηση και απομνημόνευση πληροφοριών (Oudeyer et al., 2016).

Επομένως, θεωρείται πολύ σημαντική η αναγνώριση και παρακολούθηση των συναισθημάτων των μαθητών αλλά και των πιθανών αλλαγών τους κατά τη διάρκεια μιας μαθησιακής διαδικασίας. Προς την κατεύθυνση αυτή χρησιμοποιούνται διαφορετικές προσεγγίσεις, μεταξύ των οποίων η αναγνώριση των εκφράσεων του προσώπου (Jack and Schyns, 2015), των φωνητικών εκφράσεων (Russell et al., 2003), ή των αλλαγών στις χειρονομίες (Dael et al., 2012). Η χρήση προηγμένων νέων τεχνολογιών και μεθόδων όπως είναι οι αλγόριθμοι αναγνώρισης συναισθημάτων με τη βοήθεια κάμερας ή/και βιομετρικών μετρήσεων, βοηθάει στην μοντελοποίηση (Velentza, 2021) και κατά συνέπεια στον σχεδιασμό εκπαιδευτικών παρεμβάσεων. Η πρόσφατη εμπειρία της πανδημίας, με την αναγκαστική στροφή της εκπαίδευσης σε εξ αποστάσεως διαδικασίες, απέδειξε ότι ένα από τα μεγαλύτερα προβλήματα ίσως δεν αφορούσε στο γνωστικό περιεχόμενο, αλλά στη συναισθηματική επαφή και την κοινωνική αλληλεπίδραση μεταξύ εκπαιδευτικών και μαθητών ή μεταξύ των μαθητών (Lloyd-Jones, 2021). Σε μια δια ζώσης διδασκαλία, ο εκπαιδευτικός έχει από πρώτο χέρι την αίσθηση του κλίματος της τάξης, αλλά και των μεμονωμένων αντιδράσεων των μαθητών. Προς την κατεύθυνση αυτή φαίνεται ότι η τεχνολογία, μπορεί να προσφέρει εναλλακτικές λύσεις όπως η αναγνώριση των εκφράσεων του προσώπου (Lee & Lee, 2021), γεφυρώνοντας το παρατηρούμενο χάσμα, βοηθώντας τους εκπαιδευτικούς να αντιλαμβάνονται τα συναισθήματα των μαθητών. Οι εκπαιδευτικοί θα είχαν την ευκαιρία να ανα-προσαρμόζουν τις μαθησιακές δραστηριότητες ή το στυλ του μαθήματος (Tonguç & Ozaydın Ozkara, 2020) ή ακόμη και να κάνουν ένα μικρό διάλειμμα όταν χάνεται η προσοχή τους (Krithika & Lakshmi Priya, 2016), προκειμένου να αυξήσουν τη συμμετοχή των μαθητών τους και τα θετικά τους συναισθήματα και κατ' επέκταση να ενισχύσουν τη μάθηση.
Η αναγνώριση συναισθημάτων είναι μια τεχνολογία που έχει γίνει εφικτή τα τελευταία χρόνια λόγω μεγάλων αλμάτων της τεχνολογίας σε σχέση με το υλισμικό αλλά καθώς και της εκτεταμένης έρευνας που πραγματοποιείται τα τελευταία χρόνια στον τομέα αυτό. (Dosilovic, Brcic, & Hlupic, 2018) Ο κλάδος έχει μεγάλες προοπτικές σε ένα ευρύ φάσμα εφαρμογών και κυρίως μπορεί να αποτελέσει τη βάση για σημαντικές αλλαγές στην εκπαίδευση.

Η παρούσα εργασία αποτελεί μέρος μιας ευρύτερης έρευνας η οποία αφορά την ανάπτυξη μιας πλατφόρμας δοκιμών όπου θα μπορούμε να χρησιμοποιούμε διαφορετικούς αλγόριθμους μηχανικής μάθησης καθώς και διαφορετικές τεχνικές προ-επεξεργασίας εικόνας που ως αποτέλεσμα θα έχουν την επιτυχή αναγνώριση συναισθημάτων με πάρα πολύ υψηλή ακρίβεια. Συγκεκριμένα, παρουσιάζεται η διαδικασία εγκυροποίησης της πλατφόρμας με βάση αποτελέσματα προηγούμενων ερευνών και ο τρόπος επιλογής του καλύτερου "μοντέλου" μηχανικής μάθησης.

Αναγνώριση συναισθημάτων με χρήση τεχνικών αναγνώρισης προσώπου σε εικόνα

Η κατανόηση συναισθημάτων από υπολογιστικά συστήματα είναι ένα επίκαιρο, σύνθετο πρόβλημα που προσπαθούν να επιλύσουν συστήματα μηχανικής εκμάθησης. Η μελέτη των συναισθημάτων ξεκίνησε ήδη από το 1970 με τους Ekman και Friesen (1971), οι οποίοι έθεσαν τις αρχές στην εργασία τους ορίζοντας τα έξι πιο βασικά συναισθήματα, που αποτελούνται από: τον θυμό, τον φόβο, την χαρά, την λύπη, την έκπληξη και την αηδία. Στην ίδια εργασία περιγράφουν το σύστημα που χρησιμοποίησαν για να καταγράψουν τα συναισθήματα, που αποκαλείται Facial Action Coding System (FACS) και περιγράφει πώς τμήματα του προσώπου μπορούν να βαθμολογηθούν και μετέπειτα να αξιολογηθούν ώστε να εξαχθεί ένα συμπέρασμα για την συναισθηματική κατάσταση του ατόμου. Το σύστημα που επικράτησε τελικά για την βαθμονόμηση των συναισθημάτων είναι μια παραλλαγή του FACS την οποία δημιούργησαν οι Cohn, Ambadar και Ekman (2007) σε μεταγενέστερη εργασία τους, στην οποία αναθεώρησαν το σύστημα FACS, αφαιρώντας του στοιχεία και αναθεωρώντας τους κανόνες βαθμονόμησης του συστήματος.

Σήμερα τα περισσότερα συστήματα αναπτύσσονται στο πλαίσιο διεθνών διαγωνισμών που οργανώνονται από ακαδημαϊκούς ή επιχειρηματικούς φορείς. Σε αυτούς τους διαγωνισμούς δίνεται μια βάση δεδομένων που συνήθως περιλαμβάνει ένα σετ εικόνων που έχουν χαρακτηριστεί από ανθρώπους με χρήση μεταδεδομένων και αυτό ορίζεται σαν σετ εκπαίδευσης. Επίσης δίνεται ένα δεύτερο σετ εικόνων που είναι διαφορετικές από τις αρχικές αλλά είναι πάλι χαρακτηρισμένες και αποτελεί το σετ επιβεβαίωσής. Τα συστήματα αναπτύσσονται χρησιμοποιώντας το σετ εκπαίδευσης και στην συνέχεια χρησιμοποιείται το σετ επιβεβαίωσης για να εξάγουν συμπεράσματα με βάση τα ποιοτικά κριτήρια που έχουν θέσει στον αλγόριθμο τεχνητής νοημοσύνης. Στο τέλος του διαγωνισμού δίνεται συνήθως και ένα τρίτο σετ εικόνων που χρησιμοποιείται για την αξιολόγηση των συστημάτων που συμμετέχουν στον διαγωνισμό. Τα σετ αυτά γίνονται διαθέσιμα μετά το πέρας του διαγωνισμού σε οποιονδήποτε ενδιαφερόμενο ώστε να επιβεβαιώσει τα αποτελέσματα ή να δοκιμάσει νέες τεχνικές. Επιγραμματικά αναφέρονται παρακάτω μερικά από τα πιο δημοφιλή σετ:

- AffectNet
- FER2013
- Cohn-Kanade (CK+)
- FERPlus (FER+)
- SFEW
- RAF-DB

Συνήθως ένας αλγόριθμος δοκιμάζεται σε πολλαπλά σετ ώστε να αποφευχθούν οι περιπτώσεις που ο αλγόριθμος ταυτίζεται με τον θόρυβο των δεδομένων (overfitting) και δίνει λανθασμένα αποτελέσματα.

Το σετ FERPlus (FER+)

Το FER2013 αποτελείται από 28.558 φωτογραφίες με ανάλυση 48x48 πίξελ για το σετ εκπαίδευσης (train), 3.579 φωτογραφίες για το σετ επιβεβαίωσης (validation) καθώς και 3.573 για το σετ δοκιμών (test). Για κάθε σετ περιλαμβάνεται και ένα αρχείο μεταδεδομένων που περιέχει το επικρατέστερο συναίσθημα για κάθε φωτογραφία το οποίο όμως το έκρινε ένα μόνο άτομο. Το FER+ σετ αποτελείται από τα δεδομένα του FER2013 με έξτρα μεταδεδομένα, όπου κάθε εικόνα έχει ταξινομηθεί από δέκα διαφορετικά άτομα. Συγκεντρώνοντας και συλλέγοντας δεδομένα από πολλά άτομα, βελτιώνεται το ποσοστό ευστοχίας του ανθρώπινου παράγοντα καθώς μπορούμε να χρησιμοποιήσουμε τον μέσο όρο των απαντήσεων που έδωσε κάθε άτομο ή άλλες μορφές στατιστικής ανάλυσης. Όπως αναφέρουν χαρακτηριστικά στην εργασία τους οι Barsoum et al., (2016) τα ποσοστά ευστοχίας στο σετ FER2013 είναι περίπου 65 ± 5%, ενώ στο FER+ παρατήρησαν ότι αν χρησιμοποιούσαν τα δεδομένα τριών ατόμων η ευστοχία ήταν κοντά στο 46%, ενώ με πέντε ή επτά άτομα προσέγγιζε το 67% και 80% αντίστοιχα. Επίσης έχοντας δεδομένα από δέκα άτομα τους επέτρεψε να πειραματιστούν με διαφορετικές προσεγγίσεις όσον αφορά την στατιστική κατανομή της πιθανότητας να έχουν ανιχνεύσει σωστά οι άνθρωποι το συναίσθημα. Οι προσεγγίσεις που επέλεξαν να χρησιμοποιήσουν είναι οι παρακάτω:

- Majority Voting (MV)
- Multi-Label Learning (MLL)
- Probabilistic Label Drawing (PLD)
- Cross-entropy loss (CEL)

Η κάθε προσέγγιση έχει τα δικά της πλεονεκτήματα και χρησιμοποιείται για να επιτύχει διαφορετικούς σκοπούς. Η πλειοψηφία (Majority Voting) είναι η πιο καθιερωμένη έκδοση όπου ο αλγόριθμος δέχεται ως σωστό συναίσθημα αυτό που έχουν ψηφίσει τα περισσότερα άτομα για αυτή την φωτογραφία. Ενώ ένα σύστημα πολλαπλών ετικετών (Multi-Label) έρχεται να λύσει προβλήματα όπου περισσότερα από ένα συναισθήματα εκφράζονται στην ίδια εικόνα δίνοντας την πιθανότητα του κάθε συναισθήματος.

Ερευνητικά ερωτήματα

Τα ερωτήματα της έρευνας είναι:
Στην πλατφόρμα δοκιμών που θα αναπτύξουμε:

a. Θα επιβεβαιώνονται τα αποτελέσματα της δημοσίευσης του FER+ που επιλέξαμε ως επικρατέστερο αλγόριθμο;
b. Ποιο είναι το βέλτιστο "μοντέλο" με βάση το σετ επιβεβαίωσης ώστε να το χρησιμοποιήσουμε για την εξαγωγή αποτελεσμάτων σε πραγματικό χρόνο.

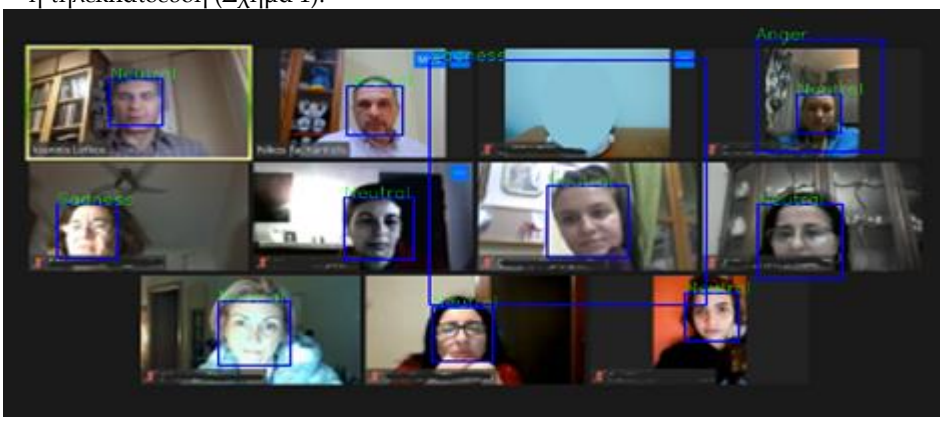
Μεθοδολογία

Για τις ανάγκες της παρούσας έρευνας, θεωρήθηκε ως καλύτερο σετ το FER+ καθώς περιέχει μεταδεδομένα από δέκα διαφορετικά άτομα ενώ επιπροσθέτως επιτρέπει να εφαρμοστούν διάφορες τεχνικές που δεν θα ήταν εφαρμόσιμες σε σετ όπως τα FER2013 ή το CK+. Επίσης τα τρία προαναφερόμενα σετ είναι τα πιο διαδεδομένα και αποδεκτά στην κοινότητα με την αρχική εργασία του FER+ να αποτελεί αναφορά πλήθους σχετικών εργασιών. Πρώτος στόχος είναι η επαλήθευση των δεδομένων των προηγούμενων ερευνών. Έτσι επιλέχθηκε η αρχική δημοσίευση του FER+ που είναι: Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution (Barsoum et al., 2016). Σε αυτή οι ερευνητές περιγράφουν την χρήση του FER+ σε συνδυασμό με ένα νευρωνικό δίκτυο VGG-13 (Simonyan & Zisserman, 2014) που αποτελείται από μια τροποποιημένη έκδοση του βασικού νευρωνικού δικτύου που αρίστευσε στον διαγωνισμό ILSVRC2014. Δεύτερος στόχος είναι να συγκρίνουμε τα αποτελέσματα των εξαγόμενων μοντέλων και αφού αυτά επαληθευτούν να επιλέξουμε το βέλτιστο μοντέλο που θα χρησιμοποιηθεί από την πλατφόρμα μας για την αναγνώριση συναισθημάτων σε πραγματικό χρόνο.

Η πλατφόρμα

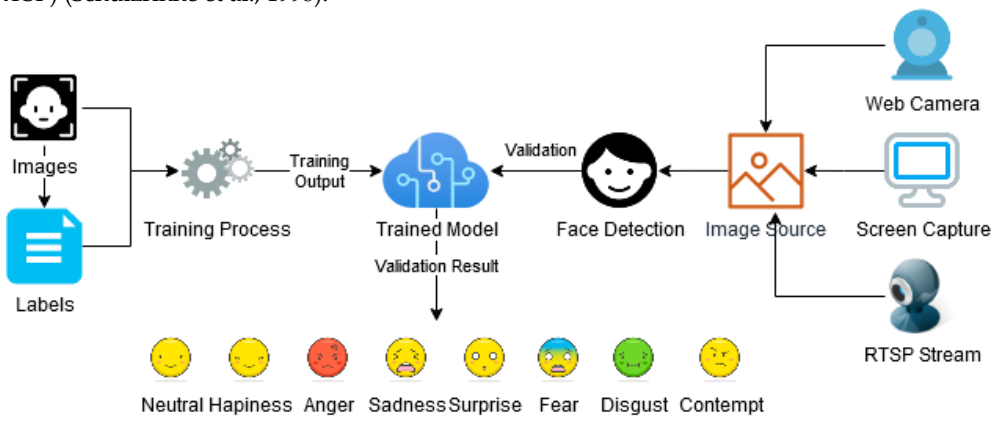
Η πλατφόρμα σχεδιάστηκε και υλοποιήθηκε με το σύστημα της τεχνητής νοημοσύνης, ώστε να καλύπτει τις παρακάτω προδιαγραφές:

- να μπορεί να δέχεται ως δεδομένο μια εικόνα ή μια συνεχόμενη ροή βίντεο και να την αξιολογεί σε σχεδόν πραγματικό χρόνο.
- να μπορεί να υποστηρίζει αναγνώριση πολλαπλών προσώπων ώστε να μπορεί να προσαρμοστεί σε ένα περιβάλλον όπου παρευρίσκονται πολλά άτομα, π.χ. Σχολική τάξη ή τηλεκπαίδευση (Σχήμα 1).



Σχήμα 1. Στιγμιότυπο από εξ αποστάσεως σύγχρονο περιβάλλον εκπαίδευσης (zoom) με ταυτόχρονη προβολή των προσώπων των συμμετεχόντων.

Για την εκπαίδευση της τεχνητής νοημοσύνης επιλέχθηκε το σετ FER+ λόγο των πλεονεκτημάτων που προαναφέρθηκαν. Για τη συγγραφή του κώδικα, επιλέχθηκε η χρήση εργαλείων που χρησιμοποιούνται ευρέως από την κοινότητα όπως την OpenCV για επεξεργασία εικόνας και την Python ως βάση για την πλατφόρμα. Τα μοντέλα θα έχουν όλα την μορφή onnx που είναι ένα στάνταρ ανοιχτού λογισμικού. Για την εκτέλεση των μοντέλων θα χρησιμοποιήσουμε την PyTorch που σχεδιάστηκε από την ομάδα του Facebook και δόθηκε για χρήση στην κοινότητα υπό την μορφή ανοιχτού κώδικα. Η πλατφόρμα θα επιτρέπει την εισαγωγή εικόνων ή ροή βίντεο από διάφορες πηγές για αυτό για την αρχική υλοποίηση επιλέξαμε να υποστηρίξουμε τις εξής ροές, βίντεο από Κάμερα Η/Υ, βίντεο ανατροφοδοτώντας την ροή της οθόνης του υπολογιστή καθώς και ροές από κάμερες υπό την μορφή Real Time Streaming Protocol (RTSP) (Schulzrinne et al., 1998).



Σχήμα 2. Σχεδιάγραμμα πλατφόρμας

Στο σχήμα 2, περιγράφονται αναλυτικά τα βήματα που ακολουθούνται κατά την εκπαίδευση αλλά και κατά την διαδικασία ανάλυσης μιας εικόνας που μπορεί να προέλθει από πολλές διαφορετικές πηγές.

Αποτελέσματα

Οι αλγόριθμοι μηχανικής μάθησης είναι από τους πιο χρονοβόρους καθώς απαιτούν προ-επεξεργασία εικόνας και στην συνέχεια υπολογισμό των βαρών σε κάθε νευρώνα του νευρωνικού δικτύου (τεχνητή νοημοσύνη). Στην συνέχεια αυτό πρέπει να επαναληφθεί 28.558 φορές όσες δηλαδή είναι οι εικόνες του παραπάνω σετ. Αυτό ονομάζεται και μια "εποχή" (Epoch) στην ορολογία, στην συνέχεια το νευρωνικό δίκτυο δοκιμάζεται και εξάγονται τα στατιστικά επιτυχίας. Όταν τα αποτελέσματα της "εποχής" είναι καλύτερα κρατούνται (ελιτισμός), αλλιώς απορρίπτονται, έτσι ώστε τελικά να επιλεχθεί η καλύτερη "εποχή", η οποία θα αποτελέσει και το "μοντέλο" (model) του νευρωνικού δικτύου. Ο αλγόριθμος μπορεί να τρέξει όσες "εποχές" του ορίσουμε, στην δική μας περίπτωση επιλέξαμε εκατό (100) εποχές.

Στον Πίνακα 1, παρουσιάζεται μια σύγκριση μεταξύ των επιδόσεων του αλγορίθμου όταν χρησιμοποιείται μόνο η κεντρική μονάδα επεξεργασίας (CPU) ή και η κάρτα γραφικών (GPU), όπου είναι εύκολο κάποιος να διακρίνει ότι η χρήση της κάρτας γραφικών είναι μονόδρομος. Η ολοκλήρωση της εκπαίδευσης χρειάζεται 49 λεπτά και 24 δευτερόλεπτα στην κάρτα γραφικών ενώ στον κεντρικό επεξεργαστή θα χρειαζόταν περίπου 33 ώρες και 30 λεπτά. Παρόμοια χαρακτηριστικά βλέπουμε και κατά την χρήση της πλατφόρμας όπου με την κάρτα γραφικών αναλύονται και τα 30 καρέ ανά δευτερόλεπτο που εξάγει η κάμερά μας ενώ όταν χρησιμοποιείται μόνο ο κεντρικός επεξεργαστής παρατηρείται πτώση και απώλεια καρέ.

Πίνακας 1. Επιδόσεις και χρόνος "εποχής" του αλγορίθμου σε CPU και GPU

Μονάδα Επεξεργασίας	Απώλεια Εκπαίδευσης	Ευστοχία Εκπαίδευσης	Ευστοχία Επαλήθευσης	Ευστοχία Δοκιμών	Χρόνος
CPU	4.146214e+04	33.08%	36.98%	34.52%	1206.689s
GPU	4.154363e+04	32.46%	36.98%	34.52%	29.640s

Κατά την εκπαίδευση του αλγορίθμου χρησιμοποιήθηκαν και οι τέσσερις προσεγγίσεις κατανομής με πέντε επαναλήψεις, εξάγοντας το μέσο όρο τον αποτελεσμάτων. Στον πίνακα 2 βλέπουμε τα αποτελέσματα της αρχικής δημοσίευσης όπου κάθε στήλη περιγράφει μια επανάληψη (trial) και το ποσοστό ευστοχίας που είχε το μοντέλο (model) του νευρωνικού δικτύου κατά αυτή την επανάληψη. Η διαφοροποίηση που παρατηρείται οφείλεται στην αρχικοποίηση του συστήματος η οποία είναι τυχαία αλλά συγκλίνει όσο το πρόγραμμα εκπαιδεύεται. Στην τελευταία στήλη περιγράφεται ο μέσος όρος της ευστοχίας των μοντέλων μετά από πέντε επαναλήψεις καθώς και η απόκλιση της.

Πίνακας 2. Δεδομένα αρχικής δημοσίευσης FER+

Κατανομή	Επανάληψη 1	Επανάληψη 2	Επανάληψη 3	Επανάληψη 4	Επανάληψη 5	Ευστοχία
MV	83.60%	84.89%	83.15%	83.39%	84.23%	83.852% ± 0.631%
MLL	83.69%	83.63%	83.81%	84.62%	84.08%	83.966% ± 0.362%
PLD	85.43%	84.65%	85.34%	85.01%	84.50%	84.986% ± 0.366%
CEL	85.01%	84.59%	84.32%	84.80%	84.86%	84.716% ± 0.239%

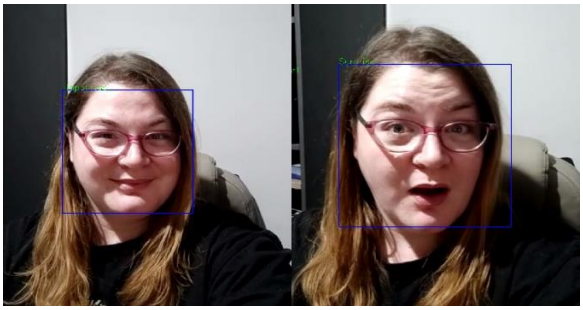
Πίνακας 3. Επαλήθευση δεδομένων χρησιμοποιώντας την πλατφόρμα μας

Κατανομή	Επανάληψη 1	Επανάληψη 2	Επανάληψη 3	Επανάληψη 4	Επανάληψη 5	Ευστοχία
MV	83.81%	82.98%	83.39%	83.39%	84.23%	83.560% ± 0.670%
MLL	83.49%	84.22%	83.36%	82.95%	83.93%	83.590% ± 0.630%
PLD	82.82%	83.10%	82.82%	83.10%	83.33%	83.034% ± 0.214%
CEL	83.55%	83.74%	83.61%	83.77%	83.39%	83.612% ± 0.222%

Από τους πίνακες 2 & 3, γίνεται φανερό ότι τα αποτελέσματα στα οποία καταλήξαμε, είναι παρόμοια με αυτά της αρχικής έρευνας και εντός του περιθωρίου σφάλματος. Έτσι επιβεβαιώνουμε ότι το σύστημα τεχνητής νοημοσύνης τρέχει με βάση τα αρχικά δεδομένα ώστε να μπορούμε να το χρησιμοποιήσουμε στην πλατφόρμα.

Η πλατφόρμα μας επιτρέπει να εισάγουμε ροές οι οποίες μπορεί να προέρχονται από κάποιο ρομπότ η ηλεκτρονική συσκευή με περιορισμένες υπολογιστικές δυνατότητες και στη συνέχεια να προωθούνται για επεξεργασία σε κάποια συστοιχία που μπορεί να περιέχει ακόμα και πολλαπλές κάρτες γραφικών για ακόμα καλύτερα αποτελέσματα.

Αναφορικά με το δεύτερο ερευνητικό μας ερώτημα, αφού επαληθεύσαμε τα αποτελέσματα της αρχικής έρευνας (Barsoum et al., 2016) και εξάγαμε τα μοντέλα, επιλέξαμε το μοντέλο της Επανάληψης 5 (Πίνακας 3) που έφερε τα καλύτερα αποτελέσματα στην κατηγορία Majority Voting. Η συγκεκριμένη κατηγορία επιλέχθηκε, γιατί επιστρέφει ένα μόνο συναίσθημα (το κυρίαρχο) καθώς και την πιθανότητα του να είναι αληθές. Με αυτό τον τρόπο μπορούμε να το προβάλουμε σε πραγματικό χρόνο πάνω στην ροή βίντεο που μας δίνεται (Σχήμα 3)



Σχήμα 3. Στιγμιότυπο από την χρήση της πλατφόρμας.

Στο σχήμα 3 παρουσιάζονται δύο στιγμιότυπα από την χρήση της πλατφόρμας όπου στο πρόσωπο της συμμετέχουσας απεικονίζεται το εκάστοτε κυρίαρχο συναίσθημα.

Συμπεράσματα

Στην εργασία αυτή παρουσιάστηκε μια πλατφόρμα με την οποία έγινε εφικτή η δοκιμή διαφόρων αλγορίθμων αναγνώρισης συναισθημάτων από ροή video σε πραγματικό χρόνο. Η δοκιμή των αλγορίθμων, αποδείχθηκε επιτυχής, καθώς επαλήθευσε τα δεδομένα από προηγούμενες έρευνες, ενώ επιλέχθηκε και το καλύτερο "μοντέλο" για την εφαρμογή σε πραγματικό χρόνο.
Καθώς βλέπουμε ήδη τα πρώτα ρομπότ σε σχολικές αίθουσες η πανεπιστήμια, η συναισθηματική εμπλοκή των μαθητών με τα ρομπότ είναι ίσως ένα από τα προφανή ερωτήματα που θα προκύψουν. Η δυνατότητα των ρομπότ να αναγνωρίζουν τα συναισθήματα του ατόμου ή ακόμα και μιας ολόκληρης σχολικής τάξης μας ανοίγει νέους ορίζοντες σε σχέση με την διεπαφή και την αλληλεπίδραση που έχουν οι μαθητές ή φοιτητές με αυτά (Belpaeme & Kennedy, 2018). Η βοήθεια που μπορεί να παρέχει στην εκπαίδευση η τεχνητή νοημοσύνη είναι ακόμα πιο επίκαιρη σήμερα, όπου λόγω των περιορισμών που έχουν επιβληθεί εξαιτίας του COVID-19 πολλές από τις δραστηριότητες γίνονται εξ 'αποστάσεως.

Σε επόμενα βήματα θα γίνει προσπάθεια βελτίωσης των αλγορίθμων που χρησιμοποιούνται όπως και συνδυασμός τους και με άλλες πηγές εισόδου όπως π.χ. ήχο, βιομετρικά στοιχεία κλπ. Επίσης, αναμένεται η εφαρμογή του αλγορίθμου σε διαδικασίες τηλεκπαίδευσης, ώστε να δοκιμαστεί σε πραγματικό περιβάλλον.

Αναφορές

Barsoum, E., Zhang, C., Ferrer, C. C., & Zhang, Z. (2016). Training deep networks for facial expression recognition with crowd-sourced label distribution. *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. https://doi.org/10.1145/2993148.2993165

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics*, *3*(21), eaat5954. https://doi.org/10.1126/scirobotics.aat5954

Carmona-Halty, M., Salanova, M., Llorens, S., & Schaufeli, W. B. (2019). Linking positive emotions and academic performance: The mediated role of academic psychological capital and academic engagement. *Current Psychology*, 1–10. https://doi.org/10.1007/s12144-019-00227-8

Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-based measurement of facial expression with the Facial Action Coding System. In J. Coan & J. Allen (Eds.), *The handbook of emotion elicitation and assessment*, 1(3), 203-221.

Dael, N., Mortillaro, M. & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion 12,* 1085–1101. doi: 10.1037/a0025737

Dosilovic, F. K., Brcic, M., & Hlupic, N. (2018). Explainable artificial intelligence: A survey. *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. doi:10.23919/mipro.2018.8400040

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124-129. doi:10.1037/h0030377

Goodfellow, I. J., Erhan, D., Luc Carrier, P., Courville, A., Mirza, M., Hamner, B., … Bengio, Y. (2015). Challenges in representation Learning: A report on three machine LEARNING CONTESTS. *Neural Networks*, *64*, 59–63. https://doi.org/10.1016/j.neunet.2014.09.005

Jack, R. E. & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Curr. Biol*. 25, R621–R634. doi: 10.1016/j.cub.2015.05.052

Krithika, L. B., & Lakshmi Priya, G. G. (2016). Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric. *Procedia Computer Science*, *85*(Cms), 767–776. https://doi.org/10.1016/j.procs.2016.05.264

Kusuma, G. P., Jonathan, J., & Lim, A. P. (2020). Emotion Recognition on FER-2013 Face Images Using Fine-Tuned VGG-16. *Advances in Science, Technology and Engineering Systems Journal*, *5*(6), 315–322. https://doi.org/10.25046/aj050638

Lanphier, R., & Rao, A. (1998). Real Time Streaming Protocol (RTSP). *RFC 2326.* https://tools.ietf.org/html/rfc2326.

Lee, H. J., & Lee, D. (2021). Study of process-focused assessment using an algorithm for facial expression recognition based on a deep neural network model. *Electronics (Switzerland)*, *10*(1), 1–14. https://doi.org/10.3390/electronics10010054

Li, L., Gow, A. D. I., & Zhou, J. (2020). The Role of Positive Emotions in Education: A Neuroscience Perspective. *Mind, Brain, and Education*, 14(3), 220-234.

Lloyd-Jones, B. (2021). Developing Competencies for Emotional, Instrumental, and Informational Student Support During the COVID-19 Pandemic: A Human Relations/Human Resource Development Approach. *Advances in Developing Human Resources*, *23*(1), 41–54. https://doi.org/10.1177/1523422320973287

Oudeyer, P.-Y., Gottlieb, J., and Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: theory and applications in educational technologies. In B. Studer & S. Knecht (Eds.), *Progress in Brain Research*. Elsevier, 229, 257–284. doi: 10.1016/bs.pbr.2016.05.005

Reisenzein, R., Horstmann, G. and Schützwohl, A. (2019). The Cognitive-Evolutionary Model of Surprise: A Review of the Evidence. *Topics in Cognitive Science,* 11(1), 50–74. doi: 10.1111/tops.12292.

Russell, J. A., Bachorowski, J.-A., and Fernández-Dols, J.-M. (2003). Facial and vocal expressions of emotion. *Annu. Rev. Psychol*. 54, 329–349. doi: 10.1146/annurev.psych.54.101601.145102

Ruvolo, P., Fasel, I., & Movellan, J. (2008). Auditory mood detection for social and educational robots. *2008 IEEE International Conference on Robotics and Automation*. doi:10.1109/robot.2008.4543754

Simonyan, K., & Zisserman, A. (2014, December 23). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv.org.* https://arxiv.org/abs/1409.1556v5.

Tonguç, G., & Ozaydin Ozkara, B. (2020). Automatic recognition of student emotions from facial expressions during a lecture. *Computers and Education,* 148, 103797. https://doi.org/10.1016/j.compedu.2019.103797

Tyng, C. M., Amin, H. U., Saad, M. N., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in psychology*, 8, 1454.

Velentza, A.M. (2021). Putting the Humans in the Middle of the CPS Design Process. In I. Papaefstathiou & A. Hatzopoulos (Eds), *Heterogeneous Cyber Physical Systems of Systems*, River Publishers, ch. 5, pp. 175–209.

Villavicencio, F. T., & Bernardo, A. B. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*, 83(2), 329-340.