

Θεσσαλονικιός πίσω από σύστημα τεχνητής νοημοσύνης της Google που «διαβάζει» τα χείλη

Από **parallaxi** - August 1, 2018



Οι κωφοί -και οι απανταχού κατάσκοποι- θα βρουν χρήσιμο ένα νέο πρόγραμμα τεχνητής νοημοσύνης που 'διαβάζει' τα χείλη των ανθρώπων εξ αποστάσεως με μεγαλύτερη ακρίβεια όχι μόνο από κάθε άλλο μέχρι σήμερα, αλλά και από τους ανθρώπους με τη σχετική εξειδίκευση.

Το πρόγραμμα ανέπτυξαν ερευνητές της Google και της βρετανικής θυγατρικής της Deep Mind, η οποία πρωτοπορεί διεθνώς στην τεχνητή νοημοσύνη, με επικεφαλής ένα νεαρό ελληνικής καταγωγής ειδικό της πληροφορικής, τον Γιάννη Ασασέλ.

Για τα εκατομμύρια των ανθρώπων που δεν μπορούν να ακούσουν, η ανάγνωση των χειλιών των άλλων αποτελεί ένα «παράθυρο» για την επικοινωνία, πέρα από τη νοηματική γλώσσα. Όμως το «διάβασμα» των χειλιών δεν είναι εύκολο και συχνά είναι ανακριβές.

Το νέο «έξυπνο» σύστημα εμφανίζει μέσο ποσοστό λαθών 41% στην ορθή κατανόηση των λέξεων που σχηματίζουν τα χείλη. Μπορεί να φαίνεται υψηλό αυτό το ποσοστό, αλλά η καλύτερη μέχρι σήμερα υπολογιστική μέθοδος είχε ποσοστό αποτυχίας 77%, συνεπώς το νέο πρόγραμμα έχει περιορίσει σχεδόν στο μισό τα λάθη.

Οι ερευνητές, με επικεφαλής τον Ασασέλ και τον Μπρένταν Σίλινγκφορντ, οι οποίοι έκαναν τη σχετική προδημοσίευση στο arXiv, σύμφωνα με το "Science", δημιούργησαν αλγόριθμους καλύτερους από κάθε άλλους στο παρελθόν, που τα καταφέρνουν αποτελεσματικότερα και από επαγγελματίες αναγνώστες χειλιών. Αν το νέο σύστημα τελειοποιηθεί και ενσωματωθεί σε ηλεκτρονικές συσκευές, όπως τα «έξυπνα» κινητά τηλέφωνα, μπορεί μελλοντικά να προσφέρει τη δυνατότητα ανάγνωσης των χειλιών στον καθένα (κάτι που μάλλον δεν θα αρέσει σε όλους...).

Η δημιουργία αλγορίθμων που να μπορούν να διαβάσουν χείλη ήταν έως τώρα τρομερά δύσκολη. Οι ερευνητές τροφοδότησαν το σύστημά τους με 140.000 ώρες βίντεο από το YouTube που έδειχναν ανθρώπους να μιλάνε στα αγγλικά, μαζί με τα αντίστοιχα απομαγνητοφωνημένα κείμενα. Στη συνέχεια, άφησαν το σύστημα μηχανικής μάθησης μόνο του να μάθει να «παντρεύει» τις διαφορετικές κινήσεις των χειλιών με τα αντίστοιχα φωνήματα και τελικά

Το σύστημα βασίζεται στα τεχνητά νευρωνικά δίκτυα, δηλαδή σε μια ομάδα αλγόριθμων που ο καθένας εκτελεί ένα διαφορετικό και απλούστερο έργο, ενώ ταυτόχρονα όλοι μαζί συνδέονται και συνεργάζονται για να επεξεργασθούν τις πληροφορίες, περίπου όπως κάνουν οι νευρώνες του ανθρώπινου εγκεφάλου.

Αφού το σύστημα «αυτο-εκπαιδεύθηκε», δοκιμάστηκε από τους ερευνητές στην ανάγνωση των χειλιών των ανθρώπων ενός 37λεπτου βίντεο, που ποτέ πριν το σύστημα δεν είχε «δει».

Το πρόγραμμα έσφαλε στο 41% των λέξεων, αλλά οι άνθρωποι -ακόμη και οι ειδικοί στο διάβασμα του στόματος- που είδαν το ίδιο βίντεο, είχαν μέσο ποσοστό αποτυχίας 93% (πάντως σε πραγματικές συνθήκες και όχι παρακολούθησης βίντεο, το ποσοστό ανθρώπινης αποτυχίας είναι κάπως μικρότερο, καθώς ο ανθρώπινος εγκέφαλος μπορεί να αξιοποιήσει και άλλα στοιχεία, όπως η γλώσσα του σώματος αυτού που μιλάει).

Σε κάθε περίπτωση, αν και συνιστά πραγματική πρόοδο, είναι σαφές ότι ένα ποσοστό αποτυχίας της τάξης του 40% στην αναγνώριση των λέξεων σημαίνει ότι το σύστημα χρειάζεται ακόμη ουσιαστική βελτίωση. Όταν αυτή υπάρξει, το σύστημα μπορεί να αξιοποιηθεί ευρύτερα και στην καθημερινότητα. Για παράδειγμα, όταν θα βλέπουμε στην τηλεόραση έναν ποδοσφαιριστή να βρίζει ένα διαιτητή ή έναν αντίπαλο παίκτη, αλλά φυσικά δεν ακούμε τι λέει, θα ξέρουμε πια τι είπε!

Ο Γιάννης Ασσαέλ υπήρξε μαθητής του Κολλεγίου Ανατόλια της Θεσσαλονίκης και σπούδασε εφαρμοσμένη πληροφορική στο Πανεπιστήμιο Μακεδονίας (2008-2013). Στη συνέχεια, έκανε μεταπτυχιακά (μάστερ) στην επιστήμη των υπολογιστών στο Πανεπιστήμιο της Οξφόρδης, όπου σήμερα ολοκληρώνει το διδακτορικό του πάνω στη μηχανική μάθηση, ενώ παράλληλα εργάζεται για τη Deep Mind της Google.

Σύνδεσμος για την επιστημονική προδημοσίευση: <https://arxiv.org/abs/1807.05162>

Πηγή: ΑΠΕ, Πάυλος Δρακόπουλος

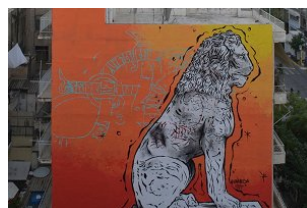
ΔΙΑΒΑΣΤΕ ΕΠΙΣΗΣ



Η προειδοποίηση του Πρίγκιπα Καρόλου που όλοι αγνόησαν



Τα τελευταία νέα για την κατάσταση της υγείας του Λεωνίδα Καλαφγιάννη



Η νέα τοχογραφία, σκέτο κόσμημα!