

Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης

Σχολή Τεχνολογικών Εφαρμογών
Τμήμα Μηχανικών Πληροφορικής



Πτυχιακή Εργασία

Τίτλος:

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

Μπολαράκης Παναγιώτης (ΑΜ: 3262)

Επιβλέπων καθηγητής: Τσικνάκης Μανώλης

Επιτροπή αξιολόγησης:

Τσικνάκης Μανώλης, Ακουμιανάκης Δημοσθένης, Βιδάκης Νικόλαος

Ημερομηνία παρουσίασης: 21/03/2017

Υπεύθυνη Δήλωση

Βεβαιώνω ότι είμαι συγγραφέας αυτής της πτυχιακής εργασίας και κάθε βοήθεια την οποία είχα για τη προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στη πτυχιακή εργασία. Έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά για τις απαιτήσεις του προγράμματος σπουδών του Τμήματος Μηχανικών Πληροφορικής του Τ.Ε.Ι. Κρήτης.

Ευχαριστίες

Η εκπόνηση της παρούσας πτυχιακής εργασίας με θέμα «Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου» στο Τμήμα Μηχανικών Πληροφορικής, Τ.Ε.Ι. Ηρακλείου πραγματοποιήθηκε υπό την επίβλεψη του καθηγητή Μανώλη Τσακνάκη.

Αρχικά θα ήθελα να εκφράσω τις ευχαριστίες μου στον καθηγητή μου Μανώλη Τσακνάκη για την εμπιστοσύνη και την κατανόηση που μου έδειξε, καθώς και την υποστήριξη που μου πρόσφερε. Επίσης θα ήθελα να ευχαριστήσω θερμά τον Ματθαίο Πεδιαδίτη για την βοήθεια που μου πρόσφερε καθ' όλη την διάρκεια της υλοποίησης της πτυχιακής εργασίας.

Θα ήθελα να ευχαριστήσω όλους τους συνεργάτες του εργαστηρίου Βιοϊατρικής Πληροφορικής & Ηλεκτρονικής Υγείας, Τ.Ε.Ι. Ηρακλείου για την βοήθεια που μου πρόσφεραν ο καθένας τους ξεχωριστά. Τέλος, ευχαριστώ θερμά την Γιώτα Δημητράκη που χωρίς την υποστήριξή και την βοήθεια της η υλοποίηση της παρούσας πτυχιακής εργασίας δεν θα ήταν εφικτή.

Abstract

Psychologist Paul Ekman, after extensive research on the facial expressions made in different populations around the world, concluded that six basic emotions are universal, meaning that they are understood and globally recognized regardless of cultural factors. These emotions are those of happiness, sadness, fear, anger, surprise and disgust. While, for people, something as simple as the detection of the face is achieved automatically and interpretation of emotions is relatively easy, for a computer or a computing system to automatically identify and recognize successfully facial expressions in real time, is a difficult task. The purpose of this thesis is to develop 5 automatic systems for the identification of the six basic emotions and the neutral expression. For this purpose, two different methods were implemented and tested for the widespread Cohn-Kanade database [1]. In the first method the algorithm proposed by J. Kalita and K. Das [2], that uses eigenvalues and eigenvectors for feature extraction from areas of the face and Euclidean distance as a decision-making technique, is implemented, and also a second variation of the same method that makes use of a “mean” eigenvector. The second method used artificial neural networks to recognize facial expressions for three different image processing techniques. Initially, without any further processing other than the accurate identification of the face, then creating and applying on the face image a Gabor filter bank and finally by using the subtraction of the neutral expression from the other expressions as a filter. While for the first method, the results were not as good as of J. Kalita and K. Das [2] presented in their paper, the recognition rate for the second method, that uses artificial neural networks and Gabor filters equals to 88% for the test images.

Keywords: Automatic emotion recognition, artificial neural networks, eigenvectors, eigenvalues, Euclidean distance, Gabor filters, Viola-Jones algorithm.

Σύνοψη

Ο ψυχολόγος Paul Ekman, μετά από εκτενείς έρευνες πάνω στις εκφράσεις του προσώπου που έγιναν σε διάφορους πληθυσμούς σε όλο τον κόσμο, κατέληξε στο συμπέρασμα ότι έξι βασικά συναισθήματα είναι καθολικά, δηλαδή γίνονται κατανοητά και αναγνωρίζονται παγκόσμια ανεξάρτητα από πολιτισμικούς παράγοντες. Τα συναισθήματα αυτά είναι αυτά της ευτυχίας, της λύπης, του φόβου, του θυμού, της έκπληξης και της αηδίας. Ενώ, για τους ανθρώπους, κάτι απλό όπως η ανίχνευση του προσώπου επιτυγχάνεται αυτόματα και η ερμηνεία των συναισθημάτων γίνεται σχετικά εύκολα. Για έναν υπολογιστή ή ένα υπολογιστικό σύστημα που πρέπει αυτόματα να εντοπίζει και να αναγνωρίζει με επιτυχία τις εκφράσεις του προσώπου, σε πραγματικό χρόνο, είναι δύσκολη υπόθεση. Στα πλαίσια αυτής της πτυχιακής εργασίας αναπτύχθηκαν 5 αυτόματα συστήματα για την αναγνώριση των έξι βασικών συναισθημάτων και της ουδέτερης έκφρασης. Για τον σκοπό αυτό δύο διαφορετικές μέθοδοι υλοποιήθηκαν και δοκιμάστηκαν στη ευρέως διαδεδομένη βάση Cohn-Kanade [1]. Στην πρώτη μέθοδο υλοποιήθηκε ο αλγόριθμος των J. Kalita και K. Das [2] που έκαναν χρήση ιδιοτιμών και ιδιοδιανυσμάτων για την εξαγωγή χαρακτηριστικών από περιοχές του προσώπου και χρήση της Ευκλείδειας απόστασης σαν τεχνική λήψης αποφάσεων, καθώς και μία δεύτερη υλοποίηση της ίδιας μεθόδου που έκανε χρήση του μέσου ιδιοδιανύσματος. Στην δεύτερη μέθοδο χρησιμοποιήθηκαν τεχνητά νευρωνικά δίκτυα για να γίνει η αναγνώριση των εκφράσεων για τρεις διαφορετικούς τρόπους επεξεργασίας της εικόνας. Αρχικά χωρίς κάποια περαιτέρω επεξεργασία, εκτός από της ακριβή αναγνώριση του προσώπου, μετά με την χρήση και εφαρμογή τράπεζας φίλτρων Gabor και τέλος με την χρήση της αφαίρεσης της ουδέτερης έκφρασης από τις υπόλοιπες εκφράσεις σαν φίλτρο. Ενώ για την πρώτη μέθοδο τα αποτελέσματα δεν ήταν αυτά που θα περιμέναμε σε σχέση με αυτά που παρουσίασαν οι J. Kalita και K. Das [2], η μέθοδος που έκανε χρήση τεχνητών νευρωνικών δικτύων και εφαρμογή φίλτρων Gabor έδωσε ποσοστό επιτυχούς αναγνώρισης ίσο με 88% για τις δοκιμαστικές εικόνες.

Λέξεις κλειδιά: Αυτόματη αναγνώριση συναισθημάτων, τεχνητά νευρωνικά δίκτυα, ιδιοδιανύσματα, ιδιοτιμές, Ευκλείδεια απόσταση, φίλτρα Gabor, Viola-Jones αλγόριθμος.

Περιεχόμενα

Υπεύθυνη Δήλωση	II
Ευχαριστίες	III
Abstract.....	IV
Σύνοψη.....	V
Λίστα Εικόνων.....	VIII
Λίστα Πινάκων	X
1. Εισαγωγή	12
1.1 Σήματα πληροφορίας και τα βασικά συναισθήματα.....	12
1.2 Ιστορική αναδρομή	14
1.3 Εφαρμογές των αυτόματων συστημάτων αναγνώρισης συναισθημάτων	15
2. Ανασκόπηση σχετικής βιβλιογραφίας.....	17
3. Μέθοδοι και βάση δεδομένων	22
3.1 Εισαγωγή στο περιβάλλον χρήσης.....	22
3.2 Βάση δεδομένων για εικόνες και βίντεο	23
4. Μέθοδος 1: Ιδιοδιανύσματα, Ιδιοτιμές και Ευκλείδεια απόσταση	25
4.1 Θεωρητικό υπόβαθρο.....	25
4.2 Σχεδιασμός υλοποίησης – Αλγόριθμος των δύο υλοποιήσεων	27
4.3 Υλοποίηση συστήματος – Κώδικας των δύο υλοποιήσεων	31
4.3.1 Κώδικας πρώτης υλοποίησης	31
4.3.2 Κώδικας δεύτερης υλοποίησης.....	44
4.4 Αποτελέσματα συστημάτων – Συζήτηση	45
4.4.1 Αποτελέσματα πρώτης υλοποίησης.....	45
4.4.2 Αποτελέσματα δεύτερης υλοποίησης.....	48
4.4.3 Συζήτηση αποτελεσμάτων για τις δύο υλοποιήσεις	48
5. Μέθοδος 2: Τεχνητά Νευρωνικά Δίκτυα	50
5.1 Θεωρητικό υπόβαθρο.....	50
5.2 Σχεδιασμός υλοποίησης – Αλγόριθμος των τριών υλοποιήσεων	52
5.3 Υλοποίηση συστήματος – Κώδικας των τριών υλοποιήσεων	55
5.3.1 Κώδικας πρώτης υλοποίησης	55
5.3.2 Κώδικας δεύτερης υλοποίησης.....	60

5.3.3 Κώδικας τρίτης υλοποίησης	61
5.4 Αποτελέσματα συστημάτων – Συζήτηση	62
5.4.1 Αποτελέσματα πρώτης υλοποίησης.....	62
5.4.2 Αποτελέσματα δεύτερης υλοποίησης.....	69
5.4.3 Αποτελέσματα τρίτης υλοποίησης.....	75
6. Συμπεράσματα	81
Βιβλιογραφία	82

Λίστα Εικόνων

Εικόνα 1: Το σύστημα εξαγωγής χαρακτηριστικών του προσώπου και ανάλυσης των εκφράσεων που δημιούργησαν οι Ioannou et al. [8]	18
Εικόνα 2: Ο πίνακας των Zhihong Zeng et al. για την αναγνώριση συναισθημάτων από συστήματα που βασίζονται σε οπτικά μέσα όπως εικόνες και βίντεο για την υλοποίησή τους [9].....	21
Εικόνα 3: παραδείγματα ορθογωνίων χαρακτηριστικών Haar	28
Εικόνα 4: Γράφημα workflow για τα βήματα που γίνονται κατά την διάρκεια επεξεργασίας για τις εικόνες εκπαίδευσης.	29
Εικόνα 5: Γράφημα workflow για τα βήματα που γίνονται κατά την διάρκεια επεξεργασίας για τις εικόνες δοκιμής.....	30
Εικόνα 6: Γράφημα workflow για τα βήματα που γίνονται κατά την διάρκεια επεξεργασίας για τις εικόνες εκπαίδευσης της δευτερης μεθόδου.	31
Εικόνα 7: Παράδειγμα ενός απλού νευρώνα.	50
Εικόνα 8: Κάποιες συναρτήσεις ενεργοποίησης που χρησιμοποιούνται στα νευρωνικά δίκτυα	51
Εικόνα 9: Το αποτέλεσμα της διαφοράς μεταξύ ουδέτερης έκφρασης και του συναισθήματος της έκπληξης.	55
Εικόνα 10: Το αποτέλεσμα της διαφοράς μεταξύ ουδέτερης έκφρασης και του συναισθήματος της αηδίας.....	55
Εικόνα 11: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [20 20].....	64
Εικόνα 12: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [50 50].....	65
Εικόνα 13: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [60 40].....	66
Εικόνα 14: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [20 10].....	67
Εικόνα 15: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50].....	68
Εικόνα 16: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [35 45].....	69
Εικόνα 17: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50], 6+1 φίλτρα Gabor 100x100 με προσανατολισμούς [0 45 90] και μήκη κύματος [4 8].....	71
Εικόνα 18: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50], 8+1 φίλτρα Gabor 100x100 με προσανατολισμούς [0 45 90 135] και μήκη κύματος [4 8].....	72
Εικόνα 19: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50], 8+1 φίλτρα Gabor 100x100 με προσανατολισμούς [0 45 90 135] και μήκη κύματος [4 8].....	74

Εικόνα 20: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [20 30], 40 φίλτρα Gabor 50x50 με προσανατολισμούς [2 4 6 8 10] και μήκη κύματος [$0 \pi/8$ $2\pi/8$ $3\pi/8$ $4\pi/8$ $5\pi/8$ $6\pi/8$ $7\pi/8$].	75
Εικόνα 21: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [24 24].	77
Εικόνα 22: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [36 36].	78
Εικόνα 23: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [100 100].	79
Εικόνα 24: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50].	80

Λίστα Πινάκων

Πίνακας 1: Κατανομή εικόνων της βάσης ανά συναίσθημα.	24
Πίνακας 2: Οι πέντε Ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το αριστερό μάτι μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.	45
Πίνακας 3: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το δεξί μάτι μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.	45
Πίνακας 4: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το στόμα μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.	46
Πίνακας 5: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για την μύτη μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.	46
Πίνακας 6: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το στόμα και την μύτη μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.	46
Πίνακας 7: Ο τελικός πίνακας ψήφων που μετράει τις ψήφους για κάθε συναίσθημα ανά περιοχή και αποφασίζει πιο είναι το συναίσθημα που εκφράζει η δοκιμαστική εικόνα.	46
Πίνακας 8: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 25 ευκλείδειες αποστάσεις.	47
Πίνακας 9: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 75 ευκλείδειες αποστάσεις.	47
Πίνακας 10: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 150 ευκλείδειες αποστάσεις.	47
Πίνακας 11: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 100 ευκλείδειες αποστάσεις χωρίς τα πέντε πρώτα ιδιοδιανύσματα κάθε περιοχής.	47
Πίνακας 12: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 25 ευκλείδειες αποστάσεις.	48
Πίνακας 13 Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 75 ευκλείδειες αποστάσεις.	48
Πίνακας 14: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 150 ευκλείδειες αποστάσεις.	48
Πίνακας 15: Ποσοστά και αριθμός εικόνων που χρησιμοποιήθηκαν για εκπαίδευση, επικύρωση και δοκιμή του νευρωνικού δικτύου για τις δύο πρώτες υλοποιήσεις.	53
Πίνακας 16: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [20 20] στην πρώτη και [50 50] στην δεύτερη.	63
Πίνακας 17: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά	

στρώματα και στις δύο, με αριθμό νευρώνων [60 40] στην πρώτη και [20 10] στην δεύτερη.	66
Πίνακας 18: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [25 50] στην πρώτη και [30 45] στην δεύτερη.	68
Πίνακας 19: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και αριθμό νευρώνων [25 50] και στις δύο.	70
Πίνακας 20: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [25 50] και συνάρτηση απόδοσης Mean Squared Error στην πρώτη και αριθμό νευρώνων [30 45] και συνάρτηση απόδοσης Cross-Entropy στην δεύτερη.....	73
Πίνακας 21: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [24 24] και συνάρτηση απόδοσης Mean Squared Error στην πρώτη και αριθμό νευρώνων [36 36] και συνάρτηση απόδοσης Cross-Entropy στην δεύτερη.....	76
Πίνακας 22: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [100 100] στην πρώτη και αριθμό νευρώνων [25 50] στην δεύτερη.	79

1. Εισαγωγή

Παράλληλα με την ραγδαία ανάπτυξη της τεχνολογίας τα τελευταία χρόνια, η ανάπτυξη συστημάτων για την αυτόματη αναγνώριση συναισθημάτων έχει επίσης συναντήσει σημαντική πρόοδο. Το πρόβλημα που καλούνται να λύσουν τα συστήματα αυτά είναι η έγκαιρη αναγνώριση των συναισθημάτων που εκφράζει ένα άτομο ώστε να ληφθούν τα κατάλληλα σε κάθε περίπτωση μέτρα. Ένα παράδειγμα εφαρμογής της αυτόματης αναγνώρισης συναισθημάτων είναι η πρόβλεψη των ενεργειών και της συμπεριφοράς ατόμων που πάσχουν από κατάθλιψη με βάση τα συναισθήματα που εκφράζουν στο πρόσωπό τους, ώστε να μπορούν οι επιβλέπων ιατροί τους να προσαρμόσουν κατάλληλα την θεραπεία τους. Τα συστήματα αυτόματης αναγνώρισης συναισθημάτων έχουν πολλές άλλες εφαρμογές στους τομείς της ιατρικής, της ηλεκτρονικής μάθησης, της παρακολούθησης ή επίβλεψης, της ψυχαγωγίας, του δικαίου ή της νομικής, του μάρκετινγκ και άλλων. Στόχος της παρούσας Πτυχιακής Εργασίας είναι η δημιουργία ενός αυτόματου συστήματος αναγνώρισης συναισθημάτων μέσω των εκφράσεων του προσώπου. Ένα αυτόματο σύστημα αναγνώρισης συναισθημάτων είναι ένα υπολογιστικό σύστημα που αξιοποιεί τεχνικές από πολλούς τομείς, όπως επεξεργασία σήματος (signal processing), μηχανική μάθηση (machine learning), υπολογιστική όραση (computer vision), για να ερμηνεύσει έναν αριθμό συναισθημάτων.

Υπάρχουν πάρα πολλοί ορισμοί για το συναίσθημα αλλά κανένας από αυτούς δεν είναι καθολικός. Σύμφωνα με τον P.Ekman [3] «ένα συναίσθημα είναι μία κατάσταση που διαρκεί από μερικά λεπτά έως και μερικές ώρες». Ο Joseph Campos οριοθετεί το συναίσθημα ως «το αίσθημα ή την συγκίνηση, που εκφράζεται όταν ένα άτομο συμμετέχει σε μία σημαίνουσα αλληλεπίδραση για αυτόν/αυτήν, ιδιαίτερα σημαντική για την ευεξία του/της» [4]. Ο P.Ekman, όπως και πολλοί άλλοι, υποστηρίζουν ότι 6 βασικά συναισθήματα είναι παγκόσμια αναγνωρίσιμα ανεξάρτητα από την κουλτούρα και την γεωγραφική τοποθεσία ενός λαού. Αυτά τα συναισθήματα είναι η ευτυχία, η θλίψη, ο θυμός, η αηδία, η έκπληξη και ο φόβος [3]. Με βάση αυτά πραγματοποιείται η προσπάθεια υλοποίησης ενός αυτόματου συστήματος αναγνώρισης συναισθημάτων, που βασίζεται στις εκφράσεις του προσώπου από στατικές εικόνες.

Σύμφωνα με τον P.Ekman [3] για την αναγνώριση ενός συναισθήματος το μόνο που χρειάζεται κάποιος να κάνει είναι να κοιτάξει το πρόσωπο του συνομιλητή του. Η εναλλαγή των εκφράσεων του προσώπου περιέχει μεγάλο όγκο πληροφορίας για το συναίσθημα το οποίο εκφράζει ένα άτομο. Παρόλα αυτά σε πολλές περιπτώσεις χρειάζεται η επίγνωση των συμφραζόμενων (context), καθώς και άλλοι τύποι πληροφορίας όπως ακουστική πληροφορία και γλώσσα του σώματος. Από τους ανθρώπους η αναγνώριση των εκφράσεων του προσώπου γίνεται αυτόματα και άμεσα χωρίς ιδιαίτερη προσπάθεια εκτός και αν το άτομο που εκφράζει κάποιο συναίσθημα προσπαθεί να το κρύψει ή να το συγκαλύψει. Αυτή η χωρίς προσπάθεια αναγνώριση των συναισθημάτων που έχουν έμφυτη οι άνθρωποι δεν την έχουν και οι υπολογιστές, κάτι που καθιστά την δημιουργία ενός αυτόματου συστήματος αναγνώρισης συναισθημάτων δύσκολη και περίπλοκη υπόθεση.

1.1 Σήματα πληροφορίας και τα βασικά συναισθήματα

Υπάρχουν τρία είδη σημάτων πληροφορίας που παρέχει το πρόσωπο, τα στατικά, τα βραδεία και τα ταχεία σήματα [3]. Τα στατικά σήματα παρέχουν πληροφορία για το χρώμα

του δέρματος, το σχήμα του προσώπου, το μέγεθος, το σχήμα και την θέση των χαρακτηριστικών του προσώπου, πληροφορία δηλαδή που δεν αλλάζει στο πρόσωπο ενός ατόμου. Τα βραδεία σήματα δίνουν πληροφορία που επέρχεται σταδιακά με το πέρασμα του χρόνου, όπως μόνιμες ρυτίδες και φακίδες, σταδιακή αλλαγή του τόνου του δέρματος κτλ. Τέλος τα ταχεία σήματα δίνουν πληροφορία για την κίνηση των μυών του προσώπου, για προσωρινές ρυτίδες και αλλαγές στην τοποθεσία των χαρακτηριστικών του προσώπου που προκαλούν οι κινήσεις αυτές και διαρκούν λίγα δευτερόλεπτα ή στην περίπτωση των μικρό-εκφράσεων (micro expressions) κλάσματα του δευτερολέπτου. Αυτά τα ταχεία σήματα πληροφορίας είναι που χρησιμοποιούνται στα αυτόματα συστήματα για την αναγνώριση συναισθημάτων.

Ο P.Ekman στο βιβλίο του “Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions” [3] [5] αναλύει λεπτομερώς κάθε ένα από τα έξι βασικά συναισθήματα. Στις επόμενες παραγράφους δίνεται μία μικρή περιγραφή του κάθε συναισθήματος και του τρόπου που αλλάζουν τα χαρακτηριστικά του προσώπου για να εκφράσουν το συναίσθημα αυτό όπως τα έχει ορίσει ο P.Ekman στο βιβλίο του.

Το συναίσθημα της έκπληξης είναι το πιο σύντομο από τα συναισθήματα. Αν υπάρχει χρόνος για να σκεφτεί κάποιος αν έχει μείνει έκπληκτος ή όχι, τότε σίγουρα δεν έχει μείνει. Η διάρκεια του συναισθήματος της έκπληξης είναι μικρή, με εξαίρεση την περίπτωση που την αρχική έκπληξη την ακολουθήσουν και άλλες καταστάσεις που προκαλούν επιπλέον εκπλήξεις. Η εξαφάνιση του συναισθήματος της έκπληξης είναι τόσο απότομη, όσο και η εμφάνιση του. Στο πρόσωπο το συναίσθημα της έκπληξης αποτυπώνεται με ανύψωση των φρυδιών, με μάτια ανοιχτά διάπλατα, με πτώση του σαγονιού και χωρισμό των χειλιών. Προφανώς σε κάποιες περιπτώσεις μπορεί να υπάρξει το συναίσθημα με κάποια από τις περιοχές να μην συμμετέχει, σε περιπτώσεις για παράδειγμα που γίνεται προσπάθεια συγκάλυψης του συναισθήματος ή όταν περισσότερα από ένα συναισθήματα λαμβάνουν χώρα ταυτόχρονα.

Το συναίσθημα του φόβου είναι αυτό που αισθάνεται κάποιος όταν βρίσκεται υπό την απειλή σωματικής ή ψυχολογικής ζημιάς. Η σωματική ζημιά μπορεί να είναι κάτι μικρό όπως ένας απλός εμβολιασμός ή κάτι πολύ πιο σοβαρό όπως απώλεια της ζωής από ένα τραύμα. Η ψυχολογική ζημιά, παρόμοια με την σωματική, έχει κλίμακα από κάτι μικρό έως κάτι πολύ σοβαρό, όπως ασήμαντες προσβολές, μέχρι και την απόρριψη της αγάπης που θέλει να δώσει ένα άτομο σε κάποιο άλλο. Η ζημιά μπορεί να είναι ταυτόχρονα σωματική και ψυχολογική αν υπάρξει μία κατάσταση, όπου η σωματική ζημιά προκαλέσει και ψυχολογική ζημιά. Στο πρόσωπο το συναίσθημα του φόβου αποτυπώνεται με ανύψωση των φρυδιών και την ένωσή τους, με τα μάτια να ανοίγουν διάπλατα, με το κάτω χέιλος να είναι σφιγμένο αλλά ταυτόχρονα και με τα δύο χείλια να τεντώνονται προς τα πίσω. Σε περιπτώσεις όπου ο φόβος είναι ακραίος κατηγοριοποιείται σαν τρόμος, όπου όλες οι παραπάνω εναλλαγές στα χαρακτηριστικά του προσώπου είναι πολύ πιο έντονες.

Το συναίσθημα της αηδίας είναι αυτό που αισθάνεται κάποιος όταν νοιώθει αποστροφή για κάτι. Αυτό το κάτι μπορεί να προκαλέσει αποστροφή με οπτική επαφή του, με την υφή του, με τον ήχο που κάνει, με την γεύση που έχει, με την μυρωδιά που εκπέμπει ή ακόμα και με την σκέψη του. Αυτό που κάποιος βρίσκει αηδιαστικό είναι εντελώς υποκειμενικό, που σημαίνει ότι κάποιος άλλος μπορεί να βρίσκει το ίδιο πράγμα ελκυστικό, αλλά το συναίσθημα της αηδίας και ο τρόπος έκφρασης του είναι καθολικός. Στο πρόσωπο το συναίσθημα της αηδίας αποτυπώνεται κυρίως στις περιοχές του στόματος και της μύτης, ενώ σε μικρότερο ποσοστό στις περιοχές των ματιών και των φρυδιών. Το άνω χέιλος

ανασηκώνεται, ενώ το κάτω χείλος είτε ανασηκώνεται είτε κατεβαίνει. Η μύτη κάνει έντονες ρυτίδες, τα κάτω βλέφαρα πάνε προς τα πάνω και τα φρύδια κατεβαίνουν προς τα κάτω.

Το συναίσθημα του θυμού είναι το πιο επικίνδυνο από όλα τα συναισθήματα. Όταν κάποιος είναι θυμωμένος υπάρχει φόβος σωματικής ζημιάς των παρευρισκόμενων. Υπάρχει κίνδυνος απώλειας ελέγχου των πράξεων ενός ανθρώπου που είναι θυμωμένος, πράγμα που κάνει τον θυμό ένα απρόβλεπτο συναίσθημα. Ο λόγος για τον οποίο ένας άνθρωπος βιώνει το συναίσθημα του θυμού είναι συνήθως οι πράξεις ή τα λόγια τρίτων που αφορούν τον άνθρωπο εν λόγο ή άτομα κοντινά σε αυτόν. Υπάρχει και αδικαιολόγητος θυμός που στην περίπτωση αυτή ο άνθρωπος που τον εκφράζει μπορεί να θεωρηθεί περιέργος ή ψυχικά ασταθής. Στο πρόσωπο το συναίσθημα του θυμού αποτυπώνεται σε όλες τις περιοχές του προσώπου. Τα φρύδια κατεβαίνουν προς τα κάτω και ενώνονται, τα βλέφαρα είναι σφιγμένα και τα μάτια κοιτάνε έντονα. Τα χείλια είτε πιέζονται σφιχτά μεταξύ τους, είτε είναι χωρισμένα σχηματίζοντας ένα τετράγωνο.

Το συναίσθημα της ευτυχίας είναι αυτό που όλοι θέλουν να βιώσουν. Ο κάθε άνθρωπος προσπαθεί να σχεδιάσει την ζωή του ώστε να βιώνει το συναίσθημα αυτό όσο πιο συχνά είναι δυνατόν. Σε σχέση με τα υπόλοιπα βασικά συναισθήματα που είναι όλα αρνητικά (εκτός από την έκπληξη που δεν είναι ούτε αρνητική ούτε θετική) και οι πιο πολλοί άνθρωποι δεν θέλουν να τα βιώνουν, το συναίσθημα της ευτυχίας το απολαμβάνουν. Υπάρχουν πολλά είδη ευτυχίας, όπως ευχαρίστηση-ηδονή, ενθουσιασμός-έξαψη, ανακούφιση και διαφορετικές εντάσεις ευτυχίας, από ήπια ευτυχία μέχρι έκσταση. Στο πρόσωπο το συναίσθημα της ευτυχίας αποτυπώνεται με τις γωνίες των χειλιών να τραβιούνται προς τα πίσω και προς τα πάνω, με το στόμα είτε κλειστό, είτε ανοικτό με τα δόντια εκτεθειμένα, συχνά με μία ρυτίδα να κατεβαίνει από την μύτη μέχρι τις άκρες των χειλιών, με τα μάγουλα ανασηκωμένα και τα κάτω βλέφαρα να σχηματίζουν ρυτίδες.

Το συναίσθημα της λύπης είναι αυτό που εκφράζεται όταν κάποιος υποφέρει σιωπηλά. Οτιδήποτε μπορεί να κάνει κάποιον λυπημένο, αλλά συνήθως το συναίσθημα της λύπης συνδέεται με την απώλεια. Απώλεια μέσω θανάτου ή απόρριψης από κάποιον αγαπημένο, απώλεια ευκαιριών ή ανταμοιβής λόγω προσωπικού λάθους ή περιστάσεων, απώλεια υλικών αγαθών, απώλεια της υγείας ενός προσώπου, μέσω αρρώστιας ή ατυχήματος. Η λύπη έχει συνήθως μεγαλύτερη διάρκεια από τα υπόλοιπα βασικά συναισθήματα και διαρκεί ώρες ή ακόμα πιο συχνά μέρες ανάλογα με το είδος και την σπουδαιότητα της απώλειας. Στο πρόσωπο το συναίσθημα της λύπης αποτυπώνεται με τις εσωτερικές γωνίες των φρυδιών δείχνουν προς τα πάνω, το δέρμα κάτω από τα φρύδια να τριγωνίζεται με την εσωτερική γωνία να δείχνει προς τα πάνω, η εσωτερική γωνία του άνω βλέφαρου σηκώνεται και οι γωνίες των χειλιών δείχνουν προς τα κάτω και τρέμουν.

1.2 Ιστορική αναδρομή

Η επιστημονική μελέτη των εκφράσεων του προσώπου και κατ' επέκταση των συναισθημάτων άρχισε με τον Κάρολο Δαρβίνο το 1872, με την δημοσίευση του με τίτλο "The Expressions of Emotions in Man and Animals", όπου υποστήριζε ότι κάποια συναισθήματα είναι καθολικά, όχι μόνο σε ανθρώπους, αλλά και σε κάποια ζώα. Ο Δαρβίνος έδωσε ένα «ερωτηματολόγιο» από 16 ερωτήσεις σε Άγγλους που ζούσαν ή ταξίδευαν σε 8 μέρη του κόσμου. (Αφρική, Αμερική, Αυστραλία, Βόρνεο, Κίνα, Ινδία, Μαλαισία και Νέα Ζηλανδία). Με βάση τις εκφράσεις των ντόπιων που παρατήρησαν, απάντησαν στις ερωτήσεις και ο Δαρβίνος βάσισε την θεωρία του για την καθολικότητα των συναισθημάτων πάνω σε αυτά τα αποτελέσματα. Όλοι έγραψαν ότι είδαν τις ίδιες εκφράσεις συναισθημάτων σε αυτούς τους τόπους με αυτές που ήδη ήξεραν από την Αγγλία. Αυτή η έρευνα όμως, δεν

ήταν αρκετή για να αποδειχτεί η καθολικότητα των συναισθημάτων για τα σημερινά δεδομένα. Το δείγμα του Δαρβίνου ήταν πολύ μικρό, οι απαντήσεις που πήρε από τους Άγγλους δεν ήταν εντελώς αξιόπιστες και ο τρόπος που είχαν διατυπωθεί οι ερωτήσεις, κατά κάποιο τρόπο καθοδηγούσαν προς την απάντηση που ήθελε να λάβει ο Δαρβίνος. [5]

Από τις αρχές της δεκαετίας του 70', οι Ekman και Friesen έκαναν εκτενείς έρευνες που αφορούσαν τις εκφράσεις του προσώπου, συλλέγοντας αποδείξεις για την καθολικότητα των συναισθημάτων που εκφράζουν χαρά, λύπη, θυμό, φόβο, αηδία και έκπληξη [6] [7]. Για να αποδείξουν την θεωρία της καθολικότητας των συναισθημάτων, πήγαν σε 21 χώρες με διαφορετικές κουλτούρες για να καταγράψουν τις εκφράσεις του προσώπου των ιθαγενών, μέσα από μια σειρά τεστ. Όταν ο Birdwhistell [8] υποστήριξε ότι η καθολικότητα των συναισθημάτων οφείλετε στην εξάπλωση κοινών πηγών, όπως ταινιών, περιοδικών και τηλεόρασης ο Ekman [5] πήγε στην Παπούα-Νέα Γουινέα για να μελετήσει την κουλτούρα του South Fore. Οι άνθρωποι αυτοί ήταν οπτικά απομονωμένοι. Δεν είχαν δει ποτέ τηλεόραση, φωτογραφία ή περιοδικό και είχαν ελάχιστη επαφή με ξένους. Ο Ekman είπε στον μεταφραστή του να τους πει μία μικρή ιστορία και μετά τους είπε να επιλέξουν μια εικόνα που πίστευαν ότι ταίριαζε με την ιστορία. Σε κάθε ιστορία κυριαρχούσε ένα από τα συναισθήματα που έδειχναν οι εικόνες. Οι άνθρωποι αυτοί που ήταν εντελώς απομονωμένοι επέλεξαν τα ίδια συναισθήματα που είχαν επιλέξει και οι άνθρωποι από τις άλλες 21 χώρες.

Οι εκφράσεις του προσώπου εξετάστηκαν και αναλύθηκαν κυρίως από ψυχολόγους, μέχρι που το 1978, οι Suwa et al. [9] επιχείρησαν για πρώτη φορά να φτιάξουν ένα αυτόματο σύστημα ανάλυσης εκφράσεων του προσώπου χρησιμοποιώντας μια ακολουθία εικόνων. Την δεκαετία του 90' η έρευνα πάνω στις εκφράσεις του προσώπου αναπτύχθηκε ακόμα περισσότερο και ωρίμασε χάρη στις προσπάθειες των Mase και Pentland [10]. Εξαιτίας των εκτενών εφαρμογών που είχε σε σχετικούς κλάδους, η αυτόματη αναγνώριση εκφράσεων προσέελκε όλο και μεγαλύτερη προσοχή και η πλειοψηφία των ερευνητών προσπαθούσαν να κατανοήσουν τεχνικές που ήταν βασισμένες σε εικόνες και ακολουθίες εικόνων (βίντεο). Κάποιοι χρησιμοποίησαν τα πλεονεκτήματα της αυτόματης αναγνώρισης των εκφράσεων του προσώπου για να αναπτύξουν διαφορετικούς αλγόριθμους που ήταν χρήσιμοι για ψυχολόγους, γιατρούς, καλλιτέχνες, ηθοποιούς, κτλ. Αργότερα οι εφαρμογές επεκτάθηκαν και στην ανάπτυξη της ρομποτικής, στα γραφικά υπολογιστών, την τεχνητή όραση, τα κινούμενα σχέδια [11] [12].

1.3 Εφαρμογές των αυτόματων συστημάτων αναγνώρισης συναισθημάτων

Η αναγνώριση των συναισθημάτων από αυτόματα συστήματα έχει πολλές εφαρμογές όπως αναφέρθηκε και παραπάνω σε διάφορους τομείς όπως στον τομέα της ιατρικής, της ηλεκτρονικής μάθησης, της παρακολούθησης ή επίβλεψης, της ψυχαγωγίας, του δικαίου ή της νομικής, του μάρκετινγκ και πολλών άλλων. Είναι σημαντικό να αναπτυχθεί ένα αξιόπιστο αυτόματο σύστημα για την αναγνώριση συναισθημάτων γιατί διαφορετικά η εφαρμογή τους στους παραπάνω τομείς δεν είναι εφικτή. Για παράδειγμα στον τομέα της ιατρικής, ένα αυτόματο σύστημα για την παρακολούθηση ηλικιωμένων που έχουν κάποιο σοβαρό πρόβλημα υγείας, πρέπει να αναγνωρίζει αξιόπιστα τα σημάδια κινδύνου που μπορεί να προβάλλονται στο πρόσωπο ή στις κινήσεις του ασθενή, όπως σημάδια πόνου ή φόβου στο πρόσωπο, απότομες, ασυνήθιστες κινήσεις κτλ.

Άλλες εφαρμογές στην ιατρική είναι η παρακολούθηση ασθενών κατά την διάρκεια φυσιοθεραπειών, την παρακολούθηση της συναισθηματικής κατάστασης του ασθενούς κατά

την διάρκεια συνεδρίων συμβουλευτικής ιατρικής ή ψυχανάλυσης και στην φροντίδα της υγείας για την καταγραφή των συναισθημάτων του ασθενή σχετικά με την θεραπεία που τους προτείνεται. Ειδικότερα σε περιπτώσεις που πάσχει ο ασθενής από κάποια γενετική ασθένεια όπως αυτισμό, μπορεί να αναπτυχθεί ένα σύστημα για την ερμηνεία των συναισθημάτων του ασθενή που δυσκολεύεται να εκφραστεί με σκοπό την θεραπεία του. Στον τομέα του μάρκετινγκ ένα αυτόματο σύστημα αναγνώρισης συναισθημάτων για να βελτιωθούν οι πωλήσεις ενός προϊόντος θα μπορούσε να καταγράφει την επιρροή που έχει μία διαφήμιση σε έναν καταναλωτή, κατά ποσό του κέντρισε το ενδιαφέρον και να προσαρμόζει της επόμενες διαφημίσεις ανάλογα, γιατί η συναισθηματική κατάσταση του καταναλωτή επηρεάζει σε μεγάλο βαθμό τις αγοραστικές του συνήθειες.

Στον τομέα της ηλεκτρονικής μάθησης ένα αυτόματο σύστημα αναγνώρισης συναισθημάτων μπορεί να διακρίνει την διάθεση ή ψυχολογική κατάσταση του μαθητή από τις εκφράσεις του και να προσαρμόζει τον τρόπο παρουσίασης της διαδικτυακής διδασκαλίας ώστε να γίνεται ποιο ενδιαφέρον για τον μαθητή που το διαβάζει. Η θετική διάθεση είναι σημαντική για την μέγιστη απόδοση στην απορρόφηση της γνώσης, το σύστημα διδασκαλίας γίνεται έτσι πιο διαδραστικό και αποδοτικό και με την χρήση ανάδρασης (feedback) από τους χρήστες μπορεί να βελτιώνεται ακόμα περισσότερο. Στον τομέα της ψυχαγωγίας μία εφαρμογή μουσικής θα μπορούσε να ανιχνεύει την διάθεση και τα συναισθήματα του χρήστη και να του προτείνει ή να παίζει την ανάλογη μουσική. Στον τομέα της νομικής ένα αυτόματο σύστημα για την ανίχνευση μικρό-εκφράσεων του προσώπου θα μπορούσε να εντοπίζει αν οι ισχυρισμοί ενός μάρτυρα ή κατηγορούμενου είναι αληθής ή ψευδής.

Στον τομέα της παρακολούθησης, αυτόματα συστήματα αναγνώρισης συναισθημάτων μέσα σε αυτοκίνητα θα μπορούν να αναγνωρίζουν την ψυχολογική κατάσταση του οδηγού, ώστε αν είναι θυμωμένος ή αν εκφράζει κάποιο άλλο συναίσθημα που μπορεί να προκαλέσει κίνδυνο στον εαυτό του ή σε άλλους οδηγούς, να ειδοποιούνται τα αυτοκίνητα γύρω του και να παρακινεί το σύστημα με μια ειδοποίηση τον οδηγό να ηρεμήσει. Μία άλλη εφαρμογή είναι στα ΑΤΜ, όπου θα γίνεται αναγνώριση συναισθημάτων του ατόμου που κάνει ανάληψη χρημάτων και αν εντοπίζεται φόβος στο πρόσωπο του, στην περίπτωση που κάποιος τον εξαναγκάζει ή τον εκβιάζει να βγάλει χρήματα από τον τραπεζικό του λογαριασμό, τότε να μην επιτρέπεται η ανάληψη χρημάτων ή να ειδοποιείται η τράπεζα του. Μια χρήσιμη εφαρμογή για ζευγάρια που έχουν σχέση εξ' αποστάσεως είναι η αναγνώριση της πραγματικής ψυχολογικής κατάστασης του συντρόφου και η ανίχνευση της στοργικότητας από τις εκφράσεις του προσώπου και την ανάλυση της χροιάς και του τόνου της φωνής.

2. Ανασκόπηση σχετικής βιβλιογραφίας

Οι Ekman και Friesen [13] ανέπτυξαν ένα σύστημα που περιγράφει όλες τις οπτικά διακριτές κινήσεις του προσώπου και το ονόμασαν “the facial action coding system” (FACS). Το FACS είναι σύστημα κωδικοποίησης προσανατολισμένο στην ανατομία και βασίζεται στον ορισμό των “action units” (AUs) του προσώπου που προκαλούν κίνηση. Κάθε action unit μπορεί να αντιπροσωπεύει έναν αριθμό μυών του προσώπου που όλοι μαζί παράγουν μία συγκεκριμένη κίνηση στο πρόσωπο. Κάποιοι μύς προκαλούν πάνω από μια κίνηση, οπότε διαφορετικά AU μπορεί να περιλαμβάνουν τον ίδιο μυ. Ένα τελικό σύνολο από 46 action units επιλέχθηκαν τα οποία είναι υπεύθυνα για τον έλεγχο των εκφράσεων του προσώπου και άλλα 12 για τον προσανατολισμό και την κατεύθυνση του βλέμματος. Οι Ekman et al. ανέπτυξαν και ένα λεξικό, ονόματι EMFACS, που περιέχει συνδυασμούς κάποιων action units με κάποιες κινήσεις για να σχηματίσουν ένα από τα 7 βασικά συναισθήματα [14].

Χρησιμοποιώντας το FACS ένας προγραμματιστής έχει την δυνατότητα να κωδικοποιήσει οποιαδήποτε ανατομική έκφραση και επειδή τα action units είναι ανεξάρτητα το ένα από το άλλο μπορούν να χρησιμοποιηθούν σε διαδικασία λήψης αποφάσεων για αναγνώριση των βασικών συναισθημάτων (Ekman et al. 2002), αναγνώριση καταστάσεων αντίληψης όπως διαφωνία και απορία (Cunningham et al. 2004), αναγνώριση ψυχολογικών καταστάσεων, όπως κατάθλιψη (Heller & Haynal, 1997) ή πόνος (Williams, 2002), αναγνώριση κοινωνικών σημάτων ή εμβλημάτων, όπως το κλείσιμο του ματιού και αναγνώριση ρυθμιστών (regulators), όπως ένα νεύμα ή ένα χαμόγελο. Άρα τα action units είναι κατάλληλα για χρήση σαν παράμετροι σε αυτόματη ανάλυση συμπεριφοράς του προσώπου, αφού με μικρό αριθμό action units μπορούν να περιγράφονται πολλές πιθανές εκφράσεις του προσώπου.

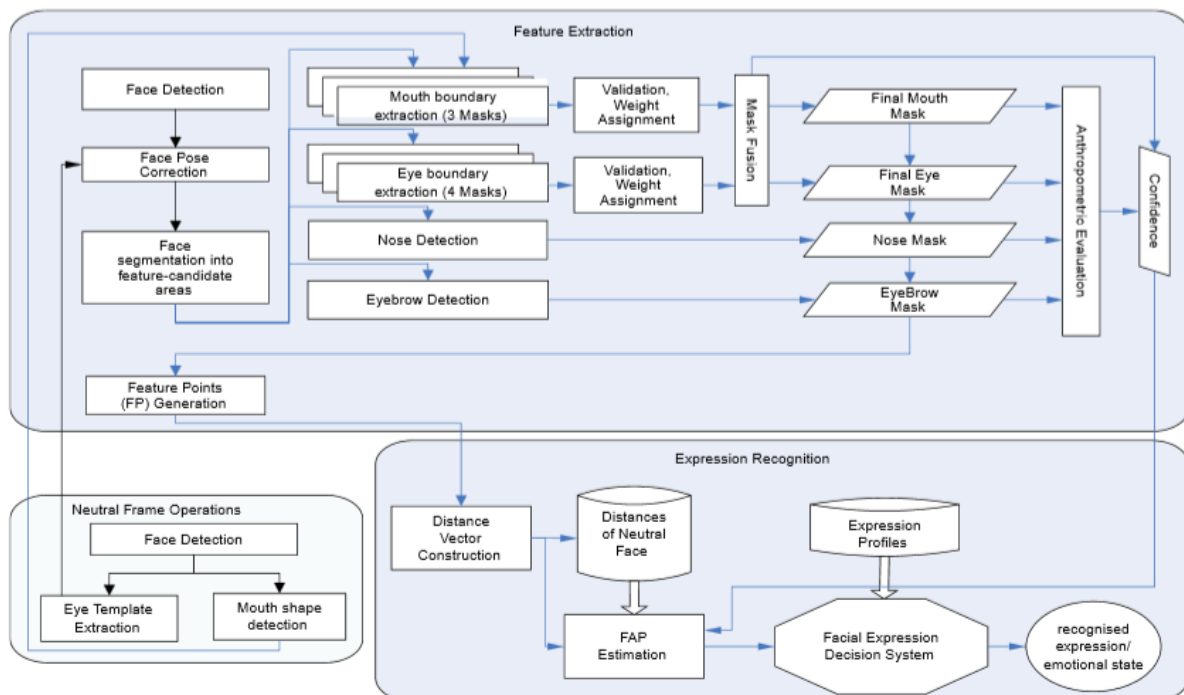
Το μοντέλο FACS χρησιμοποιήθηκε σαν πρότυπο για την παραγωγή του facial animation, που σημαίνει δημιουργία προσώπων μέσω υπολογιστή, με πρότυπο το ISO MPEG-4 [6]. Για την ακρίβεια το facial definition parameter set (FDP) και το facial animation parameter set (FAP) σχεδιάστηκαν στο MPEG-4 framework που επιτρέπει τον προσδιορισμό του σχήματος και της υφής του προσώπου, καθώς και της κίνησης του προσώπου όταν παράγονται εκφράσεις, συναισθήματα και ομιλία. Τα FAPs βασίζονται σε πολύ μικρές κινήσεις του προσώπου και συσχετίζονται με τις κινήσεις των μυών. Αντιπροσωπεύουν όλες τις δυνατές κινήσεις του προσώπου, όπως άνοιγμα και κλείσιμο των ματιών, σχηματισμό χαμόγελου ή σήκωμα των φρυδιών, που τους επιτρέπουν να παρουσιάσουν όλες σχεδόν τις φυσικές εκφράσεις.

Οι Owusu et al. [15] έφτιαξαν ένα νευρωνικό δίκτυο για αναγνώριση εκφράσεων του προσώπου βασισμένο σε AdaBoost (neural-AdaBoost based facial expression recognition system) με την χρήση και ενσωμάτωση του Viola-Jones για την αναγνώριση του προσώπου στην εικόνα. Η εικόνα με το αναγνωρισμένο πρόσωπο περνάει από υποδειγματοληψία (down-sampling) με Bessel transform για να μειωθεί το μέγεθος των χαρακτηριστικών που εξάγονται και να βελτιωθεί ο χρόνος εκτέλεσης. Γίνεται χρήση φίλτρων Gabor για την εξαγωγή χιλιάδων χαρακτηριστικών της εικόνας, τα οποία αντιπροσωπεύουν διάφορα μοτίβα παραμόρφωσης και ένας Ada-Boost αλγόριθμος δημιουργείται για την επιλογή μόνο μερικών εκατοντάδων χαρακτηριστικών για να βελτιωθεί ο χρόνος κατηγοριοποίησης. Τα επιλεγμένα χαρακτηριστικά εισάγονται σε ένα τεχνητό νευρωνικό δίκτυο με 3 στρώματα που εκπαιδεύεται με αλγόριθμο οπισθοδιάδοσης (back-propagation algorithm). Το σύστημα εκπαιδεύεται και δοκιμάζεται με τις βάσεις δεδομένων από εικόνες με εκφράσεις του

προσώπου JAFFE και Yale με μέσο όρο επιτυχούς αναγνώρισης 96.83% και 92.22% αντίστοιχα. Ο χρόνος εκτέλεσης για την αναγνώριση μίας 100*100 εικόνας είναι 14.5ms.

Οι Mohammadi et al. [16] στην δημοσίευσή τους με τίτλο “PCA-based dictionary building for accurate facial expression recognition via sparse representation” έφτιαξαν ένα λεξιλόγιο βασισμένο σε ανάλυση κύριων συνιστωσών (PCA - principal component analysis) για την αναγνώριση συναισθημάτων με την χρήση αραιής αναπαράστασης (sparse representation). Στην υλοποίησή τους κάνουν αφαίρεση μεταξύ της ουδέτερης έκφρασης ενός ατόμου και των εκφράσεων που αποτυπώνει το ίδιο άτομο για τα υπόλοιπα συναισθήματα. Μετά γίνεται ανάλυση κύριων συνιστωσών στο αποτέλεσμα της αφαίρεσης για να δημιουργηθεί ένα μοντέλο με τις διαφορές ανάμεσα σε κάθε έκφραση. Αυτές οι κύριες συνιστώσες χρησιμοποιούνται σαν στοιχεία στο λεξιλόγιο. Τα ποσοστά επιτυχούς αναγνώρισης που πήραν χρησιμοποιώντας την extended Cohn-Kanade (CK+) βάση δεδομένων με εκφράσεις του προσώπου [1], κάνοντας δοκιμές πάνω στην ίδια, στην MMI βάση δεδομένων με εκφράσεις του προσώπου (MMI facial expression database) [17] και στην Bosphorus 3D βάση προσώπων (Bosphorus 3D Face database) [18] ήταν 96.67%, 75.21% και 62.25% αντίστοιχα.

Οι Ioannou et al. [6] στην δημοσίευσή τους με τίτλο “Emotion recognition through facial expression analysis based on a neurofuzzy network” έκαναν χρήση παραμέτρων κίνησης του προσώπου (facial animation parameters FAPs) που έχουν οριστεί σύμφωνα με το πρότυπο ISO MPEG-4 και η εξαγωγή τους έχει γίνει από ένα εύρωστο σύστημα ανάλυσης χαρακτηριστικών του προσώπου, όπως φαίνεται στην Εικόνα 1 παρακάτω, για να δημιουργήσουν ένα neurofuzzy δίκτυο. Το neurofuzzy δίκτυο επιτρέπει την επιπλέον μάθηση και προσαρμογή των εκφράσεων συναισθήματος συγκεκριμένων χρηστών κάνοντας



ανάλυση με ομαδοποίηση (clustering) των νέων τιμών των παραμέτρων κίνησης του προσώπου. Το ποσοστό επιτυχούς αναγνώρισης που πέτυχαν με το σύστημα αυτό έφτασε το 78%.

Εικόνα 1: Το σύστημα εξαγωγής χαρακτηριστικών του προσώπου και ανάλυσης των εκφράσεων που δημιούργησαν οι Ioannou et al. [8]

Οι Manisha Satone και G.K. Kharate [19] στην δημοσίευση τους με τίτλο “Selection of Eigenvectors for Face Recognition” έφτιαξαν ένα σύστημα βασισμένο σε ανάλυση κύριων συνιστωσών (PCA) με την χρήση «κυματιδίων» (wavelets) για την εξαγωγή των καλύτερων ιδιοδιανυσμάτων από την περιοχή του προσώπου για την αναγνώριση συναισθημάτων. Η εξαγωγή για την επιλογή των σημαντικότερων ιδιοδιανυσμάτων έγινε με δύο διαφορετικές μεθόδους. Με την χρήση γενετικού αλγόριθμου και με την εντροπία των ιδιοδιανυσμάτων. Η χρήση μεθόδων για την επιλογή των σημαντικότερων ιδιοδιανυσμάτων γίνεται γιατί οι μεγαλύτερες ιδιοτιμές και τα ιδιοδιανύσματα που αντιστοιχούν σε αυτές σύμφωνα με τους Manisha και Kharate κωδικοποιούν ασήμαντη πληροφορία για την αναγνώριση συναισθημάτων όπως φωτισμό ή χαρακτηριστικά του προσώπου όπως μουστάκι, γένια, γυαλιά. Τα καλύτερα αποτελέσματα τους ήταν στην GTAV βάση δεδομένων [20] με ποσοστό επιτυχούς αναγνώρισης 96.73%, 98.91% και 97.82% για τα 20 μεγαλύτερα ιδιοδιανύσματα, για τα 20 ιδιοδιανύσματα που επιλέχθηκαν με γενετικό αλγόριθμο και για τα 20 ιδιοδιανύσματα που επιλέχθηκαν με εντροπία αντίστοιχα.

Η χρήση των τεχνητών νευρωνικών δικτύων (artificial neural networks) για την αναγνώριση των εκφράσεων του προσώπου αναπτύσσεται ραγδαία τα τελευταία χρόνια. Πιο συγκεκριμένα γίνεται η χρήση μίας νέας υλοποίησης τεχνητών νευρωνικών δικτύων, των νευρωνικών δικτύων συνέλιξης (Convolution neural networks) που αποτελούν ένα είδος βαθιών νευρωνικών δικτύων (deep neural networks). Η «βαθιά μάθηση» είναι μία απόπειρα μοντελοποίησης υψηλού επιπέδου αφαιρέσεων στα δεδομένα [21]. Σε ένα απλό νευρωνικό δίκτυο υπάρχουν λίγα στρώματα νευρώνων όπου όλες οι είσοδοι αποστέλλονται σε όλους του νευρώνες του επόμενου στρώματος, ενώ σε ένα βαθύ νευρωνικό δίκτυο υπάρχουν πολλά επόμενα στρώματα ανάμεσα στην αρχική είσοδο και στην τελική έξοδο. Τα νευρωνικά δίκτυα συνέλιξης είναι τα πιο ευρέως χρησιμοποιημένα για επεξεργασία οπτικών και άλλων δισδιάστατων δεδομένων όπως εικόνες προσώπων για αναγνώριση συναισθημάτων. Ένα νευρωνικό δίκτυο συνέλιξης αποτελείται από ένα ή περισσότερα στρώματα συνέλιξης με πλήρως συνδεδεμένους νευρώνες (όπως ένα «παραδοσιακό» τεχνητό νευρωνικό δίκτυο), «δεμένα» βάρη (tied weights) και στρώματα συγκέντρωσης (pooling layers). Σε σύγκριση με άλλες βαθιές αρχιτεκτονικές δικτύων τα νευρωνικά δίκτυα συνέλιξης προσφέρουν ανώτερα αποτελέσματα σε εφαρμογές.

Οι Shima Alizadeh και Azar Fazel [22] στην δημοσίευση τους με τίτλο “Convolutional Neural Networks for Facial Expression Recognition” ανέπτυξαν ένα νευρωνικό δίκτυο συνέλιξης (CNN) για την αναγνώριση εκφράσεων του προσώπου με στόχο την κατηγοριοποίηση των εικόνων σε μία από τις επτά βασικές κατηγορίες συναισθημάτων. Εκτός από την εισαγωγή ακατέργαστων εικόνων στον δίκτυο για την εκπαίδευση του, χρησιμοποίησαν χαρακτηριστικά του ιστογράμματος των προσανατολισμένων κλίσεων (Histogram of Oriented Gradients) των εικόνων σε συνδυασμό με τις ακατέργαστες εικόνες για καλύτερη εκπαίδευση. Το δίκτυο αποτελείται από M στρώματα συνέλιξης που έχουν επιπλέον spartial batch normalization, dropout, max-pooling και ένα ReLU (rectified linear unit) στρώμα μαζί με το στρώμα συνέλιξης. Η spartial batch normalization [23] όταν οι είσοδοι μετά την κανονικοποίηση (normalization) τους έχουν μηδενική ή μέση τιμή τότε το στρώμα «ζητάει» την αρχική τιμή που είχε η είσοδος πριν γίνει η κανονικοποίηση της. Η dropout [24] είναι μία τεχνική μείωσης της υπερπροσαρμογής του δικτύου, ενώ η max-pooling είναι διαδικασία υποδειγματοληψίας της εικόνας βασισμένη σε μικρά φίλτρα, ώστε να μειωθούν οι διαστάσεις της εικόνας και να εξαχθούν υποθέσεις για τα χαρακτηριστικά που δίνει η διαδικασία αυτή. Τέλος το ReLU είναι ένα στρώμα με μία συνάρτηση ενεργοποίησης $f(x)=\max(0,x)$ που αυξάνει τις μη γραμμικές ιδιότητες της συνάρτησης αποφάσεων (decision function) και συνολικά του δικτύου χωρίς να επηρεάζει την σωστή λειτουργία του στρώματος συνέλιξης [24]. Τα στρώματα συνέλιξης τα ακολουθούν N

πλήρως συνδεδεμένα στρώματα και τέλος ένα στρώμα συσχέτισης που υπολογίζει τα αποτελέσματα και την απώλεια της συνάρτησης softmax [25]. Το νευρωνικό δίκτυο συνέλιξης εκπαιδεύτηκε σε 29.000 εικόνες με ανάλυση 48x48 pixels, 4.000 εικόνες χρησιμοποιήθηκαν για την επικύρωση και άλλες 4.000 εικόνες για την δοκιμή της ακρίβειας πρόβλεψης του δικτύου. Τα αποτελέσματα για το βαθύ νευρωνικό δίκτυο συνέλιξης για το σετ επικύρωσης ήταν 65% και για το σετ δοκιμής 64%.

Οι Duncan et al. [26] στην δημοσίευση τους με τίτλο “Facial Emotion Recognition in Real Time” υλοποίησαν ένα νευρωνικό δίκτυο συνέλιξης για δυναμική αναγνώριση εκφράσεων σε πραγματικό χρόνο. Σε ένα ήδη υπάρχον νευρωνικό δίκτυο συνέλιξης που ήταν προ εκπαιδευμένο για αναγνώριση συναισθημάτων χρησιμοποίησαν μεταφορά μάθησης (transfer learning) στα πλήρως συνδεδεμένα στρωματά του. Ο ανιχνευτής προσώπου συνδεδεμένος σε ζωντανή ροή βίντεο στέλνει τα δεδομένα στο δίκτυο και αυτό, για όλα τα πρόσωπα που ανιχνεύονται, κατηγοριοποιεί τα συναισθήματα και βάζει πάνω στο πρόσωπο του κάθε ατόμου το κατάλληλο emoji. Η εκπαίδευση του δικτύου είχε ακρίβεια 90.7% και στην δοκιμή του δικτύου η προβλεπτική ακρίβεια έφτασε το 57.1%.

Οι Burkert et al. [27] στην δημοσίευση τους με τίτλο “DeXpression: Deep Convolutional Neural Network for Expression Recognition” δημιούργησαν ένα διαφορετικό νευρωνικό δίκτυο συνέλιξης που απολείπεται από 4 μέρη. Το πρώτο μέρος προεπεξεργάζεται αυτόματα τα δεδομένα εισόδου. Αυτό γίνεται με ένα στρώμα συνέλιξης, ένα στρώμα συγκέντρωσης (pooling layer) και ένα στρώμα κανονικοποίησης των δεδομένων. Τα δύο επόμενα μέρη αποτελούνται από δύο μπλοκ παράλληλης εξαγωγής χαρακτηριστικών FeatEx blocks (Parallel Feature Extraction Block) ένα για κάθε μέρος, με διαφορετικό αριθμό νευρώνων ανά μπλοκ και το τελευταίο μέρος από τον ταξινομητή. Στην extended Cohn-Kanade βάση δεδομένων το ποσοστό επιτυχούς αναγνώρισης έφτασε το 99.6% και στην MMI βάση δεδομένων το 98.63%.

Οι Gil Lavi και Tal Hassner [28] στην δημοσίευση τους με τίτλο “Emotion Recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns” παρουσιάζουν μία νέα μεθοδολογία για την υλοποίηση νευρωνικών δικτύων συνέλιξης, όπου χρησιμοποιούν τοπικό δυαδικό μοτίβο (Local Binary Pattern - LBP) στις εικόνες για την κωδικοποίηση τους ώστε να γίνει ευκολότερη η εκπαίδευση του δικτύου. Η παράμετρος της ακτίνας του LBP για κάθε εικόνα είναι διαφορετική και κάθε κωδικοποίηση μετατρέπει την ένταση της εικόνας σε μία τιμή από τουρ 256 LBP κωδικούς τιμών. Μετά οι LBP κωδικοί χαρτογραφούνται στον τρισδιάστατο χώρο χρησιμοποιώντας σαν «χάρτη» το αποτέλεσμα της εφαρμογής πολυδιάστατης κλιμάκωσης (Multi-Dimensional Scaling – MDS). Η αρχική εικόνα μαζί με τον χαρτογραφημένο κωδικό χρησιμοποιούνται για να εκπαιδευτούν πολλαπλά ξεχωριστά νευρωνικά δίκτυα συνέλιξης που έχουν ως σκοπό την πρόβλεψη μίας από τις επτά βασικές εκφράσεις. Μία τελική κατηγοριοποίηση εξασφαλίζεται από μία ζυγισμένη μέση τιμή (weighted average) που εφαρμόζεται πάνω στις εξόδους του συνόλου του δικτύου και έχει σαν αποτέλεσμα η κατηγορία πρόβλεψης του συναισθήματος να είναι εκείνη με την μέγιστη μέση πιθανότητα. Η μέθοδος τους έδειξε βελτίωση της τάξης του 15.36% σε σχέση με υλοποιήσεις παρόμοιων μεθόδων προκατόχων τους, με συνολικό ποσοστό αναγνώρισης 54.56%.

Στην εικόνα 2 παρακάτω δίνεται ένας πίνακας που έφτιαξαν οι Zeng et al. [11] στην δημοσίευση τους με τίτλο “A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions” που είναι πολύ χρήσιμος για να διατυπωθούν συγκεντρωτικά διάφοροι μέθοδοι για την αναγνώριση συναισθημάτων. Κάθε μέθοδος έχει την αναφορά της στην αντίστοιχη δημοσίευση.

References	Facial Feature	Classifier	Performance							
			exp	per	cues	rea	class	sub	samp	acc (%)
Ashraf et al. 07 [60]	AAM	SVM	S	I		N	2	21	?	Im: EER:19%
Bartlett et al. 04 [61]	Gabor wavelets	SVM+HMM	S	I		N	3 AUs	17	Vi: 230+ (OD)	Im/Vi: 75-98
Bartlett et al. 05 [62], [63]	Gabor wavelets	Adaboost SVM	S, P	I		Y	17 AUs	CK+EH: 119, RU:12	Im: 2568(CK+EH) 1689 (RU)	Im: 93.4(CK+EH), 90.5 (RU)
Cohen et al. 03 [64]	12 motion units	Tree-augmented DBN, HMM	P	D, I		Y	6	CH:5 CK:53	Vi: 30 (CH), 53 (CK)	Im: 66.53(CH), 73.22(CK) Vi: 58.63(CH)
Cohn et al. 04 [65]	shape models, Gabor wavelets	LDC	S	I	H	N	3 AUs	21	Im: 99 (OD)	Im: 76 (3-class)
El Kaliouby & Robinson 04 [66]	24 facial points	DBN	P	D	H	Y	6	30	Vi: 164 (OD)	Vi: 77.4
Fasel et al. 04 [67]	Gray-level intensity	NN	P	?	C	?	7	?	Im: 503 (CK)	Im: 38-68
Gunes & Piccardi 05 [68]	Shape features, optical flow	C4.5, Bayes-Net	P	?	B	N	8	FABO:4	Im: 206 (FABO)	Im: 80-100 (various fusion)
Ioannou et al. 05 [69]	FAPs	neurofuzzy network	S	I		N	3	?	Im: 984 (OD)	Im: 78
Ji et al. 06 [70]	Shape features	DBN	S	?	H,E,C	Y	2	8	Vi: 320min (OD)	Correlation coefficient: 95.3
Kapoor & Picard 05 [71]	Facial and head gesture	GP, SVM HMM, NN	S	?	E, P, T	?	2	8	Vi: 136 (OD)	Vi: 86
Kapoor et al. 07 [72]	Pixel difference of mouth region	Same as in [73]	S	I	E P S T M	?	2	24	Vi: 24 (OD)	Vi: 79.17 Baseline: 58
Lee & Elgammal [73]	Pixel intensity of face region	decomposable model	P	I		N	6	CK: 8 OD: 16	Vi: 48 (CK), 80 (OD)	Vi: 39.58 Im: 61.85
Littlewort et al. 07 [74]	Gabor wavelets	Adaboost SVM	S	I		Y	2	26	Vi: 312	Vi: 72
Lucey et al. 07 [75]	AAM	SVM	S,P	I		N	AUs: CK: 15 OD: 4	CK: 100 OD: ?	?	Im: 95 (CK) with 16.66% FAR, 70.47 (OD)
Pantic & Patras 06 [76]	Facial profile points	Rule-based	P	I		N	27 AUs	MMI: 19	Vi: 119 (MMI)	Vi: 86.3
Pantic & Rothkrantz 04 [77]	frontal and profile facial points	Rule-based	P	I		N	32 AUs	MMI: 25	Im: 454 (MMI)	Im: 86
Pantic & Rothkrantz [78]	same as in [103]	Rule-based, case-based	P	I	U	N	9	MMI: 8	Im: 196 (MMI)	Im: 83
Sebe et al. 04 [79]	12 motion units	kNN	S	I			4	CK: 53 SD: 28	Vi: ? (SD), 212+ (CK)	Im: 93 (CK) 95 (SD)
Tong et al. 07 [80]	Gabor wavelets	Adaboost, DBN	P	I		?	14 AUs	CK: 100 OD: 10	Im: 14000 (CK+OD)	Vi: 93.2 (OD), 93.3(CK)
Valstar et al. 04 [81]	Motion history images	SNoW kNN	P	I		N	15 AUs	MMI: 19 CK: 100	Vi: 344 (CK), 253 (MMI)	Vi: 61 (MMI) 68 (CK)
Valstar et al. 06 [82]	8 facial points	gentle boost, SVM	S, P	I		N	2	MMI:27 CK: 32 CD: 65	Vi: 60(MMI) 60(CK),70(OD)	Vi: 90.7
Valstar et al. 07 [83]	20 facial points	GentleSVM-sigmoid	S, P	I	H, B	N	2	MMI: 52	Vi: 100 (P), 102 (S)	Vi: 94%
Wang & Ahuja 03 [84]	Shape and gray-level texture	NN with HOSVD	S	?		?	7	14	Im: 110 (OD)	Im: 84.58
Wang et al. 06 [85]	3D surface labels	LDA	P	I		N	6	BU: 60	Im: 720 (BU)	Im: 83.6
Wen & Huang 03 [86]	Geometric, ratio-image	Exemplars with GMM	P	I		N	4	CK: 47	Im: 2981 (CK)	Im:75.37
Whitehill & Omlin 06 [87]	Haar features	Adaboost	P	I		Y	11 AUs	?	Im: 580 (OD)	Im: 92.35
Yeasin et al. 06 [88]	Pixel intensity of face	kNN + HMM	P, S			N	6	CK: 97 OD:21	Vi: 488 (CK) 108 (OD)	Vi: 90.7 (CK) 72-82 (OD)
Zeng et al. 06 [89]	Texture with LPP	SVDD	S	D		N	2	AAI: 2	Female:7857 Male: 5230	Im: 79(male), 87(female)

Εικόνα 2: Ο πίνακας των Zhihong Zeng et al. για την αναγνώριση συναισθημάτων από συστήματα που βασίζονται σε οπτικά μέσα όπως εικόνες και βίντεο για την υλοποίησή τους [9]. Στην στήλη exp το S σημαίνει αυθόρμητες (spontaneous) εκφράσεις και το P σημαίνει ποζαρισμένες (posed) εκφράσεις. Η στήλη per σημαίνει ανά άτομο (per person) και οι τιμές I και D σημαίνουν ανεξάρτητο (independent) και εξαρτώμενο (dependent). Η στήλη cues προσδιορίζει αν έχουν χρησιμοποιηθεί και άλλα χαρακτηριστικά εκτός του προσώπου. Αν έχει γίνει χρήση του κεφαλιού τότε H (Head), του σώματος τότε B (Body), των ματιών τότε E (Eye), του δέρματος τότε S (Skin), της στάσης του σώματος τότε P (Posture), των κλάσεων που έχουν οριστεί από τον χρήστη τότε U (User-defined classes) και άλλου πλαισίου τότε C (otherContext). Η στήλη rea με ναι (Y) ή όχι (N) προσδιορίζει εάν η μέθοδος δουλεύει σε πραγματικό χρόνο. Η στήλη class προσδιορίζει τον αριθμό των κλάσεων, η στήλη sub τον αριθμό των υποκειμένων, η

στήλη `samp` το μέγεθος του δείγματος και τέλος η στήλη `acc` την ακρίβεια της κάθε μεθόδου είτε σε στατικές εικόνες (`Im`) είτε σε βίντεο (`Vi`).

3. Μέθοδοι και βάση δεδομένων

Στο πλαίσιο της συγκεκριμένης εργασίας επιχειρείται η επιτυχής αναγνώριση των έξι βασικών συναισθημάτων (έκπληξη, φόβος, λύπη, θυμός, χαρά και αηδία) και της ουδέτερης έκφρασης. Μέσα από την βιβλιογραφία συναντήσαμε κάποιες μεθόδους που θα προσπαθήσουμε να υλοποιήσουμε με δικό μας τρόπο. Σε όλες τις προσπάθειες θα γίνει χρήση του προγράμματος MATLAB του οποίου το όνομα προέρχεται από τις αγγλικές λέξεις **matrix laboratory** και σημαίνει εργαστήριο πινάκων. Η βάση δεδομένων που χρησιμοποιήθηκε για όλες τις δοκιμές στις μεθόδους που υλοποιήθηκαν ήταν η Cohn-Kanade [1], που περιέχει εικόνες για τα έξι βασικά συναισθήματα μαζί με την ουδέτερη έκφραση.

3.1 Εισαγωγή στο περιβάλλον χρήσης

Το MATLAB είναι ένα περιβάλλον πολλαπλών χρήσεων για σύνθετους αριθμητικούς υπολογισμούς και πλησιάζει μια τέταρτης γενιάς γλώσσα προγραμματισμού. Επιτρέπει τον χειρισμό πινάκων, την κατασκευή διαγραμμάτων συναρτήσεων και δεδομένων, την δημιουργία αλγορίθμων, την δημιουργία διεπαφών χρήστη και την διασύνδεση με άλλα προγράμματα που είναι γραμμένα σε άλλες γλώσσες προγραμματισμού όπως C, C++, C#, Java, Fortran και Python [29]. Το MATLAB διαθέτει «εργαλειοθήκες» (toolboxes) όπου παρέχουν τα κατάλληλα «εργαλεία» (tools) για κάθε προγραμματιστική δουλειά που θέλει κάποιος να κάνει. Για την υλοποίηση αυτής της πτυχιακής εργασίας χρησιμοποιήθηκαν συναρτήσεις από δύο μεγάλες εργαλειοθήκες. Από την εργαλειοθήκη για επεξεργασία εικόνας (Image Processing Toolbox) [30] και από την εργαλειοθήκη για νευρωνικά δίκτυα (Neural Network Toolbox) [31] [32] [33].

Η εργαλειοθήκη για επεξεργασία εικόνας παρέχει μία ολοκληρωμένη σειρά αναφοράς τυποποιημένων αλγορίθμων, λειτουργίες και εφαρμογές για την επεξεργασία εικόνας, ανάλυση, απεικόνιση, και ανάπτυξη αλγορίθμων. Επιτρέπει την εκτέλεση ενεργειών όπως αλλαγή ανάλυσης της εικόνας, κατάτμηση της εικόνας, βελτίωση της εικόνας, μείωση του θορύβου στην εικόνα, γεωμετρικούς μετασχηματισμούς και την καταγραφή της εικόνας. Υποστηρίζει ένα μεγάλο εύρος διαφορετικών τύπων εικόνων, όπως εικόνες υψηλής δυναμικής εμβέλειας, εικόνες με ανάλυση gigapixel και τομογραφικές εικόνες. Συναρτήσεις απεικόνισης και εφαρμογές δίνουν την δυνατότητα εξερεύνησης εικόνων και βίντεο, εξέτασης περιοχών εικονοστοιχείων, προσαρμογής χρώματος και αντίθεσης, δημιουργίας περιγραμμάτων ή ιστογραμμάτων και τον χειρισμό περιοχών ενδιαφέροντος.

Πολλές από τις συναρτήσεις που θα γίνουν χρήση για την επεξεργασία των εικόνων της βάσης δεδομένων όπως την αυξομείωση του μεγέθους τους, την περιστροφή τους, την αποκοπή τους, την παρουσίαση τους σε φιγούρα, την αφαίρεση της μέσης τιμής τους και την εφαρμογή φίλτρων Gabor πάνω τους είναι μέρος της εργαλειοθήκης αυτής. Κάποιες από τις βασικές λειτουργίες της εργαλειοθήκης είναι η ανάγνωση, η εγγραφή, η απεικόνιση και η εξαγωγή εικόνων, όπως και οι γεωμετρικές μετατροπές, ο σχεδιασμός και η εφαρμογή γραμμικών φίλτρων σε εικόνες.

Η δεύτερη εργαλειοθήκη είναι αυτή των τεχνητών νευρωνικών δικτύων. Η εργαλειοθήκη αυτή παρέχει αλγόριθμους, λειτουργίες και εφαρμογές για την δημιουργία, την εκπαίδευση, την απεικόνιση και την προσομοίωση νευρωνικών δικτύων. Δίνει την δυνατότητα για την ταξινόμηση (classification), την παλινδρόμηση (regression), την ομαδοποίηση (clustering), την μείωση της διάστασης (dimensionality reduction), την δυναμική μοντελοποίηση συστήματος (dynamic system modeling), την πρόβλεψη χρονοσειρών (time-series forecasting) και τον έλεγχο (control) του δικτύου. Σημαντικά χαρακτηριστικά της εργαλειοθήκης είναι ότι παρέχει διαγράμματα, όπως για την εκπαίδευση και το λάθος πρόβλεψης, πίνακες σύγχυσης που βοηθούν στην καλύτερη κατανόηση της κατηγοριοποίησης που έχει γίνει και κώδικα για την ευκολότερη δημιουργία και προσαρμογή του δικτύου.

3.2 Βάση δεδομένων για εικόνες και βίντεο

Αρχικά πρέπει να εξεταστεί η βάση δεδομένων από εικόνες που θα χρησιμοποιηθεί, η οποία είναι η Cohn-Kanade AU-Coded Facial Expression Database [1]. Η συγκεκριμένη βάση επιλέχθηκε γιατί περιέχει ένα ικανοποιητικό αριθμό εικόνων από τις επτά εκφράσεις για την εκπαίδευση και την δοκιμή του συστήματος που θα σχεδιαστεί. Είναι ευρέως χρησιμοποιημένη και αναγνωρισμένη για εφαρμογές που κάνουν αναγνώριση εκφράσεων. Περιέχει 486 ακολουθίες εικόνων (Image Sequences) μεγέθους 640x490 pixels από 97 διαφορετικά άτομα, όπου κάθε ακολουθία αρχίζει από μία ουδέτερη έκφραση (Neutral Expression) και καταλήγει σε μία έντονη/κορυφωμένη έκφραση (Peak Expression). Αυτή η ακολουθία μπορεί να έχει μέγεθος από 10 έως 20 στιγμιότυπα (frames). Από αυτές τις ακολουθίες τα χρήσιμα για εμάς στιγμιότυπα είναι τα πρώτα που δείχνουν την ουδέτερη έκφραση και τα τελευταία που δείχνουν την έντονη έκφραση για καθένα από τα έξι βασικά συναισθήματα. Ο λόγος που τα ενδιάμεσα στιγμιότυπα απορρίπτονται είναι γιατί δείχνουν κάτι ενδιάμεσο σε ουδέτερη έκφραση και σε ένα από τα συναισθήματα, άρα θα είναι δύσκολο και ίσως αδύνατο για το σύστημα να τα αναγνωρίσει. Οπότε έγινε μία διαλογή των στιγμιότυπων που είναι κατάλληλα με αποτέλεσμα να μειωθούν από 10.000 στιγμιότυπα σε 2.418.

Το «πρόβλημα» που εμφανίζει αυτή η βάση δεδομένων είναι ότι ο αριθμός των στιγμιότυπων ανά συναίσθημα δεν είναι ο ίδιος για όλα τα συναισθήματα. Για την έκπληξη υπάρχουν 377 στιγμιότυπα από 78 άτομα, ποσοστό 15.59% της βάσης, για τη λύπη 308 στιγμιότυπα από 59 άτομα, ποσοστό 12.73% της βάσης, για τον φόβο 299 στιγμιότυπα από 58 άτομα, ποσοστό 12.36% για τον θυμό 243 στιγμιότυπα από 48 άτομα, ποσοστό 10.04%, για την αηδία 281 στιγμιότυπα από 60 άτομα, ποσοστό 11.62%, για την ευτυχία 440 στιγμιότυπα από 80 άτομα, ποσοστό 18.19% και για την ουδέτερη έκφραση 470 στιγμιότυπα από 91 άτομα, ποσοστό 19.43%, όπως φαίνεται και στον Πίνακα 1. Αυτό είναι λογικό γιατί η βάση δεν έχει ακολουθίες από όλα τα άτομα για κάθε συναίσθημα, καθώς κάποια συναισθήματα είναι πιο δύσκολα να εκφραστούν. Τα πιο κοινά συναισθήματα, που έχουν το μεγαλύτερο ποσοστό μέσα στην βάση είναι η έκπληξη και η χαρά. Η ουδέτερη έκφραση έχει επίσης μεγάλο ποσοστό γιατί όλα τα άτομα αρχίζουν από την ίδια ουδέτερη έκφραση, ανεξάρτητα από το συναίσθημα που εκφράζουν. Αυτή η ανομοιότητα των δεδομένων, θεωρητικά, μπορεί να επηρεάσει αρνητικά τα αποτελέσματα.

Συναίσθημα	Έκπληξη	Λύπη	Φόβο	Θυμό	Αηδία	Ευτυχία	Ουδέτερο
Στιγμιότυπα	377	308	299	243	281	440	470
Άτομα	78	59	58	48	60	80	91
Ποσοστό	15.59%	12.73%	12.36%	10.04%	11.62%	18.19%	19.43%

Πίνακας 1: Κατανομή εικόνων της βάσης ανά συναίσθημα.

Ο στόχος είναι να κατασκευαστεί ένα σύστημα που θα μπορεί να αναγνωρίζει ποιο συναίσθημα εκφράζεται σε μία εικόνα αυτόματα. Ο πρώτος τρόπος υλοποίησης ενός τέτοιου συστήματος, βασίζεται στην μέθοδο των Jeemoni Kalita και Karen Das [2] με την χρήση ιδιοδιανυσμάτων και ευκλείδειας απόστασης που θα αναλυθεί στο κεφάλαιο 4, ενώ ο δεύτερος τρόπος κάνει γενική χρήση τεχνητών νευρωνικών δικτύων που χρησιμοποιούνται ευρέως για τέτοιες εφαρμογές και θα αναλυθεί στο κεφάλαιο 5.

4. Μέθοδος 1: Ιδιοδιανύσματα, Ιδιοτιμές και Ευκλείδεια απόσταση

Ο πρώτος τρόπος υλοποίησης του αυτόματου συστήματος είναι αυτός των Jeemoni Kalita και Karen Das από την δημοσίευση τους «Recognition of Facial Expression Using Eigenvector Based Distributed Features and Euclidean Distance Based Decision Making Technique» [2]. Όπως αναφέρεται και στον τίτλο ο τρόπος που γίνεται η αναγνώριση των συναισθημάτων είναι με την χρήση χαρακτηριστικών που βασίζονται σε ιδιοδιανύσματα και την χρήση της Ευκλείδειας απόστασης σαν τεχνική λήψης αποφάσεων.

4.1 Θεωρητικό υπόβαθρο

Στην υλοποίηση οι Jeemoni Kalita και Karen Das [2] χρησιμοποιούν μια βάση δεδομένων από εικόνες που έχουν αποκτήσει μέσω μιας ψηφιακής κάμερας, οι οποίες έχουν άγνωστη ανάλυση και αριθμό, καθώς δεν αναφέρεται στην δημοσίευση. Οργανώνουν το dataset τους σε 6 πίνακες, έναν ανά συναίσθημα και 5 περιοχές ανά συναίσθημα. Πιο συγκεκριμένα από κάθε εικόνα παίρνουν 5 περιοχές οι οποίες είναι: αριστερό μάτι, δεξί μάτι, μύτη, στόμα και στόμα με μύτη μαζί. Για τα μάτια οι διαστάσεις είναι 40 επί 40 pixels, για την μύτη 70 επί 60 pixels, για το στόμα 60 επί 90 pixels και για το στόμα με την μύτη 110 επί 95 pixels. Μετά κάθε περιοχή εισάγεται σε ένα πίνακα $Y \times P$ (με $P = N \times M$, όπου N και M είναι οι διαστάσεις της περιοχής και Y ο συνολικός αριθμός εικόνων-περιοχών που θα εισαχθούν), οπότε έχουμε 5 πίνακες, έναν για κάθε περιοχή ανά συναίσθημα. Μετά υπολογίζουν την μέση τιμή του κάθε πίνακα και την αφαιρούν από κάθε διάσταση για κάθε πίνακα αντίστοιχα. Έτσι παίρνουν νέους πίνακες, οι οποίοι είναι κανονικοποιημένοι (normalized), δηλαδή δεν υπάρχουν πολύ ακραίες τιμές.

Μέσος όρος ή αλλιώς δειγματική μέση τιμή ενός συνόλου παρατηρήσεων αποτελεί το σπουδαιότερο και χρησιμότερο μέτρο της στατιστικής και είναι ένα μέτρο θέσης, δηλαδή δείχνει σχετικά τις θέσεις των αριθμών στους οποίους αναφέρεται. Γενικά, ορίζεται ως το άθροισμα των παρατηρήσεων δια του πλήθους αυτών. Είναι δηλαδή η μαθηματική πράξη ανεύρεσης της «μέσης απόστασης» ανάμεσα σε δύο ή περισσότερους αριθμούς [34]. Στην περίπτωση αυτή, η μέση τιμή αν υπάρχουν N στήλες θα είναι μία γραμμή $N \times 1$ όπου θα περιέχει τον μέσο όρο όλων των pixel για κάθε εικόνα στην αντίστοιχη θέση. Η τελική μέση τιμή θα είναι η μέση τιμή αυτού του διανύσματος. Μετά πρέπει να αφαιρεθεί αυτή η τιμή από κάθε στοιχείο του πίνακα. Αυτό γίνεται για να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα των δεδομένων.

Για να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα πρέπει πρώτα να βρεθεί ο πίνακας συνδιακύμανσης (covariance matrix). Στη θεωρία των πιθανοτήτων και τη στατιστική ο πίνακας συνδιακύμανσης (ή πίνακας διασποράς, ή πίνακας διακύμανσης-συνδιακύμανσης) είναι ένας πίνακας του οποίου το στοιχείο στην θέση i, j είναι η συνδιακύμανση ανάμεσα στο i^{th} και j^{th} στοιχείο ενός τυχαίου διανύσματος. Η συνδιακύμανση s_{XY} ή $\text{Cov}(X,Y)$ είναι μέτρο του βαθμού συσχέτισης δύο μεταβλητών X και Y . Ένα τυχαίο διάνυσμα είναι μία τυχαία μεταβλητή με πολλαπλές διαστάσεις. Κάθε στοιχείο του τυχαίου διανύσματος είναι μια τυχαία βαθμωτή μεταβλητή. Κάθε στοιχείο έχει, είτε έναν πεπερασμένο αριθμό από παρατηρούμενες εμπειρικές τιμές, είτε μη πεπερασμένο αριθμό από πιθανές τιμές [35]. Η εύρεση του πίνακα συνδιακύμανσης είναι αναγκαία για την εκτέλεση ορθογώνιας αποσύνθεσης (orthogonal decomposition), που λέγεται και ανάλυση κυρίων συνιστωσών (principal components analysis - PCA) και έχει ως αποτέλεσμα την εύρεση των

ιδιοδιανυσμάτων που αντιστοιχούν στις κύριες συνιστώσες και των ιδιοτιμών στην διακύμανση που εξηγούνται από τις κύριες συνιστώσες [36].

Ο ορισμός για τις ιδιοτιμές και τα ιδιοδιανύσματα προσδιορίζει ότι, αν ο A είναι ένας $n \times n$ πίνακας, τότε ένα μη μηδενικό διάνυσμα x στον R^n ονομάζεται ιδιοδιάνυσμα του A αν το Ax είναι βαθμωτό πολλαπλάσιο του x , δηλαδή αν $Ax = \lambda x$ για κάποιο βαθμωτό λ . Το βαθμωτό λ ονομάζεται ιδιοτιμή του A και λέμε ότι το x είναι ένα ιδιοδιάνυσμα του A που αντιστοιχεί στην λ [37]. Πιο απλά ένα ιδιοδιάνυσμα ενός τετραγωνικού πίνακα A είναι ένα μη μηδενικό διάνυσμα v που, όταν πολλαπλασιαστεί με τον A , ισούται με το αρχικό διάνυσμα, πολλαπλασιασμένο με έναν αριθμό λ , έτσι ώστε: $Av = \lambda v$. Ο αριθμός λ ονομάζεται ιδιοτιμή του A που αντιστοιχεί στο v [37].

Τα ιδιοδιανύσματα που συσχετίζονται με τις μεγαλύτερες ιδιοτιμές παρουσιάζουν τη μεγαλύτερη διακύμανση (ή ποικιλότητα ή διαφορετικότητα) στην εικόνα ενώ οι μικρότερες ιδιοτιμές τη μικρότερη διακύμανση. Από όλα τα ιδιοδιανύσματα επιλέγονται τα πέντε πιο σημαντικά (αυτά που συσχετίζονται με τις πέντε μεγαλύτερες ιδιοτιμές) για τις πέντε περιοχές ανά συναίσθημα. Δηλαδή, 25 ιδιοδιανύσματα για κάθε συναίσθημα και 25 για κάθε εικόνα που θα εισαχθεί για αναγνώριση. Σύμφωνα με τους J. Kalita και K. Das [2] πέντε μόνο ιδιοδιανύσματα είναι αρκετά για την επιτυχή αναγνώριση, αν και οι απόψεις δίστανται. Οι Matthew Turk και Alex Pentland [38] υποστηρίζουν ότι τα πρώτα ιδιοδιανύσματα που συσχετίζονται με τις μεγαλύτερες ιδιοτιμές περιέχουν θόρυβο και δεν είναι κατάλληλα για αναγνώριση, για αυτόν το λόγο ο αριθμός των ιδιοδιανυσμάτων που χρησιμοποιούν είναι πολύ μεγαλύτερος, μπορεί να φτάσει και τα εκατό. Χρησιμοποιώντας την μικρότερη ευκλείδεια απόσταση σαν classifier ανάμεσα στην δοκιμαστική εικόνα και στα 6 συναισθήματα υλοποιείται ένα σύστημα με «hits» όπου όποιο συναίσθημα έχει τα πιο πολλά «hit», δηλαδή τις πιο πολλές μικρότερες αποστάσεις, «κερδίζει» και η δοκιμαστική εικόνα επισημαίνεται σαν αυτό το συναίσθημα.

Ευκλείδεια απόσταση ή Ευκλείδεια μετρική είναι η απόσταση μεταξύ δύο σημείων στον Ευκλείδειο χώρο. Ο Ευκλείδειος χώρος μπορεί να είναι ο δισδιάστατος χώρος, ο τρισδιάστατος και τα λοιπά. Ο όρος «Ευκλείδειος» χρησιμοποιείται για να ξεχωρίσει αυτούς τους χώρους από άλλους τύπους χώρων που υπάρχουν στην σύγχρονη γεωμετρία. Η ευκλείδεια απόσταση μεταξύ δύο σημείων στον δισδιάστατο χώρο είναι το μήκος της ευθείας γραμμής που τα συνδέει. Σε καρτεσιανές συντεταγμένες για δύο σημεία σε ένα ευκλείδειο n -διάστατο χώρο η απόσταση μεταξύ τους δίνεται από το Πυθαγόρειο θεώρημα. Η νόρμα που συσχετίζεται με την ευκλείδεια απόσταση λέγεται ευκλείδεια νόρμα ή δεύτερη νόρμα (L^2 norm) [39] [40]. Όταν βρεθεί η Ευκλείδεια απόσταση ανάμεσα στα ιδιοδιανύσματα χρησιμοποιείται η μικρότερη απόσταση ανάμεσα στα 7 συναισθήματα και στην δοκιμαστική εικόνα, γιατί αυτή προσδιορίζει ότι τα δυο ιδιοδιανύσματα που συγκρίθηκαν έχουν τα περισσότερα κοινά χαρακτηριστικά.

Στην υλοποίηση της εργασίας δοκιμάστηκε και μία εναλλακτική προσέγγιση εκτός από αυτή των J. Kalita και K. Das [2], στην οποία αντί να γίνουν όλες οι περιοχές στήλες και να μουν σε έναν πίνακα, έμειναν ως έχουν και εκτελέστηκαν τα ίδια βήματα. Αυτό σημαίνει ότι α) υπολογίστηκε η μέση τιμή της εικόνας, β) η μέση αυτή τιμή αφαιρέθηκε από την εικόνα, γ) υπολογίστηκε στη συνέχεια ο πίνακας συνδιακύμανσης (covariance matrix) και δ) στη συνέχεια υπολογίστηκαν τα 5 ιδιοδιανύσματα για τις 5 περιοχές, αλλά αυτή τη φορά κάθε εικόνα ξεχωριστά. Στο τέλος έγιναν πράξεις για να βγει ένας μέσος όρος από 25 ιδιοδιανύσματα που έχουν τον ίδιο ρόλο, όπως και στον πρώτο τρόπο, για να βρεθεί η μικρότερη ευκλείδεια απόσταση σε σχέση με την δοκιμαστική εικόνα.

4.2 Σχεδιασμός υλοποίησης – Αλγόριθμος των δύο υλοποιήσεων

Στον κόσμο της πληροφορικής και του προγραμματισμού είναι αναγκαίο, πριν γίνει η οποιαδήποτε απόπειρα καταγραφής κώδικα, να υπάρχει ένας αλγόριθμος για την μέθοδο που θα υλοποιηθεί. Ως αλγόριθμος ορίζεται μια πεπερασμένη σειρά ενεργειών, αυστηρά καθορισμένων και εκτελέσιμων σε πεπερασμένο χρόνο, που στοχεύουν στην επίλυση ενός προβλήματος. Πιο απλά, αλγόριθμος ονομάζεται μία σειρά από εντολές που έχουν αρχή και τέλος, είναι σαφείς και ως σκοπό έχουν την επίλυση κάποιου προβλήματος. Ο σχεδιασμός του συστήματος «στο χαρτί» γίνεται, ώστε να κατανοηθεί ευκολότερα η μέθοδος και να αποφευχθούν λογικά λάθη, πριν αρχίσει η υλοποίηση του κώδικα στο matlab.

Για τον πρώτο τρόπο, με την χρήση ιδιοδιανυσμάτων και ευκλείδειας απόστασης, σαν πρώτο βήμα γίνεται ο διαχωρισμός της βάσης δεδομένων σε επτά κατηγορίες συναισθημάτων και μία κατηγορία από δοκιμαστικές εικόνες, που περιέχει ένα ποσοστό εικόνων από κάθε συναίσθημα (βήμα 1). Στην συνέχεια οι εικόνες εισάγονται μία προς μία στην επιφάνεια εργασίας του matlab για να αρχίσει η επεξεργασία τους. Η εισαγωγή τους γίνεται μέσω interface, δηλαδή με εντολή του matlab θα ανοίγει παράθυρο για να γίνει η επιλογή ενός φακέλου, όπου αυτός ο φάκελος θα περιέχει όλες τις εικόνες ενός συναισθήματος (βήμα 2). Το πρόγραμμα θα πρέπει να είναι σχεδιασμένο κατάλληλα, ώστε να διαβάζει μόνο αρχεία εικόνων, όπως png, tif, jpg κτλ. (βήμα 3).

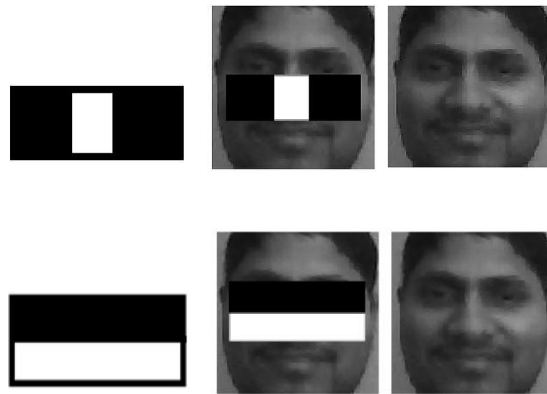
Για να ανιχνευτεί η περιοχή του προσώπου στην εικόνα γίνεται χρήση του αλγόριθμου Viola-Jones (βήμα 4). Το Viola-Jones πλαίσιο ανίχνευσης (detection framework) [41] είναι το πρώτο πλαίσιο ανίχνευσης που μπορεί να ανιχνεύει αντικείμενα σε πραγματικό χρόνο με μεγάλο ποσοστό επιτυχίας. Μπορεί να εκπαιδευτεί για τον εντοπισμό διαφόρων αντικειμένων, αλλά κυρίως χρησιμοποιείται για τον εντοπισμό προσώπων σε μία εικόνα. Η λειτουργία του αλγόριθμου γίνεται σε πραγματικό χρόνο με πολύ υψηλό ποσοστό επιτυχίας αναφορικά με τον εντοπισμό του προσώπου. Πρέπει να σημειωθεί ότι ο αλγόριθμος κάνει αποκλειστικά εντοπισμό και όχι αναγνώριση του προσώπου. Έχει τέσσερα στάδια λειτουργίας τα οποία είναι:

- a) Επιλογή Haar χαρακτηριστικών (Haar Feature Selection)
- b) Δημιουργία μίας ενσωματωμένης εικόνας (Creating an Integral Image)
- c) Εκπαίδευση Adaboost (Adaboost Training)
- d) Ταξινομητές με υπερχείλιση (Cascading Classifiers).

Τα χαρακτηριστικά Haar προσομοιώνουν τα κοινά χαρακτηριστικά που έχουν όλα τα ανθρώπινα πρόσωπα. Για παράδειγμα η περιοχή των ματιών είναι πιο σκοτεινή σε σχέση με τα άνω μάγουλα, ενώ η περιοχή της μύτης είναι πιο φωτεινή σε σχέση με την περιοχή των ματιών, όπως φαίνεται και στην Εικόνα 3. Η ενσωματωμένη εικόνα είναι μια αναπαράσταση της εικόνας που αξιολογεί τα ορθογώνια χαρακτηριστικά σε σταθερό χρόνο, που τους δίνει ένα σημαντικό πλεονέκτημα ταχύτητας/χρόνου εκτέλεσης σε σχέση με εναλλακτικά, πιο πολύπλοκα, χαρακτηριστικά. Επειδή για κάθε χαρακτηριστικό η ορθογώνια περιοχή είναι πάντα γειτονική με άλλο ένα ορθογώνιο, κάθε χαρακτηριστικό που αποτελείται από 2 ορθογώνια μπορεί να υπολογιστεί με έξι πίνακες παραπομπών (array references). Η ενσωματωμένη εικόνα στη θέση (x, y) συμπεριλαμβάνει το άθροισμα των pixels πάνω και στα αριστερά του (x, y) .

Επειδή ακόμα και για μία πολύ μικρή εικόνα τα χαρακτηριστικά που πρέπει να αξιολογηθούν ανέρχονται σε εκατοντάδες χιλιάδες, διαδικασία που είναι υπολογιστικά ακριβή, γίνεται χρήση μίας παραλλαγής του αλγόριθμου εκμάθησης AdaBoost για να

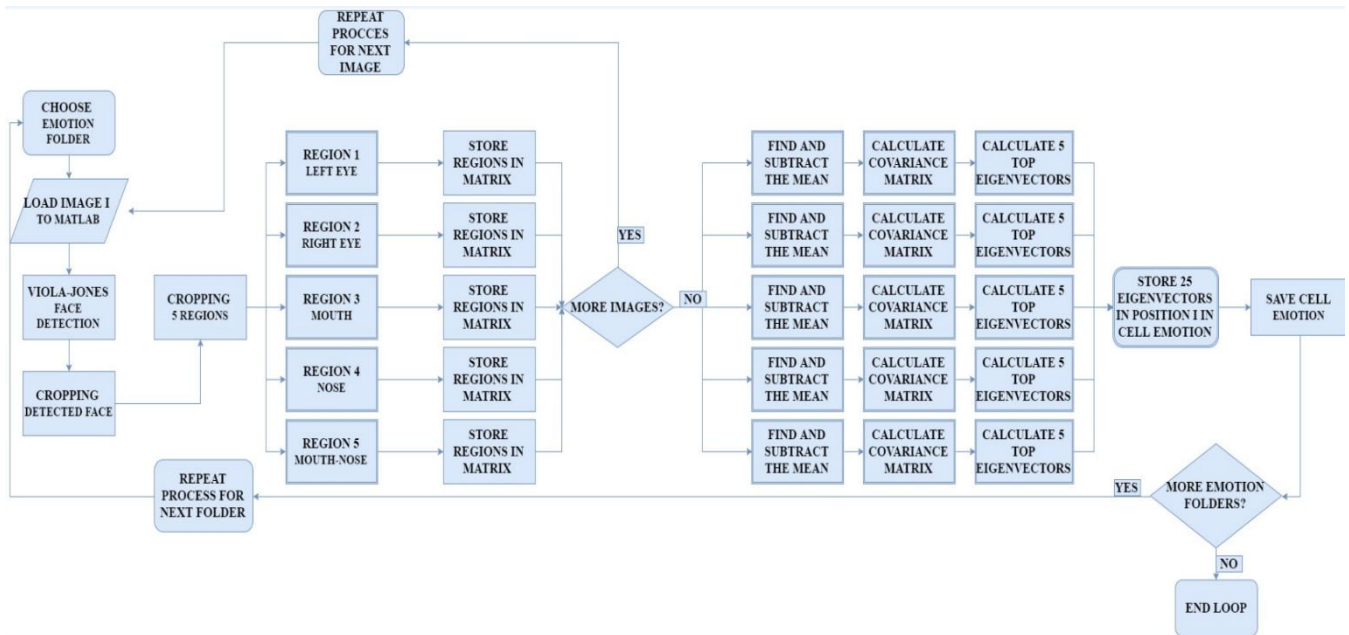
επιλεχθούν τα καλύτερα χαρακτηριστικά και να εκπαιδευτούν οι ταξινομητές που τα χρησιμοποιούν. Ο αλγόριθμος κατασκευάζει έναν δυνατό ταξινομητή, συνδυάζοντας γραμμικά έναν αριθμό από αδύναμους απλούς ταξινομητές στους οποίους έχουν δοθεί τα σχετικά βάρη. Τέλος, στην υπερχείλιση κάθε στάδιο της αποτελείται από ένα δυνατό ταξινομητή. Οπότε όλα τα χαρακτηριστικά χωρίζονται σε αρκετά στάδια που κάθε στάδιο έχει ένα συγκεκριμένο αριθμό χαρακτηριστικών. Σε κάθε στάδιο πρέπει να αποφασιστεί αν το δοθέν χαρακτηριστικό είναι πιθανώς πρόσωπο ή οριστικά μη-πρόσωπο και έτσι να απορριφθεί άμεσα σε οποιοδήποτε στάδιο και αν βρίσκεται η υπολογιστική διαδικασία.



Εικόνα 3: παραδείγματα ορθογωνίων χαρακτηριστικών Haar

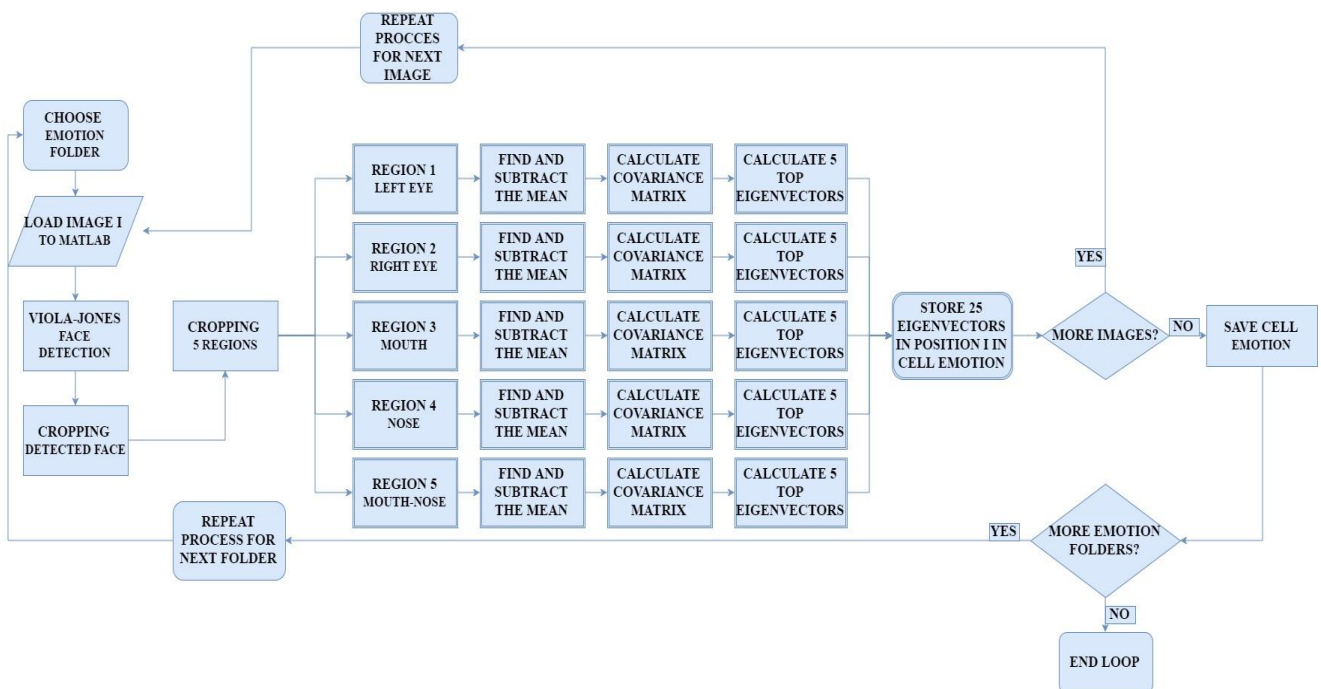
Μετά την εκτέλεση του Viola-Jones αλγόριθμου, που παράγει ως αποτέλεσμα ένα ορθογώνιο με τις συντεταγμένες της περιοχής του προσώπου, αποκόπτεται το πρόσωπο από την αρχική εικόνα (βήμα 5) και γίνεται στατική αποκοπή των πέντε περιοχών, που είναι αριστερό μάτι, δεξί μάτι, μύτη, στόμα και μύτη με στόμα. Οι διαστάσεις είναι 40 επί 40 pixels, 40 επί 40, 70 επί 60, 60 επί 90 και 110 επί 95 αντίστοιχα. Η αποκοπή των περιοχών είναι στατική, δηλαδή με συγκεκριμένες συντεταγμένες, γιατί, ενώ ο Viola-Jones αλγόριθμος δίνει την δυνατότητα ανίχνευσης των περιοχών αυτών, θέλει ρύθμιση χειροκίνητα ενός κατωφλίου (threshold) για κάθε εικόνα και περιοχή ξεχωριστά κάτι που είναι πρακτικά αδύνατο για ένα αυτόματο σύστημα αναγνώρισης (βήμα 6). Η κάθε περιοχή πρέπει να εισαχθεί σε ένα πίνακα, που θα αποτελείται από Y περιοχές ίδιου τύπου με $P=M \times N$ διαστάσεις ανά περιοχή και θα έχει αποτέλεσμα έναν $Y \times P$ πίνακα (βήμα 7). Τα βήματα 1-7 επαναλαμβάνονται για όλες τις εικόνες ενός συναισθήματος.

Όταν δημιουργηθούν οι πέντε τελικοί πίνακες για ένα συναίσθημα, βρίσκεται η μέση τιμή (ή το μέσο διάνυσμα) του καθενός και μετά αφαιρείται από κάθε διάσταση του αντίστοιχου πίνακα (βήμα 8). Όταν ολοκληρωθεί αυτή η διαδικασία βρίσκεται ο πίνακας συνδιακύμανσης για κάθε περιοχή (βήμα 9) και χρησιμοποιώντας αυτόν βρίσκονται οι ιδιοτιμές και τα ιδιοδιανύσματα κάθε περιοχής (βήμα 10). Επιλέγονται τα πέντε ιδιοδιανύσματα που συσχετίζονται με τις πέντε μεγαλύτερες ιδιοτιμές και αποθηκεύονται σε έναν πίνακα (βήμα 11), καταλήγοντας σε έναν πίνακα που περιέχει 25 ιδιοδιανύσματα, πέντε από κάθε περιοχή. Η σειρά με την οποία μπαίνουν στον πίνακα είναι αριστερό μάτι, δεξί μάτι, μύτη, στόμα και μύτη με στόμα. Η ίδια διαδικασία (βήματα 1-11) γίνεται και για τα υπόλοιπα 6 συναισθήματα (βήμα 12).



Εικόνα 4: Γράφημα workflow για τα βήματα που γίνονται κατά την διάρκεια επεξεργασίας για τις εικόνες εκπαίδευσης.

Για τις δοκιμαστικές εικόνες η διαδικασία είναι η ίδια, μόνο που γίνεται για κάθε εικόνα ξεχωριστά. (βήμα 13). Αυτό σημαίνει ότι κάθε δοκιμαστική εικόνα, έχει τις πέντε περιοχές της, που γίνονται στήλες, αφαιρείται η μέση τιμή, υπολογίζεται ο πίνακας συνδιακύμανσης και μέσω αυτού οι ιδιοτιμές και τα ιδιοδιανύσματα, αλλά μόνο για μία δοκιμαστική εικόνα κάθε φορά και όχι για το σύνολο των δοκιμαστικών εικόνων. Κάθε εικόνα πρέπει να έχει μία ετικέτα που προσδιορίζει το συναίσθημα που προβάλλεται σε αυτή για να γίνει η επαλήθευση στο τέλος (βήμα 14). Όταν γίνει αυτό για όλες τις δοκιμαστικές

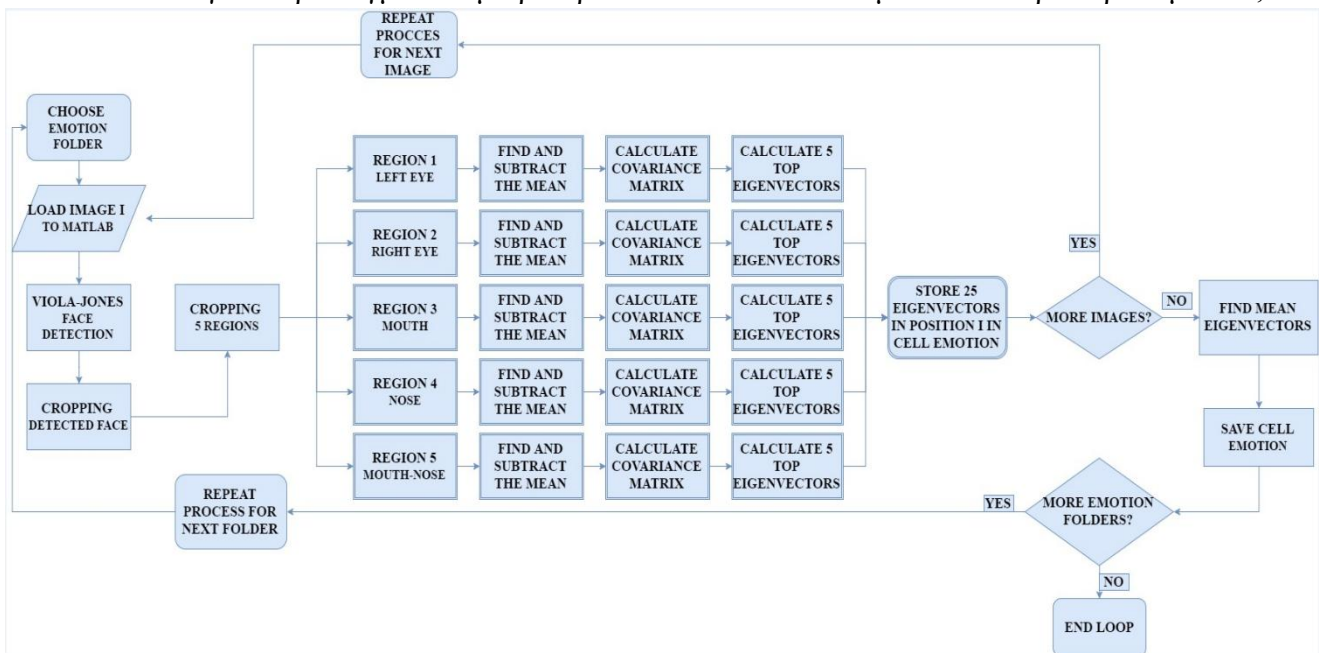


εικόνες, εισάγονται στην επιφάνεια εργασίας του matlab τα ιδιοδιανύσματα των 7 συναισθημάτων και των δοκιμαστικών εικόνων (βήμα 15).

Εικόνα 5: Γράφημα workflow για τα βήματα που γίνονται κατά την διάρκεια επεξεργασίας για τις εικόνες δοκιμής.

Υπολογίζονται οι ευκλείδειες αποστάσεις για κάθε δοκιμαστική εικόνα σε σχέση με τα επτά συναισθήματα. Κάθε δοκιμαστική εικόνα έχει 175 ευκλείδειες αποστάσεις, 7 ανά ιδιοδιανύσμα. Από αυτές τις 7 επιλέγεται η μικρότερη και σε έναν πίνακα προστίθεται μία «ψήφος» ανάλογα με το συναίσθημα στο οποίο αντιστοιχεί η απόσταση αυτή (βήμα 16). Όταν εκτελεστεί το βήμα 16 για τα 25 ιδιοδιανύσματα μία καταμέτρηση των «ψηφών» αποφασίζει με πλειοψηφία το συναίσθημα που εκφράζει η δοκιμαστική εικόνα (βήμα 17). Για παράδειγμα αν οι «ψηφοί» είναι 10 για χαρά, 5 για έκπληξη, 3 για ουδέτερη έκφραση, 2 για λύπη, 2 για θυμό, 2 για αηδία και 1 φόβο, τότε η δοκιμαστική εικόνα προβάλλει το συναίσθημα της χαράς, σύμφωνα με το σύστημα. Για να επαληθευτεί αυτό, γίνεται σύγκριση της ετικέτας που έχει η δοκιμαστική εικόνα με αυτή του αποτελέσματος (βήμα 18). Αν συμφωνούν τότε το σύστημα έχει κάνει σωστή κατηγοριοποίηση.

Για την εναλλακτική προσέγγιση τα βήματα 1-6 είναι ακριβώς τα ίδια. Αφού αποκοπούν οι περιοχές αποθηκεύονται σε έναν πίνακα πινάκων (βήμα 7β). Ένας πίνακας πινάκων είναι ένας ταξινομημένος πίνακας δεικτών που μέσα σε αυτόν είναι δυνατή η αποθήκευση των διευθύνσεων μνήμης πολλών πινάκων. Στο matlab ένας τέτοιος πίνακας λέγεται κελί (cell). Για της πέντε περιοχές υπολογίζεται η μέση τιμή κάθε περιοχής και αφαιρείται από την ίδια, κρατώντας το αποτέλεσμα της αφαίρεσης (βήμα 8β). Υπολογίζεται ο πίνακας συνδιακύμανσης για κάθε περιοχή (βήμα 9β) και γίνεται χρήση αυτού για την εύρεση των ιδιοτιμών και των ιδιοδιανυσμάτων (βήμα 10β). Επιλέγονται τα πέντε ιδιοδιανύσματα που συσχετίζονται με τις μεγαλύτερες ιδιοτιμές από κάθε περιοχή, αποθηκεύονται σε ένα άλλο κελί (βήμα 11β) και αυτό γίνεται για όλες τις εικόνες ενός συναισθήματος. Όταν όλα τα ιδιοδιανύσματα από όλες τις εικόνες έχουν αποθηκευτεί, κάθε ιδιοδιανύσμα προστίθεται με το αντίστοιχο του (βήμα 12β). Δηλαδή το πρώτο ιδιοδιανύσμα, που είναι για παράδειγμα το μικρότερο από τα ιδιοδιανύσματα του αριστερού ματιού,



προστίθεται μαζί με όλα τα αντίστοιχα των υπόλοιπων εικόνων, το δεύτερο με τα υπόλοιπα δεύτερα κτλ, μέχρι να γίνουν ένα σύνολο από 25 ιδιοδιανύσματα. Τέλος, γίνεται διαίρεση

του κάθε τελικού ιδιοδιανύσματος με τον αριθμό των εικόνων που έχουν εισαχθεί στο κελί για να δημιουργηθούν 25 «μέσα» ιδιοδιανύσματα (βήμα 13β).

Εικόνα 6: Γράφημα workflow για τα βήματα που γίνονται κατά την διάρκεια επεξεργασίας για τις εικόνες εκπαίδευσης της δευτερης μεθόδου.

Η λογική πίσω από αυτή την προσέγγιση είναι ότι κάθε εικόνα που προβάλλει ένα συναίσθημα και συνεπώς τα ιδιοδιανύσματα που παράγονται μέσω αυτής έχουν κάποιες ομοιότητες, αλλά σίγουρα θα υπάρχουν και κάποιες ακραίες τιμές που θα επηρεάζουν το τελικό αποτέλεσμα. Οπότε βγάζοντας έναν μέσο όρο από όλα τα ιδιοδιανύσματα, πιθανόν να μειωθούν αυτές οι ακραίες διαφορές. Τα βήματα αυτά γίνονται για όλα τα συναισθήματα, ενώ για τις δοκιμαστικές εικόνες εκτελούνται τα βήματα 1-6 και 7β-11β, χωρίς να γίνουν αλλαγές στον κώδικα. Τέλος εκτελούνται τα βήματα 15-18 που τώρα γίνεται χρήση των «μέσων» ιδιοδιανυσμάτων και των ιδιοδιανυσμάτων των δοκιμαστικών εικόνων που δεν έχουν γίνει στήλες.

4.3 Υλοποίηση συστήματος – Κώδικας των δύο υλοποιήσεων

4.3.1 Κώδικας πρώτης υλοποίησης

Η πρώτη υλοποίηση όπως προαναφέρθηκε είναι αυτή των Jeemoni Kalita και Karen Das [2]. Στην υλοποίηση αυτή, έγινε προσπάθεια να δημιουργηθεί από την αρχή το αυτόματο σύστημα αναγνώρισης συναισθημάτων σύμφωνα με τον αλγόριθμο που παρουσιάζεται στο εν λόγω άρθρο. Το αρχείο **EigenvectorsTrainVer001.m** περιέχει τον εκτελέσιμο κώδικα για την ανάγνωση όλων των εικόνων από κάθε συναίσθημα, την επεξεργασία τους για να βρεθούν τα ιδιοδιανύσματα και την αποθήκευση αυτών για περαιτέρω χρήση. Το σύνολο εικόνων κάθε συναισθήματος διαβάζεται σαν φάκελος και ο κώδικας πρέπει να εκτελεστεί μία φορά για κάθε συναίσθημα, οπότε με μία επαναληπτική διαδικασία (for loop) εκτελείται ο κώδικας εφτά φορές. Η εντολή `close all;` κλείνει όλα τα ανοικτά παράθυρα του matlab (όπως γραφήματα), η εντολή `clear all;` καθαρίζει το workspace του matlab (διαγράφει όλες τις μεταβλητές που έχουν δηλωθεί) και η εντολή `clc;` καθαρίζει το command window από της εντολές και τα αποτελέσματα που έχουν εκτυπωθεί.

Η `ii` είναι μία μεταβλητή που χρησιμοποιείται για να είναι δυνατή η αποθήκευση των εικόνων σε διαφορετικές θέσεις ενός πίνακα και πριν την εισαγωγή κάθε νέας εικόνας αυξάνεται κατά ένα. Η `eigsum` είναι μία μεταβλητή που καθορίζει πόσα ιδιοδιανύσματα θα χρησιμοποιηθούν τελικά. Αν το `eigsum` είναι ίσο με ένα τότε θα χρησιμοποιηθούν 5 ιδιοδιανύσματα, αν το `eigsum` είναι ίσο με δύο τότε θα χρησιμοποιηθούν 10 ιδιοδιανύσματα, αν είναι ίσο με τρία τότε 15 ιδιοδιανύσματα κτλ.

Η μεταβλητή `topLevelFolder = uigetdir(start_path);` κρατάει τη διαδρομή (path) του αρχικού φακέλου που δίνεται μέσω της συνάρτησης `uigetdir`, η οποία παίρνει σαν όρισμα μία διαδρομή την `start_path`. Αν η αρχική διαδρομή δεν περιέχει εικόνες δεν εκτελεί καμία ενέργεια, γιατί αν δεν υπάρχουν εικόνες δεν μπορεί για να συνεχιστεί η εκτέλεση του προγράμματος. Η εντολή `A=imread(fullFileName);` διαβάζει στην μεταβλητή `A` την εικόνα που βρίσκεται στην διαδρομή (path) που δίνεται από την μεταβλητή `fullFileName`, όπου `fullFileName = fullfile(thisFolder, baseFileNames(f).name);` και όπου `thisFolder`

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

η διαδρομή του φακέλου και `baseFileNames(f).name` μία δομή (structure) με πέντε πεδία με το `name` να δίνει το πλήρες όνομα της εικόνας που δείχνει το `(f)` (`f = 1 : numberOfImageFiles`). Παρακάτω παραθέτεται το κομμάτι του κώδικα που χρησιμοποιείται για την ανάγνωση των εικόνων.

```
% ii is used to in order to add the next image in the right position.
ii=1;
% eigsum is a variable that determines how many eigenvectors will
eventually be used.
eigsum=1;
% Define a starting folder.
start_path = fullfile('C:\Users\P.Bolarakis\Desktop');
% Ask user to confirm or change.
topLevelFolder = uigetdir(start_path);
if topLevelFolder == 0
    return;
end
% Get list of all subfolders.
allSubFolders = genpath(topLevelFolder);
% Parse into a cell array.
remain = allSubFolders;
listOfFolderNames = {};
while true
    [singleSubFolder, remain] = strtok(remain, ';');
    if isempty(singleSubFolder)
        break;
    end
    listOfFolderNames = [listOfFolderNames singleSubFolder];
end
numberOfFolders = length(listOfFolderNames)

% Process all image files in those folders.
for k = 1 : numberOfFolders
    % Get this folder and print it out.
    thisFolder = listOfFolderNames{k};
    fprintf('Processing folder %s\n', thisFolder);

    % Get PNG files.
    filePattern = sprintf('%s/*.png', thisFolder);
    baseFileNames = dir(filePattern);
    % Add on TIF files.
    filePattern = sprintf('%s/*.tif', thisFolder);
    baseFileNames = [baseFileNames; dir(filePattern)];
    % Add on JPG files.
    filePattern = sprintf('%s/*.jpg', thisFolder);
    baseFileNames = [baseFileNames; dir(filePattern)];
    numberOfImageFiles = length(baseFileNames);
    % Now we have a list of all files in this folder.

    if numberOfImageFiles >= 1
        % Go through all those image files.
        for f = 1 : numberOfImageFiles
            fullFileName = fullfile(thisFolder, baseFileNames(f).name);
            A=imread(fullFileName);
        end
    else
        fprintf('    Folder %s has no image files in it.\n', thisFolder);
    end
end
```



```
end
```

Μέσα στην `for f = 1 : numberOfImageFiles` δηλώνεται το `FDetect = vision.CascadeObjectDetector;` που είναι ένας ανιχνευτής αντικειμένου Cascade που κάνει χρήση του Viola Jones αλγορίθμου [41], όπως αναφέρθηκε παραπάνω και είναι συνάρτηση του computer vision toolbox του matlab. Η συνάρτηση `step (BB = step (FDetect, A);)` που παίρνει σαν ορίσματα την εικόνα `A` που έχει διαβαστεί και το `CascadeObjectDetector, FDetect` δίνει σαν αποτέλεσμα το πλαίσιο οριοθέτησης `BB` (bounding box), οπού κρατάει τις συντεταγμένες των γωνιών του πλαισίου που βρίσκεται μέσα το πρόσωπο. Με μία `if` διασφαλίζεται ότι δεν θα υπάρχουν πολλαπλά `BB` στην εικόνα, δηλαδή δεν θα έχει αναγνωρίσει λάθος ένα μη πρόσωπο σαν πρόσωπο. `if size(BB,1) == 1 { εντολές } else ii=ii+1; fprintf(' Multiple detections on file %s\n', fullFileName);`

Η συνάρτηση `I2=imcrop(I)` δημιουργεί ένα διαδραστικό εργαλείο περικοπής εικόνας για την εικόνα `I` και αποθηκεύει την περιοχή που επιλέχθηκε με το εργαλείο στο `I2`. Το εργαλείο περικοπής είναι ένα μετακινούμενο ορθογώνιο με δυνατότητα αλλαγής μεγέθους με την χρήση του ποντικιού. Σε αυτή την περίπτωση γίνεται χρήση της `imcrop` με άλλο ένα όρισμα το `rect` όπου (`rect == BB`) είναι ένα διάνυσμα θέσης τεσσάρων στοιχείων της μορφής [`xmin ymin ύψος πλάτος`] που καθορίζει το μέγεθος και τη θέση του ορθογωνίου περικοπής [42]. `FDetect=imcrop(A, BB);` Η επόμενη εντολή `X = imresize(FDetect, 'OutputSize', [300 300]);` κάνει χρήση της συνάρτησης για αλλαγή μεγέθους μίας εικόνας `imresize` που παίρνει τρία ορίσματα. Το πρώτο όρισμα καθορίζει έναν πίνακα, την εικόνα που θα γίνει η αυξομείωση του μεγέθους, το όρισμα 'OutputSize' που δηλώνει ότι το μέγεθος της εικόνας εξόδου θα δίνεται από το διάνυσμα δύο μεταβλητών στην μορφή [`numrows numcols`] που είναι το τρίτο όρισμα [43].

Επόμενο βήμα είναι να γίνει αποκοπή των περιοχών και η εισαγωγή τους σε έναν πίνακα ανά περιοχή. Γίνεται και πάλι χρήση της `imresize` αυτή την φορά σε συνδυασμό με την `imcrop` για την αποκοπή και την αλλαγή του μεγέθους των περιοχών των ματιών, της μύτης, του στόματος και της μύτης και του στόματος μαζί. Για την εισαγωγή όλων των περιοχών μίας κατηγορίας στον δικό τους αντίστοιχο πίνακα χρησιμοποιούνται οι ιδιότητες του matlab για τον χειρισμό πινάκων. Αρχικά εισάγεται η πρώτη εικόνα που έχει χωριστεί σε πέντε περιοχές στους πέντε αντίστοιχους πίνακες περιοχών και με μια `if` καθορίζουμε ποια θέση στους πίνακες περιοχών θα πάρουν οι επόμενες εικόνες. Οι πίνακες `LeftEye, RightEye, Mouth, Nose, MouthNose` στο τέλος της επανάληψης `for` θα περιέχουν αντίστοιχα τις περιοχές για το αριστερό μάτι, το δεξί μάτι, το στόμα, την μύτη και το στόμα με μύτη για όλες τις εικόνες ενός συναισθήματος. Παρακάτω δίνεται ο κώδικας μέσα στην `for`.

```
if numberOfImageFiles >= 1
    % Go through all those image files.
    for f = 1 : numberOfImageFiles

        fullFileName = fullfile(thisFolder, baseFileNames(f).name);
        A=imread(fullFileName);
        A=im2double(A);
        % Create a Viola-Jones, cascade object detector
        FDetect = vision.CascadeObjectDetector;
        BB = step (FDetect, A);
```

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

```
if size(BB,1) == 1

    FDetect=imcrop(A,BB);
    X = imresize(FDetect,'OutputSize',[300 300]);
    % Cropped Left eye
    if ii==1

        LeftEye=imresize(imcrop(X,[40 60 100 105]),[40 40]);
        % Cropped Right eye
        RightEye = imresize(imcrop(X,[150 60 100 105]),[40 40]);
        % Cropped Mouth
        Mouth=imresize(imcrop(X,[90 200 130 100]),[90 60]);
        % Cropped Nose
        Nose = imresize(imcrop(X,[100 140 100 70]),[70 60]);
        % Cropped Mouth and Nose
        MouthNose = imresize(imcrop(X,[100 145 110 155]),[95 110]);
    else

        LeftEyeTemp=imresize(imcrop(X,[40 60 100 105]),[40 40]);
        LeftEye = [LeftEye,LeftEyeTemp];
        % Cropped Right eye
        RightEyeTemp = imresize(imcrop(X,[150 60 100 105]),[40 40]);
        RightEye= [RightEye,RightEyeTemp];
        % Cropped Mouth
        MouthTemp=imresize(imcrop(X,[90 200 130 100]),[90 60]);
        Mouth=[Mouth,MouthTemp];
        % Cropped Nose
        NoseTemp = imresize(imcrop(X,[100 140 100 70]),[70 60]);
        Nose=[Nose,NoseTemp];
        % Cropped Mouth and Nose
        MouthNoseTemp = imresize(imcrop(X,[100 145 110 155]),[95 110]);
        MouthNose=[MouthNose,MouthNoseTemp];
    end

    ii=ii+1;
    fprintf('        Processing image file %s\n', fullFileName);

else

    ii=ii+1;
    fprintf('        Multiple detections on file %s\n', fullFileName);

end
end
else
    fprintf('        Folder %s has no image files in it.\n', thisFolder);
end
```

Μετά την δημιουργία των πέντε πινάκων των περιοχών βρίσκεται η μέση τιμή του καθενός και αφαιρείται από κάθε διάσταση του. Αυτό γίνεται με την χρήση της συνάρτησης `mean` που στην απλή της μορφή παίρνει σαν όρισμα έναν πίνακα και επιστρέφει ένα διάνυσμα που έχει όσες στήλες έχει και ο πίνακας εισόδου και μία γραμμή με τον μέσο όρο κάθε στήλης του πίνακα [44]. $X1 = \text{mean}(\text{LeftEye}, 2)$; Για να βρεθεί η μέση τιμή χρησιμοποιείται ξανά η συνάρτηση `mean` στο $X1$ αυτή την φορά. Η εντολή $Y1 = \text{size}(\text{LeftEye})$; κάνει χρήση της συνάρτησης `size` που δίνει σαν αποτέλεσμα τον αριθμό των γραμμών και των στηλών του ορίσματος που παίρνει. Δηλαδή το $Y1(1)$ περιέχει τον αριθμό των γραμμών και το $Y1(2)$ τον αριθμό των στηλών του `LeftEye`. Γίνεται η αφαίρεση της μέσης τιμής της κάθε περιοχής από όλες τις διαστάσεις του

αντίστοιχου πίνακα με μία `for` και μετά βρίσκεται ο πίνακας συνδιακύμανσης με την χρήση της συνάρτησης `cov` [45]. Η συνάρτηση αυτή, με τον τρόπο που χρησιμοποιείται εδώ, παίρνει σαν όρισμα ένα πίνακα και δίνει σαν έξοδο τον πίνακα συνδιακύμανσης του.

Τα ιδιοδιανύσματα και οι ιδιοτιμές δίνονται με την χρήση της συνάρτησης `eigs` [46], `[V1,D1]=eigs(X2,5*eigsum);` όπου `D1` οι ιδιοτιμές και `V1` τα ιδιοδιανύσματα. Το πρώτο όρισμα το `X2` πρέπει να είναι τετραγωνικός πίνακας (εδώ είναι ένας πίνακας συνδιακύμανσης) και το δεύτερο όρισμα μας δίνει έναν συγκεκριμένο αριθμό ιδιοδιανυσμάτων (σε αυτή την περίπτωση πέντε), αυτά που συσχετίζονται με τις μεγαλύτερες ιδιοτιμές. Έτσι δεν χρειάζεται να αποθηκευτεί στην μνήμη ο πίνακας με όλα τα ιδιοδιανύσματα, παρά μόνο αυτά που χρειάζονται. Πριν την τελική αποθήκευση των ιδιοδιανυσμάτων δημιουργείται ένα κελί (`cell`), που είναι ένας πίνακας πινάκων και με μία `for` τα ιδιοδιανύσματα κάθε περιοχής τοποθετούνται στην σωστή θέση. Γίνεται η τελική αποθήκευση του κελιού `Emotion` για ένα συναίσθημα, με την χρήση της `uisave('Emotion');` που ανοίγει ένα παράθυρο για να επιλεγεί η τοποθεσία αποθήκευσης καθώς και το όνομα που θα δοθεί στο κελί. Για τα συναισθήματα της έκπληξης, της ευτυχίας, του φόβου, του θυμού, της αηδίας, της λύπης και της ουδέτερης έκφρασης δόθηκαν τα ονόματα `SurpriseVal.mat`, `HappinesVal.mat`, `FearVal.mat`, `AngerVal.mat`, `DisgustVal.mat`, `SadnessVal.mat`, `NeutralVal.mat` αντίστοιχα και θα χρειαστούν όταν βρεθεί η ευκλείδεια απόσταση.

```

X1=mean(LeftEye,2);
X1=mean(X1);
Y1=size(LeftEye);
for i=1:Y1(2)*Y1(1)
    LeftEye(i)=LeftEye(i)-X1;
end
X1=mean(RightEye,2);
X1=mean(X1);
Y1=size(RightEye);
for i=1:Y1(2)*Y1(1)
    RightEye(i)=RightEye(i)-X1;
end
X1=mean(Mouth,2);
X1=mean(X1);
Y1=size(Mouth);
for i=1:Y1(2)*Y1(1)
    Mouth(i)=Mouth(i)-X1;
end
X1=mean(Nose,2);
X1=mean(X1);
Y1=size(Nose);
for i=1:Y1(2)*Y1(1)
    Nose(i)=Nose(i)-X1;
end
X1=mean(MouthNose,2);
X1=mean(X1);
Y1=size(MouthNose);
for i=1:Y1(2)*Y1(1)
    MouthNose(i)=MouthNose(i)-X1;
end

X2=cov(LeftEye');
[V1,D1]=eigs(X2,5*eigsum);
clear X2;
X2=cov(RightEye');

```

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

```
[V2,D2]=eigs(X2,5*eigsum);
clear X2;
X2=cov(Mouth');
[V3,D3]=eigs(X2,5*eigsum);
clear X2;
X2=cov(Nose');
[V4,D4]=eigs(X2,5*eigsum);
clear X2;
X2=cov(MouthNose');
[V5,D5]=eigs(X2,5*eigsum);
clear X2;
Em=cell(1,25*eigsum);
for i=1:5*eigsum
    Emotion{1,i}=V1(:,i);
    Emotion{1,i+(5*eigsum)}=V2(:,i);
    Emotion{1,i+(10*eigsum)}=V3(:,i);
    Emotion{1,i+(15*eigsum)}=V4(:,i);
    Emotion{1,i+(20*eigsum)}=V5(:,i);
end

uisave('Emotion');
```

Για τις δοκιμαστικές εικόνες, δηλαδή εκείνες τις οποίες θα επιχειρήσει να αναγνωρίσει αυτόματα το σύστημα ο κώδικας είναι παρόμοιος με αυτόν που χρησιμοποιείται παραπάνω, με κάποιες βασικές διαφορές. Το αρχείο **EigenvectorsTestVer001.m** περιέχει τον εκτελέσιμο κώδικα για την ανάγνωση όλων των εικόνων από κάθε συναισθημα, την επεξεργασία τους για να βρεθούν τα ιδιοδιανύσματα και την αποθήκευση αυτών για περαιτέρω χρήση. Αρχικοποιείται ένα κελί (cell) το `Labels=cell(1000,1);` καθώς και ένα ακόμα το `Emotion=cell(1000,25*eigsum);`. Το πρώτο χρησιμοποιείται για να κρατάει ετικέτες των εικόνων που διαβάζει για να είναι γνωστό σε ποια κατηγορία εκφράσεων ανήκουν με την εντολή `Labels{ii,1}=thisFolder;`. Το δεύτερο κελί κρατάει τα 25 ιδιοδιανύσματα κάθε εικόνας για να χρησιμοποιηθούν αργότερα για τον υπολογισμό των ευκλείδειων αποστάσεων της κάθε εικόνας.

Για κάθε περιοχή γίνεται ο υπολογισμός και η αφαίρεση της μέσης τιμής από κάθε διάσταση της περιοχής και ο υπολογισμός του πίνακα συνδιακύμανσης για να γίνει τελικά η εξαγωγή των ιδιοτιμών και τον ιδιοδιανυσμάτων. Η διαφορά με τον κώδικα του **EigenvectorsTrainVer001.m** είναι ότι αντί για ένα σύνολο περιοχής (πολλές εικόνες τις ίδιας περιοχής δίπλα δίπλα) γίνεται επεξεργασία σε μία περιοχή (μία μοναδική εικόνα) και τα ιδιοδιανύσματα αυτής αποθηκεύονται στο κελί `Emotion`. Στο `Emotion{ii,i}` η θέση `ii` καθορίζει την εικόνα και η θέση `i` καθορίζει το ιδιοδιάνυσμα `i` της εικόνας. Εφόσον η επεξεργασία πρέπει να γίνεται για κάθε εικόνα οι εντολές για την μέση τιμή, τον πίνακα συνδιακύμανσης, για τα ιδιοδιανύσματα και την εισαγωγή τους στον πίνακα `Em`, βρίσκονται μέσα στην `for f = 1 : numberOfImageFiles`. Ο πίνακας αποθηκεύεται σαν **Values_Test.m** με την χρήση της συνάρτησης `uisave`. Παρακάτω δίνεται ο κώδικας που πραγματοποιεί την επεξεργασία των δοκιμαστικών εικόνων (ο κώδικας μέσα στην `for f = 1`).

```
for f = 1 : numberOfImageFiles
    fullFileName = fullfile(thisFolder, baseFileNames(f).name);
    Labels{ii,1}=thisFolder;
    A=imread(fullFileName);
    A=im2double(A);
    FDetect = vision.CascadeObjectDetector;
    BB = step(FDetect,A);
```

```

if size(BB,1) == 1

    FDetect=imcrop(A,BB);
    X = imresize(FDetect,'OutputSize',[300 300]);

    % Cropped Left eye
    LeftEye{1,ii}=imresize(imcrop(X,[40 60 100 105]),[40 40]);
    X1=mean(LeftEye{1,ii},2);
    X1=mean(X1);
    Y1=size(LeftEye{1,ii});
    for i=1:Y1(2)*Y1(1)
        LeftEye{1,ii}(i)=LeftEye{1,ii}(i)-X1;
    end
    X2=cov(LeftEye{1,ii}');
    [V1,D1]=eigs(X2,5*eigsum);
    clear X2;
    for i=1:5*eigsum
        Emotion{ii,i}=V1(:,i);
    end

    % Cropped Right eye
    RightEye{1,ii} = imresize(imcrop(X,[150 60 100 105]),[40 40]);
    X1=mean(RightEye{1,ii},2);
    X1=mean(X1);
    Y1=size(RightEye{1,ii});
    for i=1:Y1(2)*Y1(1)
        RightEye{1,ii}(i)=RightEye{1,ii}(i)-X1;
    end
    X2=cov(RightEye{1,ii}');
    [V2,D2]=eigs(X2,5*eigsum);
    clear X2;
    for i=1:5*eigsum
        Emotion{ii,i+(5*eigsum)}=V2(:,i);
    End

    % Cropped Mouth
    Mouth{1,ii}=imresize(imcrop(X,[90 200 130 100]),[90 60]);
    X1=mean(Mouth{1,ii},2);
    X1=mean(X1);
    Y1=size(Mouth{1,ii});
    for i=1:Y1(2)*Y1(1)
        Mouth{1,ii}(i)=Mouth{1,ii}(i)-X1;
    end
    X2=cov(Mouth{1,ii}');
    [V3,D3]=eigs(X2,5*eigsum);
    clear X2;
    for i=1:5*eigsum
        Emotion{ii,i+(10*eigsum)}=V3(:,i);
    End

    % Cropped Nose
    Nose{1,ii} = imresize(imcrop(X,[100 140 100 70]),[70 60]);
    X1=mean(Nose{1,ii},2);
    X1=mean(X1);
    Y1=size(Nose{1,ii});
    for i=1:Y1(2)*Y1(1)
        Nose{1,ii}(i) =Nose{1,ii}(i) -X1;
    end
    X2=cov(Nose{1,ii}');
    [V4,D4]=eigs(X2,5*eigsum);
    clear X2;

```

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

```
for i=1:5*eigsum
    Emotion{ii,i+(15*eigsum)}=V4(:,i);
end

% Cropped Mouth and Nose
MouthNose{1,ii} = imresize(imcrop(X,[100 145 110 155]),[95 110]);
X1=mean(MouthNose{1,ii},2);
X1=mean(X1);
Y1=size(MouthNose{1,ii});
for i=1:Y1(2)*Y1(1)
    MouthNose{1,ii}(i)=MouthNose{1,ii}(i)-X1;
end
X2=cov(MouthNose{1,ii}');
[V5,D5]=eigs(X2,5*eigsum);
clear X2;
for i=1:5*eigsum
    Emotion{ii,i+(20*eigsum)}=V5(:,i);
end
ii=ii+1;
fprintf('    Processing image file %s\n', fullFileName);
else
    ii=ii+1;
    fprintf('    Multiple detections on file %s\n', fullFileName);
end
end
end
```

Για να βρεθούν οι Ευκλείδειες αποστάσεις για κάθε δοκιμαστική εικόνα αρχικά εισάγονται τα αποθηκευμένα αρχεία με τα ιδιοδιανύσματα κάθε έκφρασης και των δοκιμαστικών εικόνων. Αυτό γίνεται με την συνάρτηση `importdata` που παίρνει σαν όρισμα την διαδρομή του αρχείου που θα διαβαστεί στην μνήμη του matlab μαζί με το όνομα και τον τύπο αρχείου που είναι. Αν δεν δοθεί διαδρομή, παρά μόνο το όνομα και ο τύπος του αρχείου, τότε επιχειρείται η ανάγνωση του αρχείου με βάση το όνομα του, στην τρέχουσα διαδρομή του matlab. Το αρχείο που θα διαβαστεί πρέπει να οριστεί σε μία μεταβλητή για να κρατηθεί σαν δεδομένο. Το αρχείο **findEuclideanDistancesVer001.m** περιέχει τον εκτελέσιμο κώδικα για τον υπολογισμό της ευκλείδειας απόστασης για κάθε δοκιμαστική εικόνα.

Για να βρεθεί η Ευκλείδεια απόσταση για κάθε έκφραση σε σχέση με τις δοκιμαστικές εικόνες χρειάζεται μία διπλή επανάληψη `for z=1:s(1)` και `for i=1:25*eigsum`, όπου `s=size(Values_test)`; δηλαδή `s(1)` ο αριθμός των δοκιμαστικών εικόνων και το `i` για τα 25 ιδιοδιανύσματα. Οι μεταβλητές `temp2` και `temp3` παίρνουν ένα ιδιοδιάνυσμα από μία έκφραση και το αντίστοιχο ιδιοδιάνυσμα από μία δοκιμαστική εικόνα και με την χρήση της συνάρτησης `norm` [47], που επιτρέπει τον υπολογισμό της L2 νόρμας ή αλλιώς της Ευκλείδειας νόρμας που ισούται με την Ευκλείδεια απόσταση δύο σημείων στο καρτεσιανό επίπεδο, υπολογίζει την Ευκλείδεια απόσταση των δύο ιδιοδιανυσμάτων. Ο τύπος της ευκλείδειας απόστασης στο matlab για δύο σημεία ή στην περίπτωση μας για δύο διανύσματα δίνεται και από τις εντολές `sqrt(sum((temp3-temp2).^2))`; που είναι το Πυθαγόρειο θεώρημα.

Ο πίνακας `EuclideanDistances` κρατάει τις ευκλείδειες αποστάσεις, που είναι αποτέλεσμα όλων των δοκιμαστικών εικόνων και μίας έκφρασης, δηλαδή 25 ανά δοκιμαστική εικόνα. Όταν τελειώσει η διπλή επανάληψη για αυτή την έκφραση ο πίνακας `EuclideanDistances` αποθηκεύεται με τον όνομα της έκφρασης που χρησιμοποιήθηκε για την εξαγωγή των αποστάσεων. Μετά ελευθερώνεται ο πίνακας

EuclideanDistances για να οριστεί ξανά στην επόμενη for επανάληψη για το επόμενο συναίσθημα. Για την έκφραση του θυμού ο πίνακας αποθηκεύεται ως anger.mat, για την έκφραση της ευτυχίας ως happy.mat, για την έκφραση του φόβο ως fear.mat, για την ουδέτερη έκφραση ως neutral.mat, για την έκφραση της έκπληξης ως surprise.mat, για την έκφραση της αηδίας ως disgust.mat και για την έκφραση της λύπης ως sadness.mat. Ο εκτελέσιμος κώδικας για τον υπολογισμό της ευκλείδειας απόστασης παρατίθεται παρακάτω.

```

Values_test=importdata('Values_test.mat');
angerVal=importdata('AngerVal.mat');
surpriseVal=importdata('SurpriseVal.mat');
happinesVal=importdata('HappinessVal.mat');
neutralVal=importdata('NeutralVal.mat');
fearVal=importdata('FearVal.mat');
disgustVal=importdata('DisgustVal.mat');
sadnessVal=importdata('SadnessVal.mat');

eigsum=1;
s=size(Values_test);

for z=1:s(1)
    for i=1:25*eigsum
        temp2=angerVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        %Euclidean distance can also be given given by the Pythagorean
        formula: EuDis{z,i} = sqrt(sum((temp3-temp2).^2));
        EuclideanDistances{z,i} = norm(temp3 - temp2);
    end
end
uisave('EuclideanDistances','anger');
clear EuclideanDistances;

for z=1:s(1)
    for i=1:25*eigsum
        temp2=happinesVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        EuclideanDistances{z,i} = norm(temp3 - temp2);
    end
end
uisave('EuclideanDistances','happy');
clear EuclideanDistances;

for z=1:s(1)
    for i=1:25*eigsum
        temp2=fearVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        EuclideanDistances{z,i} = norm(temp3 - temp2);
    end
end
end

```

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

```
uisave('EuclideanDistances','fear');
clear EuclideanDistances;

for z=1:s(1)
    for i=1:25*eigsum
        temp2=neutralVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        EuDis{z,i} = norm(temp3 - temp2);
    end
end
uisave('EuclideanDistances','neutral');
clear EuclideanDistances;

for z=1:s(1)
    for i=1:25*eigsum
        temp2=surpriseVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        EuDis{z,i} = norm(temp3 - temp2);
    end
end
uisave('EuclideanDistances','surprise');
clear EuclideanDistances;

for z=1:s(1)
    for i=1:25*eigsum
        temp2=disgustVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        EuDis{z,i} = norm(temp3 - temp2);
    end
end
uisave('EuclideanDistances','disgust');
clear EuclideanDistances;

for z=1:s(1)
    for i=1:25*eigsum
        temp2=sadnessVal{1,i};
        temp3=Values_test{z,i};
        if isempty(temp3) == 1
            break;
        end
        EuDis{z,i} = norm(temp3 - temp2);
    end
end
uisave('EuclideanDistances','sadness');
clear EuclideanDistances;
```

Εφόσον υπάρχουν τα αρχεία με τις Ευκλείδειες αποστάσεις των δοκιμαστικών εικόνων για κάθε έκφραση το μόνο που μένει είναι να δημιουργηθεί κώδικας που να συγκρίνει τις αποστάσεις αυτές και να μετράει τις ψήφους σύμφωνα με την μικρότερη απόσταση. Κάθε δοκιμαστική εικόνα έχει 25 Ευκλείδειες αποστάσεις ανά έκφραση, με 7

εκφράσεις συνολικά, δηλαδή 175 αποστάσεις. Το αρχείο με τον κώδικα σύγκρισης των αποστάσεων και της καταμέτρησης των ψήφων κάθε εικόνας έχει ονομαστεί **importEuclideanDistances.m**.

Αρχικά φορτώνονται στην μνήμη οι Ευκλείδειες αποστάσεις ανά συναίσθημα, οι ετικέτες των δοκιμαστικών εικόνων και ορίζονται τρία κελιά τα A, B και C. Αρχίζοντας ανάποδα, το κελί C κρατάει την μικρότερη ευκλείδεια απόσταση για κάθε δοκιμαστική εικόνα κάνοντας συγκρίσεις ανάμεσα στις αποστάσεις των επτά εκφράσεων μέχρι να βρεθεί η μικρότερη. Κάθε δοκιμαστική εικόνα καταλαμβάνει 2 γραμμές και 25 στήλες του κελιού C. Η πρώτη γραμμή έχει σε κάθε στήλη της τις 25 τιμές των Ευκλείδειων αποστάσεων και η δεύτερη μία συμβολοσειρά που δηλώνει σε ποια έκφραση ανήκει η απόσταση αυτή. Όταν συγκρίνονται δύο αποστάσεις και αφού βρεθεί η μικρότερη απόσταση, η αντίστοιχη συμβολοσειρά για την έκφραση εισάγεται στην δεύτερη γραμμή.

Το κελί A με την χρήση μίας δήλωσης switch συγκρίνει τις συμβολοσειρές που υπάρχουν στην δεύτερη γραμμή του κελιού C για κάθε δοκιμαστική εικόνα και μετράει τις ψήφους που έχει συνολικά κάθε έκφραση. Κάθε δοκιμαστική εικόνα καταλαμβάνει 2 στήλες και 8 γραμμές του κελιού A. Στην πρώτη στήλη στις 7 πρώτες γραμμές υπάρχουν συμβολοσειρές για κάθε έκφραση. Στην δεύτερη στήλη στις 7 πρώτες γραμμές υπάρχουν οι ψήφοι για την αντίστοιχη έκφραση της πρώτης στήλης. Στην τελευταία γραμμή (8^η) στην πρώτη στήλη αναγράφεται η συμβολοσειρά 'testing for' και στην δεύτερη η πραγματική έκφραση που αποτυπώνεται στην δοκιμαστική εικόνα, που έχει δοθεί από τον πίνακα Labels με τις ετικέτες.

Με μία if και την χρήση της strcmp, η οποία επιστρέφει άσσο εάν τα ορίσματα της που είναι δυο συμβολοσειρές είναι πανομοιότυπα και μηδέν αν δεν είναι, συγκρίνεται η συμβολοσειρά της όγδοης γραμμής του A που δείχνει την έκφραση, με την συμβολοσειρά με τις περισσότερες ψήφους και αν η strcmp επιστρέψει άσσο, προστίθεται συν ένα σε αντίστοιχη σειρά του κελιού B. Το κελί B δηλαδή κρατάει το σύνολο των επιτυχώς αναγνωρισμένων εικόνων ανά έκφραση. Ο κώδικας για την εύρεση της μικρότερης ευκλείδειας απόστασης και την καταμέτρηση των ψήφων δίνεται παρακάτω.

```
anger=importdata('anger.mat');
surprise=importdata('surprise.mat');
happiness=importdata('happy.mat');
neutral=importdata('neutral.mat');
fear=importdata('fear.mat');
disgust=importdata('disgust.mat');
sadness=importdata('sadness.mat');
Labels=importdata('Labels.mat');
A=cell(8,5000);
eigsum=1;
s=size(anger);
```

```
for z=0:s(1)-1
    A{1,1+2*z}={'Surprise'};
    A{2,1+2*z}={'Happiness'};
    A{3,1+2*z}={'Sadness'};
    A{4,1+2*z}={'Neutral'};
    A{5,1+2*z}={'Fear'};
    A{6,1+2*z}={'Disgust'};
    A{7,1+2*z}={'Anger'};
    A{1,2+2*z}=0;
    A{2,2+2*z}=0;
    A{3,2+2*z}=0;
```

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

```
A{4,2+2*z}=0;
A{5,2+2*z}=0;
A{6,2+2*z}=0;
A{7,2+2*z}=0;

for i=1:25*eigsum
    C{1+2*z,i}=min(anger{1+z,i},surprise{1+z,i});
    if anger{1+z,i} < surprise{1+z,i}
        C{2+2*z,i}='7.anger';
    else
        C{2+2*z,i}='1.surprise';
    end
    C{1+2*z,i}=min(C{1+2*z,i},happiness{1+z,i});
    if C{1+2*z,i} < happiness{1+z,i}
        C{2+2*z,i}=C{2+2*z,i};
    else
        C{2+2*z,i}='2.happiness';
    end
    C{1+2*z,i}=min(C{1+2*z,i},neutral{1+z,i});
    if C{1+2*z,i} < neutral{1+z,i}
        C{2+2*z,i}=C{2+2*z,i};
    else
        C{2+2*z,i}='4.neutral';
    end
    C{1+2*z,i}=min(C{1+2*z,i},fear{1+z,i});
    if C{1+2*z,i} < fear{1+z,i}
        C{2+2*z,i}=C{2+2*z,i};
    else
        C{2+2*z,i}='5.fear';
    end
    C{1+2*z,i}=min(C{1+2*z,i},disgust{1+z,i});
    if C{1+2*z,i} < disgust{1+z,i}
        C{2+2*z,i}=C{2+2*z,i};
    else
        C{2+2*z,i}='6.disgust';
    end
    C{1+2*z,i}=min(C{1+2*z,i},sadness{1+z,i});
    if C{1+2*z,i} < sadness{1+z,i}
        C{2+2*z,i}=C{2+2*z,i};
    else
        C{2+2*z,i}='3.sadness';
    end
    switch C{2+2*z,i}
        case '1.surprise'
            A{1,2+2*z}=A{1,2+2*z}+1;
        case '2.happiness'
            A{2,2+2*z}=A{2,2+2*z}+1;
        case '3.sadness'
            A{3,2+2*z}=A{3,2+2*z}+1;
        case '4.neutral'
            A{4,2+2*z}=A{4,2+2*z}+1;
        case '5.fear'
            A{5,2+2*z}=A{5,2+2*z}+1;
        case '6.disgust'
            A{6,2+2*z}=A{6,2+2*z}+1;
        case '7.anger'
            A{7,2+2*z}=A{7,2+2*z}+1;
    end
end
end
l=size(Labels);
```

Μπολαράκης Παναγιώτης

```
for x=1:s(1)
    if strfind(Labels{x,1}, 'Surprise')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Surprise'};
    end
    if strfind(Labels{x,1}, 'Happiness')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Happiness'};
    end
    if strfind(Labels{x,1}, 'Sadness')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Sadness'};
    end
    if strfind(Labels{x,1}, 'Neutral')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Neutral'};
    end
    if strfind(Labels{x,1}, 'Fear')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Fear'};
    end
    if strfind(Labels{x,1}, 'Disgust')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Disgust'};
    end
    if strfind(Labels{x,1}, 'Anger')
        A{8,2*x-1}={ 'Testing for'};
        A{8,2*x}={ 'Anger'};
    end
end

B=zeros(7,1);
for z=0:s(1);
    if isempty(A{1,2+2*z}) == 1
        break;
    end
    if strcmp(A{8,2+2*z}, 'Surprise') && (A{1,2+2*z}>
A{2,2+2*z}) && (A{1,2+2*z}> A{3,2+2*z}) && (A{1,2+2*z}>
A{4,2+2*z}) && (A{1,2+2*z}>
A{5,2+2*z}) && (A{1,2+2*z}>A{6,2+2*z}) && (A{1,2+2*z}>A{7,2+2*z})
        B(1)=B(1)+1;
    end
    if strcmp(A{8,2+2*z}, 'Happiness') && (A{2,2+2*z}>
A{1,2+2*z}) && (A{2,2+2*z}> A{3,2+2*z}) && (A{2,2+2*z}>
A{4,2+2*z}) && (A{2,2+2*z}>
A{5,2+2*z}) && (A{2,2+2*z}>A{6,2+2*z}) && (A{2,2+2*z}>A{7,2+2*z})
        B(2)=B(2)+1;
    end
    if strcmp(A{8,2+2*z}, 'Sadness') && (A{3,2+2*z}> A{2,2+2*z}) && (A{3,2+2*z}>
A{1,2+2*z}) && (A{3,2+2*z}> A{4,2+2*z}) && (A{3,2+2*z}>
A{5,2+2*z}) && (A{3,2+2*z}>A{6,2+2*z}) && (A{3,2+2*z}>A{7,2+2*z})
        B(3)=B(3)+1;
    end
    if strcmp(A{8,2+2*z}, 'Neutral') && (A{4,2+2*z}> A{2,2+2*z}) && (A{4,2+2*z}>
A{3,2+2*z}) && (A{4,2+2*z}> A{1,2+2*z}) && (A{4,2+2*z}>
A{5,2+2*z}) && (A{4,2+2*z}>A{6,2+2*z}) && (A{4,2+2*z}>A{7,2+2*z})
        B(4)=B(4)+1;
    end
    if strcmp(A{8,2+2*z}, 'Fear') && (A{5,2+2*z}> A{2,2+2*z}) && (A{5,2+2*z}>
A{3,2+2*z}) && (A{5,2+2*z}> A{4,2+2*z}) && (A{5,2+2*z}>
A{1,2+2*z}) && (A{5,2+2*z}>A{6,2+2*z}) && (A{5,2+2*z}>A{7,2+2*z})
```

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

```
B(5)=B(5)+1;
end
if strcmp(A{8,2+2*z},'Disgust') && (A{6,2+2*z}> A{2,2+2*z}) && (A{6,2+2*z}>
A{3,2+2*z}) && (A{6,2+2*z}> A{4,2+2*z}) && (A{6,2+2*z}>
A{5,2+2*z}) && (A{6,2+2*z}>A{1,2+2*z}) && (A{6,2+2*z}>A{7,2+2*z})
B(6)=B(6)+1;
end
if strcmp(A{8,2+2*z},'Anger') && (A{7,2+2*z}> A{2,2+2*z}) && (A{7,2+2*z}>
A{3,2+2*z}) && (A{7,2+2*z}> A{4,2+2*z}) && (A{7,2+2*z}>
A{5,2+2*z}) && (A{7,2+2*z}>A{6,2+2*z}) && (A{7,2+2*z}>A{1,2+2*z})
B(7)=B(7)+1;
end
end
```

4.3.2 Κώδικας δεύτερης υλοποίησης

Για την δεύτερη υλοποίηση της πρώτης μεθόδου με την χρήση ιδιοδιανυσμάτων και της μικρότερης Ευκλείδειας απόστασης κατά την οποία γίνεται η εξαγωγή των 25 μέσων ιδιοδιανυσμάτων, ο κώδικας που θα δώσει τα μέσα ιδιοδιανύσματα για τις 7 εκφράσεις των εικόνων που θα χρησιμοποιηθούν ως «εκπαιδευτικές εικόνες» είναι ίδιος με αυτόν που βρίσκεται στο αρχείο **EigenvectorsTestVer001.m**, με την διαφορά ότι έχει κάποιες επιπλέον γραμμές κώδικα για να βρεθεί ο μέσος όρος των ιδιοδιανυσμάτων και το κελί που κρατάει τις ετικέτες δεν χρειάζεται οπότε διαγράφεται. Το αρχείο με τον κώδικα έχει ονομαστεί **EigenvectorsTrainVer002.m**.

Με μία διπλή φωλιασμένη `for`, `for i=2:Y1(1)` και `for j=1:Y1(2)` όπου `Y1=size(EmotionTemp);` το ιδιοδιάνυσμα που βρίσκεται στην γραμμή `i` και στην στήλη `j` προστίθεται στο ιδιοδιάνυσμα της θέσης 1 και στήλης `j` του κελιού `EmotionTemp` με την εντολή `EmotionTemp{1,j}=EmotionTemp{1,j}+EmotionTemp{i,j};`. Αυτό δίνει 25 σύνολα ιδιοδιανυσμάτων που το καθένα από αυτά πρέπει να διαιρεθεί με τον συνολικό αριθμό εικόνων που έχουν εισαχθεί σαν «εκπαιδευτικές» εικόνες. Ορίζεται ένα νέο κελί το `EmotionFinal`, με `EmotionFinal=cell(1,Y1(2));` και με μία `for` γίνεται η διαίρεση με τα αποτελέσματα να εισάγονται στο καινούργιο κελί `EmotionFinal`. Το κελί αποθηκεύεται με την χρήση της συνάρτησης `uisave` και ο κώδικας εκτελείται συνολικά 7 φορές μία για κάθε έκφραση. Το τμήμα του κώδικα που αλλάζει δίνεται παρακάτω.

```
Y1=size(EmotionTemp);
for i=2:Y1(1)
    for j=1:Y1(2)
        EmotionTemp{1,j}=EmotionTemp{1,j}+EmotionTemp{i,j};
    end
end

EmotionFinal=cell(1,Y1(2));

for j=1:Y1(2)
    EmotionFinal{1,j}=EmotionTemp{1,j}/Y1(1);
end

uisave('EmotionFinal');
```

Οι κώδικες που βρίσκουν τα ιδιοδιανύσματα για τις δοκιμαστικές εικόνες, τις Ευκλείδειες αποστάσεις για κάθε έκφραση και την μικρότερη ευκλείδεια απόσταση (που


συμπεριλαμβάνει και τον κώδικα καταμέτρησης των ψήφων) δεν αλλάζουν οπότε χρησιμοποιούνται τα αρχεία κώδικα **EigenvectorsTestVer001.m**, **findEuclideanDistancesVer001.m** και **importEuclideanDistances.m**.

4.4 Αποτελέσματα συστημάτων – Συζήτηση


4.4.1 Αποτελέσματα πρώτης υλοποίησης

Στους Πίνακες 2-6 δίνεται ένα παράδειγμα για το πώς εκτελείται η διαδικασία καταμέτρησης των ψήφων για μία δοκιμαστική εικόνα στην οποία κάθε περιοχή της έχει πέντε Ευκλείδειες αποστάσεις, οι οποίες είναι αποτέλεσμα των πέντε ιδιοδιανυσμάτων κάθε περιοχής και των πέντε ιδιοδιανυσμάτων εκπαίδευσης κάθε περιοχής. Επιλέχθηκαν τα πέντε μεγαλύτερα ιδιοδιανύσματα γιατί, όπως αναφέρθηκε και παραπάνω, σύμφωνα με τους J. Kalita και K. Das [2], αυτά είναι τα ιδιοδιανύσματα που κωδικοποιούν την πιο σημαντική πληροφορία στην εικόνα, όπως την αλλαγή των χαρακτηριστικών για να εκφραστεί ένα συναίσθημα.


Οι τιμές από κάθε στήλη ED (από το Euclidean Distance) που είναι πιο έντονα τονισμένες είναι οι μικρότερες ευκλείδειες αποστάσεις. Ο Πίνακας 7 παρουσιάζει την τελική καταμέτρηση των ψήφων από όλες τις περιοχές για κάθε συναίσθημα και το συναίσθημα με τις περισσότερες ψήφους είναι αυτό που εκφράζει η εικόνα.

Testing Image	Emotion	ED1	ED2	ED3	ED4	ED5
	Surprise	1.9906	1.2516	1.3464	1.6667	0.9981
	Happiness	1.9905	1.1143	1.4732	1.1612	1.5228
	Fear	1.9833	1.1089	1.2353	1.6888	1.5514
	Anger	1.9815	1.0962	1.4485	1.4953	1.6974
	Sadness	1.9870	1.6690	1.5015	0.9589	1.2232
	Disgust	1.9860	1.0365	1.2941	1.4568	1.1461
	Neutral	1.9899	1.1308	1.2717	1.6973	1.6224


Πίνακας 2: Οι πέντε Ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το αριστερό μάτι μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.

Testing Image	Emotion	ED1	ED2	ED3	ED4	ED5
	Surprise	1.9306	1.5698	1.5921	1.2060	1.7815
	Happiness	1.9457	1.5581	1.0398	1.6077	1.5053
	Fear	1.9308	1.5897	1.0640	1.2664	1.0894
	Anger	1.9402	1.5598	1.7803	1.4225	1.5214
	Sadness	1.9392	1.5482	1.0158	0.9372	1.6930
	Disgust	1.9353	1.6185	1.7467	1.4511	1.4544
	Neutral	1.9386	1.5340	0.9024	1.2841	1.6304


Πίνακας 3: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το δεξί μάτι μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.

Testing Image	Emotion	ED1	ED2	ED3	ED4	ED5
	Surprise	1.4593	1.3752	1.8279	1.1704	1.1852
	Happiness	1.4779	1.4197	1.4926	1.4946	1.3709
	Fear	1.4823	1.4214	1.6348	1.2584	1.5771
	Anger	1.4961	1.5042	1.5958	1.4181	1.3163


Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

	Sadness	1.4874	1.4027	1.4058	1.3428	1.6014
	Disgust	1.4792	1.4207	1.4854	1.3985	1.5896
	Neutral	1.4956	1.4112	1.6227	1.2421	1.4619


Πίνακας 4: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το στόμα μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.

Testing Image	Emotion	ED1	ED2	ED3	ED4	ED5
	Surprise	0.4180	0.4098	1.9189	1.4117	1.1135
	Happiness	0.4153	1.0391	1.0671	1.7652	1.5740
	Fear	0.3981	1.0904	1.1791	1.7025	1.5842
	Anger	0.3722	0.7364	0.8564	1.0046	1.6147
	Sadness	1.9673	0.7806	0.8887	1.6600	1.6041
	Disgust	1.8749	1.5480	1.4308	1.3408	1.3028
	Neutral	0.3800	0.5594	0.7001	1.0333	1.2133

Πίνακας 5: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για την μύτη μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.

Testing Image	Emotion	ED1	ED2	ED3	ED4	ED5
	Surprise	1.7810	1.3073	1.5672	1.5387	1.3583
	Happiness	1.7524	1.3247	1.4047	1.2294	1.4322
	Fear	1.7584	1.4305	1.4850	1.4989	1.3485
	Anger	1.7611	1.4696	0.9757	1.5303	1.1302
	Sadness	1.7803	1.3755	1.0045	1.3474	1.4568
	Disgust	1.8847	1.3910	1.4598	1.5602	1.5593
	Neutral	1.7786	1.5746	1.2822	1.4373	1.4444

Πίνακας 6: Οι πέντε ευκλείδειες αποστάσεις για όλα τα συναισθήματα για το στόμα και την μύτη μίας δοκιμαστικής εικόνας. Οι τιμές που έχουν έντονο μαύρο είναι οι μικρότερες ευκλείδειες αποστάσεις.

Test Image	Features	Number of votes for the selected features							
		Emotion	Surprise	Happiness	Fear	Anger	Sadness	Disgust	Neutral
	Left eye		1	0	1	1	1	1	0
	Right eye		1	0	0	1	1	0	2
	Mouth		4	0	0	0	0	1	0
	Nose		2	0	0	2	0	0	1
	MouthNose		1	2	0	2	0	0	0
	Total		9	2	1	6	2	2	3

Πίνακας 7: Ο τελικός πίνακας ψήφων που μετράει τις ψήφους για κάθε συναίσθημα ανά περιοχή και αποφασίζει πιο είναι το συναίσθημα που εκφράζει η δοκιμαστική εικόνα.

Στους Πίνακες 8-10 δίνονται τα αποτελέσματα των δοκιμών που έγιναν σε 886 εικόνες για 25, 75 και 150 ευκλείδειες αποστάσεις. Όπως φαίνεται και στους πίνακες το ποσοστό επιτυχούς αναγνώρισης είναι πολύ χαμηλό ξεπερνώντας ελάχιστα το 20% στην δοκιμή με τις 150 ευκλείδειες αποστάσεις (30 αποστάσεις ανά περιοχή). Μια τελευταία προσπάθεια αποτυπώνεται στον Πίνακα 11, όπου τα πέντε ιδιοδιανύσματα που συσχετίζονται με τις πέντε μεγαλύτερες ιδιοτιμές αφαιρέθηκαν σε περίπτωση που κωδικοποιούν πληροφορία όπως φωτεινότητα [19] και τα επόμενα 20 χρησιμοποιήθηκαν για την εξαγωγή των ευκλείδειων αποστάσεων και τελικά για την καταμέτρηση των ψήφων.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	43,65	34.78	6.66	8.78	6.20	12.3	13.26	19.30
Επιτυχία (εικόνες)	55	64	8	13	8	10	13	171
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 8: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 25 ευκλείδειες αποστάσεις.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	42.85	10.86	13.33	6.75	11.62	41.97	6.12	17.49
Επιτυχία (εικόνες)	54	20	16	10	15	34	6	155
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 9: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 75 ευκλείδειες αποστάσεις.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	34.92	28.26	19.16	12.16	9.30	29.62	20.40	21.78
Επιτυχία (εικόνες)	44	52	23	18	12	24	20	193
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 10: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 150 ευκλείδειες αποστάσεις.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	23.01	15.21	14.16	6.08	9.30	27.16	19.38	15.57
Επιτυχία (εικόνες)	29	28	17	9	12	24	19	138
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 11: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 100 ευκλείδειες αποστάσεις χωρίς τα πέντε πρώτα ιδιοδιανύσματα κάθε περιοχής.

4.4.2 Αποτελέσματα δεύτερης υλοποίησης

Τα αποτελέσματα της δεύτερης υλοποίησης με την χρήση του «μέσων ιδιοδιανυσμάτων» ήταν σχετικά καλύτερα με συνολική βελτίωση στις 25 Ευκλείδειες αποστάσεις της τάξης του 9% σε σχέση με την πρώτη υλοποίηση, όπως φαίνεται στους Πίνακες 12-14.

Στις 75 ευκλείδειες αποστάσεις, τα αποτελέσματα ήταν ελάχιστα χαμηλότερα σε σχέση με τις 25 αποστάσεις της ίδιας μεθόδου, αλλά και πάλι καλύτερα σε σχέση με της πρώτης μεθόδου που ήταν περίπου 11% χαμηλότερα. Τέλος το ποσοστό επιτυχούς αναγνώρισης στις 150 αποστάσεις ήταν χαμηλότερο σε σχέση με τα αποτελέσματα της πρώτης μεθόδου κατά 1.5%.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	65.07	31.52	22.50	31.75	3.87	24.69	12.24	28.32
Επιτυχία (εικόνες)	82	58	27	47	5	20	12	251
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 12: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 25 ευκλείδειες αποστάσεις.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	52.38	25.00	22.50	20.94	20.93	41.97	17.34	27.99
Επιτυχία (εικόνες)	66	46	27	31	27	34	17	248
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 13 Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 75 ευκλείδειες αποστάσεις.

	Έκπληξη	Ευτυχία	Λύπη	Ουδέτερη	Φόβος	Αηδία	Θυμός	Σύνολο
Επιτυχία (%)	34.92	10.32	27.50	6.08	17.05	34.56	24.48	20.20
Επιτυχία (εικόνες)	44	19	33	9	22	28	17	179
Συνολικές Εικόνες	126	184	120	148	129	81	98	886

Πίνακας 14: Ο πίνακας με το ποσοστό % και ποσοστό σε εικόνες για όλες τις σωστά αναγνωρισμένες δοκιμαστικές εικόνες με 150 ευκλείδειες αποστάσεις.

4.4.3 Συζήτηση αποτελεσμάτων για τις δύο υλοποιήσεις

Τα αποτελέσματα που βρέθηκαν και με τις δύο υλοποιήσεις από τα συστήματα με χρήση χαρακτηριστικών που βασίζονται σε ιδιοδιανύσματα και την χρήση την Ευκλείδειας απόστασης σαν τεχνική λήψης αποφάσεων δεν ήταν όσο καλά όσο αναμενόταν. Ενώ οι J. Kalita και K. Das [2] στην ερευνά τους πέτυχαν ποσοστό αναγνώρισης 95%, στη παρούσα

εργασία το ποσοστό επιτυχούς αναγνώρισης των εκφράσεων ήταν τόσο χαμηλό που θα μπορούσε να θεωρηθεί τυχαίο.

Μια ερμηνεία του γεγονότος αυτού είναι ότι μπορεί να δουλεύει καλά ένα τέτοιο σύστημα όταν υπάρχει μικρός όγκος δεδομένων, λιγότερες από 100 εικόνες για παράδειγμα όπως χρησιμοποίησαν οι Kalita και Das και να μην δουλεύει καλά για μεγάλο όγκο όπως 2.500 εικόνες. Μία δεύτερη ερμηνεία είναι ότι επειδή οι εικόνες δεν ευθυγραμμίστηκαν όλες μεταξύ τους με κάποιο αλγόριθμο ώστε π.χ. η κόρη του αριστερού ματιού να είναι ακριβώς στην ίδια θέση για όλες τις εικόνες, είχε σαν αποτέλεσμα να επηρεαστούν αρνητικά τα δεδομένα. Ακόμη και αν αυτό ίσχυε βέβαια, θα περιμέναμε ένα ποσοστό αναγνώρισης 50%-60% και όχι τόσο κοντά στο μηδέν.

Επίσης θα μπορούσε να υπάρχει κάποια διαφοροποίηση στην υλοποίηση του αλγορίθμου που χρησιμοποίησαν οι J. Kalita και K. Das, καθώς στην δημοσίευσή τους υπήρχαν κάποιες ασάφειες στον τρόπο υλοποίησης του συστήματος. Ο τρόπος που εισάγονται τα δεδομένα για επεξεργασία καθώς και ο όγκος των δεδομένων που χρησιμοποιούν δεν ορίζεται ξεκάθαρα (αναφέρεται μόνο ότι δοκιμάζουν το σύστημα με 10 εικόνες για κάθε έκφραση, συνολικά δηλαδή 60 εικόνες). Τέλος το χαμηλό ποσοστό επιτυχίας αναγνώρισης των εκφράσεων πιθανά να οφείλεται στο γεγονός ότι αυτό το σύστημα δουλεύει μόνο υπό κατάλληλες συνθήκες σε ελεγχόμενο περιβάλλον και δεν μπορεί να αποδώσει σε δεδομένα με διακυμάνσεις που πιθανά να υπάρχουν στο υπό ανάλυση σύνολο δεδομένων.

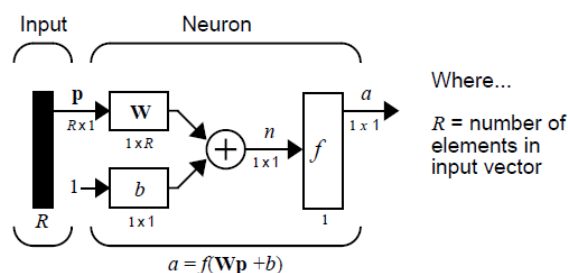
5. Μέθοδος 2: Τεχνητά Νευρωνικά Δίκτυα

Όπως έγινε αναφορά και στο κεφάλαιο 3.2 η δεύτερη υλοποίηση του αυτόματου συστήματος αναγνώρισης συναισθημάτων μπορεί να γίνει με την χρήση των τεχνητών νευρωνικών δικτύων. Τα τεχνητά νευρωνικά δίκτυα χρησιμοποιούνται ολοένα και περισσότερο για την αυτόματη αναγνώριση συναισθημάτων γιατί είναι ευέλικτα, εύκολα στην κατανόηση και προσφέρουν υψηλά ποσοστά αναγνώρισης αν εκπαιδευτούν σωστά.

5.1 Θεωρητικό υπόβαθρο

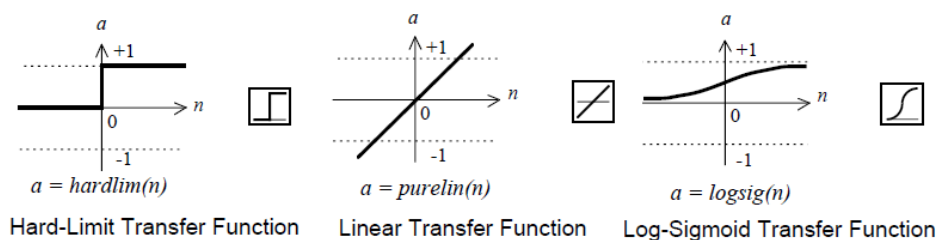
Ένα τεχνητό νευρωνικό δίκτυο είναι μία προσπάθεια προσομοίωσης της λειτουργίας του ανθρώπινου εγκεφάλου σε υπολογιστικό περιβάλλον. Όπως ο ανθρώπινος εγκέφαλος αποτελείται από ένα σύνολο από νευρώνες (neurons) που δουλεύουν παράλληλα/ταυτόχρονα και παράγουν συνάψεις (synapses) ώστε να στείλουν ηλεκτρικό ή χημικό σήμα σε γειτονικούς νευρώνες [48], έτσι και στο τεχνητό νευρωνικό δίκτυο υπάρχει ένα σύνολο στοιχείων που τα λένε τεχνητούς νευρώνες που αποστέλλουν στους γειτονικούς νευρώνες σήματα πληροφορίας για την επεξεργασία τους, ώστε να έχουν σαν αποτέλεσμα μια λογική έξοδο.

Οι νευρώνες είναι πολλοί σε αριθμό και συνδέονται μεταξύ τους, ενώ οι συνάψεις, που στα τεχνητά νευρωνικά δίκτυα λέγονται βάρη (weights), προστίθενται στην είσοδο καθώς περνάει από νευρώνα σε νευρώνα για να δώσουν ένα λογικό τελικό αποτέλεσμα. Αυτό που κάνει ο κάθε νευρώνας, που είναι μια απλή μονάδα επεξεργασίας, είναι να παίρνει τις εισόδους και να προσθέτει το βάρος που συσχετίζεται με τον συγκεκριμένο νευρώνα και μία σταθερά που λέγεται bias. Αφού βρεθεί το αποτέλεσμα της πράξης μεταφέρεται μέσω μίας συνάρτησης ενεργοποίησης στον επόμενο στρώμα νευρώνων ή στην τελική έξοδο. Η υλοποίηση ενός νευρώνα σχηματικά φαίνεται στην Εικόνα 5.



Εικόνα 7: Παράδειγμα ενός απλού νευρώνα.

Κάθε τεχνητό νευρωνικό δίκτυο αποτελείται από πολλούς τεχνητούς νευρώνες και είναι συνδεδεμένοι μεταξύ τους σύμφωνα με κάποια συγκεκριμένη αρχιτεκτονική. Οι νευρώνες κάθε αρχιτεκτονικής μπορούν να έχουν διαφορετικές συναρτήσεις ενεργοποίησης όπως αυτές που φαίνονται στην Εικόνα 6 που είναι είτε γραμμικές είτε μη γραμμικές και έναν αριθμό από κρυφά στρώματα (hidden layers). Ένα κρυφό στρώμα μαθαίνει να επανακωδικοποιεί τις εισόδους και μία αρχιτεκτονική νευρωνικών δικτύων με περισσότερα κρυφά στρώματα είναι συνήθως πιο ισχυρή υπολογιστικά από μια αρχιτεκτονική με μόνο ένα στρώμα.



Εικόνα 8: Κάποιες συναρτήσεις ενεργοποίησης που χρησιμοποιούνται στα νευρωνικά δίκτυα

Υπάρχουν νευρωνικά δίκτυα με τροφοδότηση προς τα εμπρός (feed-forward neural networks) και με ανάδραση/ανατροφοδότηση (feedback) [49] [50]. Τα feed-forward δίκτυα επιτρέπουν την αποστολή των δεδομένων μόνο προς τα εμπρός, από την είσοδο στην έξοδο και η έξοδος ενός νευρώνα δεν μπορεί να επηρεάσει τον ίδιο. Τα feedback δίκτυα στέλνουν αρχικά τα δεδομένα προς τα εμπρός και μετά τα δεδομένα εξόδου μπορούν να σταλούν προς τα πίσω σε οποιοδήποτε από τα άλλα στρώματα. Επειδή μπορούν να στείλουν τα δεδομένα εξόδου και στις δύο κατευθύνσεις δημιουργώντας βρόγχους γίνονται αρκετά πολύπλοκα. Είναι δυναμικά και αλλάζουν συνεχώς μέχρι να φτάσουν σε μια κατάσταση ισορροπίας, όπου μένουν εκεί μέχρι να αλλάξει η είσοδος και να πρέπει να βρεθεί μια νέα κατάσταση ισορροπίας.

Το σημαντικότερο πλεονέκτημα των τεχνητών νευρωνικών δικτύων είναι ότι μπορούν να εκπαιδευτούν σε ένα συγκεκριμένο τύπο δεδομένων και να μάθουν να κατηγοριοποιούν επιτυχώς νέα δεδομένα χωρίς αρχικά να ξέρουν σε πια κατηγορία ανήκουν. Η εκπαίδευση μπορεί να γίνει με δύο τρόπους, με επίβλεψη (supervised learning) ή χωρίς επίβλεψη (unsupervised learning). Στην εκπαίδευση με επίβλεψη δίνονται στο δίκτυο οι εισόδους και οι εξόδους. Το δίκτυο επεξεργάζεται τις εισόδους και συγκρίνει τα αποτελέσματα με τις επιθυμητές εξόδους. Υπολογίζει το σφάλμα της κάθε εξόδου, δηλαδή την απόκλιση που έχει η τελική έξοδος σε σχέση με την επιθυμητή έξοδο και με αυτό το σφάλμα προσαρμόζονται τα βάρη του δικτύου. Η διαδικασία επαναλαμβάνεται μέχρι το σφάλμα να γίνει μηδενικό ή να φτάσει πολύ κοντά στο μηδέν ή να τελειώσει ο αριθμός των επαναλήψεων που έχει οριστεί. Κατά την διάρκεια της εκπαίδευσης τα ίδια δεδομένα επεξεργάζονται ξανά και ξανά μέχρι τα βάρη να φτάσουν σε ένα επιθυμητό αποτέλεσμα. Στην εκπαίδευση χωρίς επίβλεψη το δίκτυο εφοδιάζεται μόνο με τις εισόδους και πρέπει να βρει από μόνο του ποια χαρακτηριστικά θα χρησιμοποιήσει για να ομαδοποιήσει (clustering) τα δεδομένα εισόδου που θεωρεί ότι είναι ίδιου τύπου [50]. Αυτό συνήθως αναφέρεται σαν αυτό-οργάνωση ή προσαρμογή.

Ο σχεδιασμός του τεχνητού νευρωνικού δικτύου γίνεται ακολουθώντας τα εξής βήματα. Αρχικά δίνονται κάποιες εισόδους που μπορεί να έχουν επεξεργαστεί ή όχι και μετά αποφασίζεται η αρχιτεκτονική του δικτύου. Το πρώτο στρώμα καθορίζεται από τον αριθμό των εισόδων, μετά καθορίζονται τα κρυφά στρώματα, πόσα θα είναι και με πόσους νευρώνες το κάθε ένα (κάθε κρυφό στρώμα μπορεί να έχει διαφορετικό αριθμό νευρώνων από τα υπόλοιπα) και τέλος το στρώμα εξόδου που συνήθως ακολουθεί τον κανόνα, αν υπάρχουν X κατηγορίες (κλάσεις) δεδομένων τότε το στρώμα εξόδου θα έχει X εξόδους. Η συνάρτηση μεταφοράς καθορίζεται για κάθε στρώμα όταν αποφασίζεται η αρχιτεκτονική του δικτύου.

Σαν πρώτο βήμα της εκπαίδευσης του δικτύου είναι να περάσουν τα δεδομένα εισόδου από το νευρωνικό δίκτυο για πρώτη φορά, ώστε να υπολογιστεί το σφάλμα, που δίνεται αφαιρώντας την πραγματική έξοδο από την επιθυμητή έξοδο και να γίνει χρήση του σφάλματος αυτού για να ρυθμιστούν τα βάρη μέσω κάποιου ρυθμού μάθησης (learning rate). Η πλέον διαδεδομένη μέθοδος ενημέρωσης των βαρών είναι η back-propagation. Η

διαδικασία αυτή επαναλαμβάνεται για έναν αριθμό επαναλήψεων που λέγονται εποχές (epochs) μέχρι το σφάλμα να φτάσει μία επιθυμητή τιμή ή να ολοκληρωθεί ο αριθμός των εποχών που έχουν οριστεί.

Οι εποχές μαζί με μία άλλη παράμετρο που λέγεται έλεγχος επαλήθευσης (validation check) χρησιμοποιούνται για να αποτρέψουν το νευρωνικό δίκτυο να μπει σε ατέρμον βρόχο σε περίπτωση που δεν καταφέρει να φτάσει το επιθυμητό σφάλμα, δηλαδή αν το επιθυμητό σφάλμα δεν είναι εφικτό να σταματήσουν οι επαναλήψεις και να μην συνεχιστούν για πάντα. Το επιθυμητό σφάλμα ορίζεται συνήθως μηδέν ή μία τιμή πολύ κοντά στο μηδέν, ώστε οι πραγματικές εξοδοί των δεδομένων εκπαίδευσης να έχουν όσο το δυνατόν πιο κοντινή τιμή γίνεται στις επιθυμητές εξόδους. Ο έλεγχος επαλήθευσης παρακολουθεί την αλλαγή του σφάλματος και αν εντοπίσει ότι η τιμή του αλλάζει ελάχιστα ή αυξομειώνεται γύρω από τις ίδιες τιμές για X εποχές, που X ορίζεται από τον χρήστη, τότε σταματάει την εκπαίδευση του δικτύου. Ο ρυθμός μάθησης, που αναφέρθηκε στην προηγούμενη παράγραφο, είναι ένα βήμα που καθορίζει το πόσο γρήγορά ή αργά θα εκπαιδεύεται ένα δίκτυο. Αν είναι πολύ μεγάλος, τότε το δίκτυο θα εκπαιδεύεται πολύ γρήγορα, αλλά και πιθανώς εσφαλμένα, ενώ αν είναι πολύ μικρός, τότε το δίκτυο θα εκπαιδεύεται πολύ αργά, γιατί τα βάρη θα αλλάζουν ελάχιστα και το σφάλμα δεν θα φτάνει την επιθυμητή τιμή. Η σωστή επιλογή του ρυθμού μάθησης είναι σημαντική για τη σωστή λειτουργία του νευρωνικού δικτύου.

Στα τεχνητά νευρωνικά δίκτυα δεν υπάρχει συγκεκριμένη μεθοδολογία για το πως θα επιτευχθεί η υλοποίηση ενός συστήματος για ένα συγκεκριμένο σκοπό. Ανάλογα με το πρόβλημα που αντιμετωπίζεται, το μέγεθος των δεδομένων και την εμπειρία που έχει στον τομέα των τεχνητών νευρωνικών δικτύων ο προγραμματιστής πρέπει να αποφασίσει για την επιλογή της συνάρτησης μεταφοράς, των εισόδων που χρειάζεται το δίκτυο καθώς και των εξόδων, των κρυμμένων στρωμάτων και των νευρώνων ανά στρώμα. Υπάρχουν βέβαια κάποιοι γενικοί τρόποι υλοποίησης ενός συστήματος, αλλά οι τελικές τιμές για τα παραπάνω πεδία δίνονται μετά από διαδικασία δοκιμής και λάθους (trial and error procedure), που σημαίνει ότι με πολλές δοκιμές σταδιακά προσαρμόζονται οι παραμέτροι του συστήματος ώστε να βρεθούν τα καλύτερα δυνατά αποτελέσματα.

5.2 Σχεδιασμός υλοποίησης – Αλγόριθμος των τριών υλοποιήσεων

Για την δεύτερη μέθοδο, με την χρήση τεχνητών νευρωνικών δικτύων, σαν πρώτο βήμα γίνεται ο διαχωρισμός της βάσης δεδομένων σε 3 σύνολα. Ένα σύνολο συναισθημάτων και άλλα δύο σύνολα από «δοκιμαστικές» εικόνες, που περιέχουν ένα ποσοστό εικόνων από κάθε συναίσθημα όπως φαίνεται στον Πίνακα 15 (βήμα 1). Το πρώτο σύνολο θα χρησιμοποιηθεί για την εκπαίδευση του νευρωνικού δικτύου (training set). Το ένα σύνολο «δοκιμαστικών» εικόνων χρησιμοποιείται μόνο για τον έλεγχο της τελικής λύσης, προκειμένου να επιβεβαιώσει την πραγματική προβλεπτική ικανότητα του δικτύου (testing set). Το δεύτερο σύνολο λέγεται και σύνολο επικύρωσης (validation set) και αυτό το σύνολο δεδομένων χρησιμοποιείται για την ελαχιστοποίηση υπερπροσαρμογής (overfitting). Δε γίνεται προσαρμογή των βαρών του δικτύου με αυτό το σύνολο δεδομένων, αλλά εξακριβώνεται ότι οποιαδήποτε αύξηση της ακρίβειας επί των δεδομένων εκπαίδευσης, πραγματικά δίνει μια αύξηση της ακρίβειας σε ένα σύνολο δεδομένων που δεν έχει δει ποτέ το δίκτυο πριν, ή τουλάχιστον το δίκτυο δεν έχει εκπαιδευτεί σε αυτό το σύνολο. Εάν η ακρίβεια κατά τη διάρκεια της εκπαίδευσης αυξάνεται, αλλά η ακρίβεια πάνω στο σύνολο δεδομένων επικύρωσης παραμένει η ίδια ή μειώνεται, τότε γίνεται υπερπροσαρμογή του νευρωνικού δικτύου. Σε αυτή την περίπτωση θα πρέπει να σταματήσει η εκπαίδευση, να

αναλυθούν τα αποτελέσματα και να αναθεωρηθεί η αρχιτεκτονική του δικτύου αν χρειάζεται.

Συναίσθημα	Έκπληξη	Λύπη	Φόβος	Θυμός	Αηδία	Χαρά	Ουδέτερο
Συνολικά Στιγμιότυπα	377	308	299	243	281	440	470
Συνολικά Άτομα	78	59	58	48	60	80	91
Στιγμιότυπα για Training	251	188	170	145	200	256	322
Ποσοστό για Training	66.5%	61%	56.8%	59.6%	71.1%	58.1%	68.5%
Στιγμιότυπα για Validation	52	48	53	55	41	81	60
Ποσοστό για Validation	13.7%	15.5%	17.7%	22.6%	14.5%	18.4%	12.7%
Στιγμιότυπα για Testing	74	72	76	43	40	103	88
Ποσοστό για Testing	19.6%	23.3%	25.4%	17.6%	14.2%	23.4%	18.7%

Πίνακας 15: Ποσοστά και αριθμός εικόνων που χρησιμοποιήθηκαν για εκπαίδευση, επικύρωση και δοκιμή του νευρωνικού δικτύου για τις δύο πρώτες υλοποιήσεις.

Πριν δημιουργηθεί το νευρωνικό δίκτυο πρέπει να γίνει μία προ-επεξεργασία στις εικόνες της βάσης δεδομένων για να μπορούν να εισαχθούν στο δίκτυο. Αρχικά οι εικόνες εισάγονται μία προς μία στην επιφάνεια εργασίας του matlab για να αρχίσει η επεξεργασία τους. Η εισαγωγή τους γίνεται μέσω interface, δηλαδή με εντολή του matlab θα ανοίγει παράθυρο για να γίνει η επιλογή ενός φακέλου, όπου αυτός ο φάκελος θα περιέχει όλες τις εικόνες, για την εκπαίδευση, την επικύρωση και τη δοκιμή (βήμα 2). Το πρόγραμμα θα πρέπει να είναι σχεδιασμένο κατάλληλα, ώστε να διαβάζει μόνο αρχεία εικόνων, όπως png, tif, jpg κτλ (βήμα 3). Ορίζεται ένας πίνακας που κρατεί μία κωδικοποιημένη ετικέτα για κάθε εικόνα που επεξεργάζεται το matlab ανάλογα με το συναίσθημα που αντιπροσωπεύει (βήμα 4). Με την χρήση του Viola-Jones αλγόριθμου (βλ.4.2.1) γίνεται η ανίχνευση του προσώπου μέσα στην εικόνα (βήμα 5). Μετά η ανιχνευμένη εικόνα αποκόπτεται από την αρχική με τελικό μέγεθος 300x300 pixels (βήμα 6). Η εικόνα με αρχικές διαστάσεις 300x300 μετατρέπεται σε στήλη μεγέθους 1x90000 και εισάγεται σε έναν πίνακα (Face) (βήμα 7). Όλες οι εικόνες από κάθε σύνολο (training, validation, testing) θα μπουν στον ίδιο πίνακα αφού ο αριθμός των εικόνων ανά σύνολο είναι γνωστός, τότε το νευρωνικό δίκτυο μπορεί να χρησιμοποιήσει τον ίδιο πίνακα για διαφορετικό σκοπό. Τέλος, αφού έχουν εισαχτεί όλες οι εικόνες, ο πίνακας Face και ο πίνακας με τις ετικέτες που ονομάζεται RealLabels αποθηκεύονται για μελλοντική χρήση όταν κατασκευαστεί το νευρωνικό δίκτυο (βήμα 8).

Το τεχνητό νευρωνικό δίκτυο που υλοποιήθηκε είναι ένα δίκτυο που εκπαιδεύεται με επίβλεψη (supervised learning) αφού του δίνονται οι εισοδοί (ο πίνακας Face) και οι εξοδοί (ο πίνακας RealLabels), ενώ ο τρόπος εκπαίδευσης αρχικά γίνεται με scaled conjugate gradient backpropagation που είναι τρόπος εκπαίδευσης με οπισθοδιάδοση (back-propagation). Ο στόχος της εκπαίδευσης είναι να βρεθεί μηδενικό λάθος (performance goal). Η εκπαίδευση γίνεται για 1000 εποχές (epochs), χωρίς να υπάρχει χρονικός περιορισμός, με ελάχιστη βαθμίδα απόδοσης ίση με $1e-6$ (minimum performance gradient) και μέγιστη αποτυχία επικύρωσης ίση με 10 (maximum validation failures).

Οι πρώτες δοκιμές που έγιναν ήταν για τα επτά συναισθήματα, χωρίς να έχουν υποβληθεί σε κάποια επεξεργασία εκτός από την μετατροπή των προσώπων σε στήλη. Αυτό σημαίνει ότι αρχικά το νευρωνικό δίκτυο θα έχει σαν πρώτο στρώμα από νευρώνες 90.000

εισόδους και συνολικά 2418 εικόνες. Από αυτές οι πρώτες 1532 θα χρησιμοποιηθούν για την εκπαίδευση (training) του τεχνητού νευρωνικού δικτύου, οι 390 από αυτές για την επικύρωση (validation) και οι υπόλοιπες 496 για δοκιμές (testing). Το matlab έχει ειδικό toolbox για την ευκολότερη υλοποίηση τεχνητών νευρωνικών δικτύων.

Τα κρυμμένα στρώματα θα είναι από ένα ως δύο και ο αριθμός των νευρώνων ανά στρώμα θα διαφέρει από δοκιμή σε δοκιμή. Δεν είναι γνωστό από πριν ποιο θα είναι το πιο αποδοτικό δίκτυο και ο μόνος τρόπος για να βρεθεί είναι με πάρα πολλές δοκιμές και συνεχείς αλλαγές στα κρυμμένα στρώματα. Θα δοκιμαστούν κρυμμένα στρώματα που θα έχουν από 10 έως 150 νευρώνες το καθένα μέχρι να βρεθεί αυτό που προσφέρει τα καλύτερα αποτελέσματα.

Το στρώμα εξόδου έχει 7 νευρώνες όπου ανάλογα με το αποτέλεσμα έξι από αυτούς θα δίνουν έξοδο κοντά στο μηδέν και ένας από αυτούς θα δίνει έξοδο κοντά στο ένα. Ανάλογα με την σειρά των τιμών από τις εξόδους, ένα από τα 7 συναισθήματα αντιπροσωπεύεται. Για την έκφραση της έκπληξης το νευρωνικό δίκτυο πρέπει να έχει έξοδο '000001', για την έκφραση της ευτυχίας '000010', για την έκφραση της λύπης '0000100', για την ουδέτερη έκφραση '0001000', για την έκφραση του φόβου '0010000', για την έκφραση της αηδίας '0100000' και για την έκφραση του θυμού '1000000'.

Όπως αναφέρθηκε παραπάνω οι εικόνες δεν έχουν υποβληθεί σε περαιτέρω επεξεργασία εκτός από την μετατροπή τους από δύο διαστάσεις σε μία. Ο λόγος που έγινε αυτό είναι για να διαπιστωθεί πόσο καλά μπορεί να αποδώσει το νευρωνικό δίκτυο αν του δοθούν ακατέργαστα δεδομένα και πόσο μπορεί να βελτιωθεί η ακρίβεια του μόλις υποβληθούν τα ίδια δεδομένα σε επεξεργασία. Τρεις διαφορετικές υλοποιήσεις πραγματοποιήθηκαν για την βελτίωση της ακρίβειας του νευρωνικού δικτύου στην παρούσα πτυχιακή εργασία. Χρήση φίλτρου Gabor για την δημιουργία τράπεζας Gabor. Χρήση φίλτρου μέσης τιμής σε συνδυασμό με το φίλτρο Gabor και της τράπεζας Gabor και τέλος χρήση της ουδέτερης έκφρασης σαν φίλτρο. Στην τελευταία μέθοδο το νευρωνικό δίκτυο θα μπορεί να βρίσκει μόνο έξι εκφράσεις καθώς η ουδέτερη έκφραση χρησιμοποιείται σαν φίλτρο.

Το φίλτρο Gabor είναι ένα γραμμικό φίλτρο που χρησιμοποιείται για ανίχνευση γωνιών. Η αναπαράσταση της συχνότητας (frequency) και του προσανατολισμού (orientation) του φίλτρου Gabor είναι παρόμοιες με εκείνες του ανθρώπινου οπτικού συστήματος και έχει βρεθεί ιδιαίτερα κατάλληλο για την αναπαράσταση και διάκριση υφής [51]. Η τράπεζα Gabor αποτελείται από τους μοναδικούς συνδυασμούς του κάθε προσανατολισμού και του κάθε μήκους κύματος, όπου μήκος κύματος ημιτονοειδούς ορίζεται ως το αριθμητικό διάνυσμα σε pixels ανά κύκλο. Για παράδειγμα εάν το μήκος κύματος είναι ένα διάνυσμα μήκους δύο και ο προσανατολισμός είναι ένα διάνυσμα μήκους τρία (δηλαδή με τρεις προσανατολισμούς), τότε η έξοδος θα είναι ένα διάνυσμα (μια τράπεζα Gabor) μήκους έξι [52]. Η επιλογή των διανυσμάτων μήκους κύματος και προσανατολισμού για μέγιστη απόδοση του νευρωνικού δικτύου είναι σημαντική.

Το φίλτρο μέσης τιμής (δεύτερη υλοποίηση) έχει οριστεί στο κεφάλαιο 4 και εδώ χρησιμοποιείται πάλι για να αφαιρεθούν οι ακραίες τιμές από την εικόνα πριν γίνει χρήση του φίλτρου Gabor. Η τελευταία υλοποίηση κάνει χρήση της ουδέτερης έκφρασης ενός ατόμου σαν φίλτρο αφαιρώντας την από όλες τις εκφράσεις που παρουσιάζει με αποτέλεσμα μίας εικόνας όπως φαίνεται στις Εικόνες 7 και 8 παρακάτω. Την χρήση μιας τέτοιας τεχνικής κάνουν οι Mohammadi et al. [16] για μια δική τους διαφορετική υλοποίηση, όπου για τα έξι βασικά συναισθήματα το ποσοστό επιτυχούς αναγνώρισης έφτασε το 95.67%.



Εικόνα 9: Το αποτέλεσμα της διαφοράς μεταξύ ουδέτερης έκφρασης και του συναισθήματος της έκπληξης.



Εικόνα 10: Το αποτέλεσμα της διαφοράς μεταξύ ουδέτερης έκφρασης και του συναισθήματος της αηδίας.

5.3 Υλοποίηση συστήματος – Κώδικας των τριών υλοποιήσεων

5.3.1 Κώδικας πρώτης υλοποίησης

Η πρώτη υλοποίηση με τεχνητά νευρωνικά δίκτυα ήταν για να προσδιοριστεί με πόση ακρίβεια μπορεί το δίκτυο να αναγνωρίσει τις επτά εκφράσεις του προσώπου αν δεν έχει υποβληθεί η εικόνα σε εκτενή επεξεργασία. Το αρχείο για την αρχική επεξεργασία των εικόνων λέγεται **Method1_NeuralNetworkSet.m**. Όπως και στις υπόλοιπες μεθόδους γίνεται χρήση του κώδικα για να διαβαστούν οι εικόνες στην μνήμη του matlab. Αυτή την φορά δεν χρειάζεται να εκτελεστεί ο κώδικας με μία `for i=1:7` γιατί ο κώδικας θέλει σαν είσοδο τον φάκελο EMOTIONS, που περιέχει τους υποφακέλους Surprise, Sadness, Anger, Neutral, Happiness, Disgust, Fear, TestData και ValidationData. Οι φάκελοι TestData και ValidationData περιέχουν σαν υποφακέλους πάλι έναν φάκελο για κάθε συναίσθημα.

Δηλώνεται ένα κελί, το `Labels{ii,1}=thisFolder;` που παίρνει σε κάθε επανάληψη το όνομα της διαδρομής του φακέλου στον οποίο βρίσκεται αυτή την στιγμή ο κώδικας. Με την χρήση της συνάρτησης `strfind()` συγκρίνουμε τα περιεχόμενα του `Labels{ii,1}` με μία συμβολοσειρά για να κωδικοποιήσουμε την ετικέτα για κάθε συναίσθημα. Η κωδικοποίηση για την έκφραση της έκπληξης θα είναι '0000001', για την έκφραση της ευτυχίας '0000010', για την έκφραση της λύπης '0000100', για την ουδέτερη έκφραση '0001000', για την έκφραση του φόβου '0010000', για την έκφραση της αηδίας '0100000' και για την έκφραση του θυμού '1000000'. Ο πίνακας `RealLabels` κρατάει της κωδικοποιήσεις αυτές με ένα "bit" ανά γραμμή, που σημαίνει ότι είναι ένας πίνακας $N \times 7$, αν N ο τελικός αριθμός εικόνων.

Οι εικόνες διαβάζονται κάνοντας χρήση της `imread`, γίνεται μετατροπή των εικόνων σε διπλής ακρίβειας (`double precision`) με την χρήση της συνάρτησης `im2double()` και με την χρήση του Viola-Jones αλγορίθμου ανιχνεύεται η πιο ακριβής περιοχή του προσώπου. (`FDetect = vision.CascadeObjectDetector;` `BB = step(FDetect, A);` Εκτελείται αποκοπή της περιοχής του προσώπου με την χρήση της

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

συνάρτησης `imcrop()` και του `bounding box` και μετά αλλαγή του μεγέθους της εικόνας σε `300x300` με την χρήση της συνάρτησης `imresize()`. Από `300x300` η εικόνα μετατρέπεται σε `1x90000` με την συνάρτηση `reshape()` και εισάγεται στον πίνακα `Face` στην θέση `ii` με την εντολή `Face(ii,:)=reshape(X,300*300,1);` Όταν τελειώσει η `for` για όλες τις εικόνες, ο πίνακας `Face` και ο πίνακας `RealLabels` αποθηκεύονται με την χρήση της συνάρτησης `uisave()`. Ο κώδικας που υλοποιεί την προ επεξεργασία των εικόνων δίνεται παρακάτω.

```
% Process all image files in those folders.
for k = 1 : numberOfFolders
    % Get this folder and print it out.
    thisFolder = listOfFolderNames{k};
    fprintf('Processing folder %s\n', thisFolder);

    % Get PNG files.
    filePattern = sprintf('%s/*.png', thisFolder);
    baseFileNames = dir(filePattern);
    % Add on TIF files.
    filePattern = sprintf('%s/*.tif', thisFolder);
    baseFileNames = [baseFileNames; dir(filePattern)];
    % Add on JPG files.
    filePattern = sprintf('%s/*.jpg', thisFolder);
    baseFileNames = [baseFileNames; dir(filePattern)];
    numberOfImageFiles = length(baseFileNames);
    % Now we have a list of all files in this folder.

    if numberOfImageFiles >= 1
        % Go through all those image files.
        for f = 1 : numberOfImageFiles
            fullFileName = fullfile(thisFolder, baseFileNames(f).name);

            Labels{ii,1}=thisFolder;
            if strfind(Labels{ii,1}, 'Surprise')
                RealLabels(ii,:)= [0;0;0;0;0;0;0;1];
            end
            if strfind(Labels{ii,1}, 'Happiness')
                RealLabels(ii,:)= [0;0;0;0;0;0;1;0];
            end
            if strfind(Labels{ii,1}, 'Sadness')
                RealLabels(ii,:)= [0;0;0;0;1;0;0;0];
            end
            if strfind(Labels{ii,1}, 'Neutral')
                RealLabels(ii,:)= [0;0;0;1;0;0;0;0];
            end
            if strfind(Labels{ii,1}, 'Fear')
                RealLabels(ii,:)= [0;0;1;0;0;0;0;0];
            end
            if strfind(Labels{ii,1}, 'Disgust')
                RealLabels(ii,:)= [0;1;0;0;0;0;0;0];
            end
            if strfind(Labels{ii,1}, 'Anger')
                RealLabels(ii,:)= [1;0;0;0;0;0;0;0];
            end

            A=imread(fullFileName);
            A=im2double(A);
            FDetect = vision.CascadeObjectDetector;
            BB = step(FDetect,A);
```



```

        if size(BB,1) == 1
            FDetect=imcrop(A,BB);
            X = imresize(FDetect,'OutputSize',[300 300]);
            Face(ii,:)=reshape(X,300*300,1);
            ii=ii+1;
            fprintf('        Processing image file %s\n', fullFileName);
        else
            ii=ii+1;
            fprintf('        Multiple detections on file %s\n',
fullFileName);
            delete(fullFileName);

        end
    end
else
    fprintf('        Folder %s has no image files in it.\n', thisFolder);
end

end
uisave('Face','Face');
uisave('RealLabels','RealLabels');

```

Ο κώδικας για την υλοποίηση του τεχνητού νευρωνικού δικτύου βρίσκεται στο αρχείο **Method1_NeuralNetwork.m**. Γίνεται χρήση του Neural Network Toolbox του matlab, το οποίο παρέχει αλγόριθμους, λειτουργίες και εφαρμογές για την δημιουργία, την εκπαίδευση, την απεικόνιση και προσομοίωση νευρωνικών δικτύων. Αυτό σημαίνει ότι αντί να δημιουργηθεί από το μηδέν ένα τεχνητό νευρωνικό δίκτυο, το matlab δίνει την δυνατότητα με συναρτήσεις του toolbox να δημιουργηθεί ένα αρχικό δίκτυο που με σωστή ρύθμιση των παραμέτρων του να προσαρμοστεί για την επίλυση του προβλήματος της αναγνώρισης εκφράσεων.

Παρακάτω γίνεται η επεξήγηση των συναρτήσεων του Neural Network Toolbox και τα βήματα για την δημιουργία του. Αρχικά με την χρήση της συνάρτησης `importdata()` εισάγονται οι πίνακες `Face` και `RealLabels` στην μνήμη και γίνεται αντιμετάθεση σε αυτούς (οι γραμμές γίνονται στήλες και οι στήλες γραμμές). Το αποτέλεσμα εισάγεται σε δύο νέους πίνακες τους `X` και `R` αντίστοιχα. Ο πίνακας `hiddenLayerSize` κρατάει ένα διάνυσμα που μας δίνει το μέγεθος των κρυφών στρωμάτων του δικτύου, καθώς και το πλήθος αυτών. Αν `hiddenLayerSize = [25 50 100]` σημαίνει ότι το δίκτυο θα έχει τρία κρυφά στρώματα, από τα οποία το πρώτο θα έχει μέγεθος 25 νευρώνων, το δεύτερο 50 νευρώνων και το τρίτο μέγεθος 100 νευρώνων. Το μέγεθος και ο αριθμός των νευρώνων ανά κρυφό στρώμα θα αλλάξει πολλές φορές (trial and error) για να βρεθεί το δίκτυο που παρουσιάζει το καλύτερο ποσοστό αναγνώρισης για τις επτά εκφράσεις του προσώπου.

Η εντολή `net=patternnet(hiddenLayerSize);` δημιουργεί ένα feedforward δίκτυο αναγνώρισης προτύπων που μπορεί να εκπαιδευτεί για να ταξινομήσει εισόδους σύμφωνα με τις κατηγορίες εξόδου. Τα δεδομένα εξόδου για το δίκτυο αναγνώρισης προτύπων θα πρέπει να αποτελούνται από μηδενικές τιμές, εκτός από έναν άσσο το στοιχείο `i`, όπου `i` είναι η κατηγορία που εκπροσωπεί [53]. Για αυτό τον λόγο έχει γίνει η «κωδικοποίηση» των συναισθημάτων στον πίνακα `RealLabels` σε μία ακολουθία από 6 μηδενικά και έναν άσσο. Οι εντολές `net.input.processFcns = { 'removeconstantrows', 'mapminmax' };` και `net.output.processFcns = { 'removeconstantrows', 'mapminmax' };` αφαιρούν τις γραμμές εισόδου

και εξόδου που έχουν σταθερές τιμές και επεξεργάζονται τα δεδομένα χαρτογραφώντας την μικρότερη τιμή της σειράς με -1 και την μεγαλύτερη με 1.

Επειδή έχουμε συγκεκριμένο αριθμό εικόνων θέλουμε να τις χωρίσουμε σε τρεις κατηγορίες. Μία κατηγορία για να πραγματοποιηθεί η εκπαίδευση του δικτύου, μία κατηγορία για την επικύρωση του δικτύου και μία για την δοκιμή της προβλεπτικής ικανότητας του δικτύου. Για να χωριστούν με συγκεκριμένη σειρά τα δεδομένα γίνεται η χρήση της `net.divideFcn = 'divideind'`; όπου `net.divideFcn` ορίζει τον τρόπο που χωρίζονται τα δεδομένα εισόδου. Με την παράμετρο `divideind` χωρίζονται τα δεδομένα σύμφωνα με προεπιλεγμένους δείκτες. Οι εντολές `net.divideParam.trainInd = 1:1532`; `net.divideParam.valInd = 1533:1922`; και `net.divideParam.testInd = 1923:2418`; προσδιορίζουν ποια δεδομένα θα χρησιμοποιηθούν για την εκπαίδευση, την επικύρωση και την δοκιμή του δικτύου. Η εντολή `net.divideMode = 'sample'`; ορίζει με το όρισμα `sample` ότι θα χωριστούν τα δεδομένα ανά δείγμα.

Η συνάρτηση εκπαίδευσης του δικτύου ορίζεται με την εντολή `net.trainFcn = 'trainscg'`; και η συνάρτηση εκπαίδευσης που κάνουμε χρήση είναι η `'trainscg'` που είναι Scaled conjugate gradient backpropagation. Η συνάρτηση της επίδοσης του δικτύου ορίζεται με την εντολή `net.performFcn='crossentropy'`; και η συνάρτηση επίδοσης είναι η `'crossentropy'`. Η cross-entropy υπολογίζει την απόδοση του δικτύου, αφού της έχουν δοθεί οι εξόδοι και οι κατηγορίες στόχοι, με τα βέλτιστα βάρη επίδοσης και άλλες παραμέτρους. Η συνάρτηση αυτή επιστρέφει ένα αποτέλεσμα που τιμωρεί τον βαθμό επίδοσης αν τα δεδομένα εξόδου είναι ανακριβή (y κοντά στο $1-t$), αλλά έχει πολύ μικρές απώλειες αν τα δεδομένα έχουν σχετικά σωστές κατηγοριοποιήσεις (y κοντά στο t) [54]. Εκτός από την cross-entropy δοκιμάστηκε εναλλακτικά και η συνάρτηση επίδοσης μέσω τετραγωνικού κανονικοποιημένου σφάλματος (Mean squared normalized error) `'mse'` [55].

Η συνάρτηση που ορίζει ποια διαγράμματα θα υλοποιήσει το εργαλείο του matlab είναι η `net.plotFcns`. Το `'plotconfusion'` δίνει το διάγραμμα ταξινόμησης των συγχυσμένων δεδομένων. Από αυτό το διάγραμμα μπορούμε να δούμε εύκολα τι ποσοστό επιτυχούς εκπαίδευσης, επικύρωσης και δοκιμής έχει παραχθεί από το νευρωνικό δίκτυο, καθώς και τον αριθμό των εικόνων που έχουν κατηγοριοποιηθεί λανθασμένα.

Πριν αρχίσει η εκπαίδευση του νευρωνικού δικτύου ορίζονται μία σειρά από παραμέτρους που είναι οι προκαθορισμένοι (default) για την συνάρτηση εκπαίδευσης (`'trainscg'`) εκτός από την `max_fail` που είναι ο μέγιστος αριθμός εποχών αποτυχίας επικύρωσης, που αντί για 6 εποχές έγινε 10. Με την συνάρτηση `train(net,X,R)`; αρχίζει η εκπαίδευση του νευρωνικού δικτύου. Στην δοκιμή του δικτύου η `e = gsubtract(R,y)`; υπολογίζει το λάθος με την συνάρτηση της γενικευμένης αφαίρεσης, αφαιρώντας τις εξόδους του δικτύου από τις επιθυμητές εξόδους (πραγματικές τιμές). Η εντολή `performance = perform(net,R,y)` υπολογίζει την απόδοση του δικτύου σαν σύνολο, ενώ οι εντολές `trainPerformance = perform(net, trainTargets, y)`, `valPerformance = perform(net, valTargets, y)`, `testPerformance = perform(net, testTargets, y)` υπολογίζουν την απόδοση των δεδομένων εκπαίδευσης, επικύρωσης και δοκιμής αντίστοιχα. Ο κώδικας υλοποίησης του νευρωνικού δικτύου δίνεται παρακάτω.

```
% Solve a Pattern Recognition Problem with a Neural Network
```

Μπολαράκης Παναγιώτης

```
Face=importdata('Face.mat');
X = Face';
clear Face;
RealLabels=importdata('RealLabels.mat');
R = RealLabels';
clear RealLabels;

% Create a Pattern Recognition Network
hiddenLayerSize = [25 50];
net = patternnet(hiddenLayerSize);

% Choose Input and Output Pre/Post-Processing Functions
% For a list of all processing functions type: help nnprocess
net.input.processFcns = {'removeconstantrows','mapminmax'};
net.output.processFcns = {'removeconstantrows','mapminmax'};

% Setup Division of Data for Training, Validation, Testing
% For a list of all data division functions type: help nndivide
% net.divideFcn = 'divideint'; % Divide data randomly
% % Divide targets into three sets using specified indices
net.divideFcn = 'divideind'; % Divide data randomly
net.divideMode = 'sample'; % Divide up every sample
% Training70%
net.divideParam.trainInd = 1:1532;
net.divideParam.valInd = 1533:1922;
net.divideParam.testInd = 1923:2418;

% For help on training function 'trainscg' type: help trainscg
% For a list of all training functions type: help nntrain
net.trainFcn = 'trainscg'; % Scaled conjugate gradient

% Choose a Performance Function
% For a list of all performance functions type: help nnperformance
% net.performFcn = 'crossentropy'; % Cross-entropy
net.performFcn = 'mse';

% Choose Plot Functions
% For a list of all plot functions type: help nnplot
net.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...
    'plotregression','plotfit','plotconfusion'};

% Choose Training Parameters
net.trainParam.goal = 0;
net.trainParam.max_fail = 10;
net.trainParam.epochs = 1000;
net.trainParam.showCommandLine = true;
net.trainParam.showWindow = true;
net.trainParam.lr = 0.1;
net.trainParam.min_grad = 1e-6;
net.trainParam.show = 25;

% Train the Network
[net,tr] = train(net,X,R);
% Test the Network
y = net(X);
e = gsubtract(R,y);
tind = vec2ind(R);
yind = vec2ind(y);
percentErrors = sum(tind ~= yind)/numel(tind);
performance = perform(net,R,y)
```

```
% Recalculate Training, Validation and Test Performance
trainTargets = R .* tr.trainMask{1};
valTargets = R .* tr.valMask{1};
testTargets = R .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,y)
valPerformance = perform(net,valTargets,y)
testPerformance = perform(net,testTargets,y)
```

5.3.2 Κώδικας δεύτερης υλοποίησης

Η δεύτερη υλοποίηση κάνει χρήση του φίλτρου Gabor για την δημιουργία μίας τράπεζας Gabor (Gabor bank). Το μέγεθος της εξαρτάται από το πόσα μήκη κύματος και προσανατολισμοί θα επιλεχθούν. Η συνάρτηση `gabor` δημιουργεί την τράπεζα φίλτρων `gabor` και η συνάρτηση `imgaborfilt` εφαρμόζει το φίλτρο πάνω στην εικόνα `X`. Ο `gaborMag` είναι ένας τρισδιάστατος πίνακας που περιέχει τις τιμές των εννιά φίλτρων της εικόνας `X`. Τα φίλτρα ανασχηματίζονται για να γίνουν στήλες και μαζί με την αρχική εικόνα, που επίσης γίνεται στήλη, εισάγονται στον πίνακα `Face` ο οποίος αποθηκεύεται μαζί με τις ετικέτες `RealLabels` των εικόνων.

Το πρόβλημα που αντιμετωπίζεται με τη μέθοδο αυτή είναι ότι τα φίλτρα καταλαμβάνουν αρκετό χώρο στην μνήμη και ο αριθμός τους πρέπει να είναι τέτοιος ώστε ο τελικός πίνακας να μην ξεπερνάει το όριο του 2GB που ορίζει το `matlab`. Αυτό περιορίζει το μέγεθος που θα έχουν τα φίλτρα και από 300x300 που ήταν αρχικά έγιναν σμίκρυνση (`resize`) σε 100x100. Σημειώνεται ότι πρώτα υπολογίζονται τα φίλτρα στην 300x300 εικόνα και μετά γίνονται σμίκρυνση (`resize`), γιατί αν γίνει πρώτα η εικόνα σμίκρυνση (`resize`) και μετά υπολογιστούν τα φίλτρα χάνεται ακόμα περισσότερη πληροφορία. Ο κώδικας για την επεξεργασία των εικόνων δίνεται από το αρχείο **Method2_NeuralNetworkSetGaborV1.m** για εννιά φίλτρα Gabor διαστάσεων 100x100 και για 40 φίλτρα Gabor διαστάσεων 50x50 από το αρχείο **Method2_NeuralNetWorkSetGaborV2.m**.

```
if numberOfImageFiles >= 1
    % Go through all those image files
    for f = 1 : numberOfImageFiles
        fullFileName = fullfile(thisFolder, baseFileNames(f).name);
        A=imread(fullFileName);
        A=im2double(A);
        FDetect = vision.CascadeObjectDetector;
        BB = step(FDetect,A);

        Labels{ii,1}=thisFolder;
        if strfind(Labels{ii,1}, 'Surprise')
            RealLabels(ii,:)=[0;0;0;0;0;0;0;1];
        end
        if strfind(Labels{ii,1}, 'Happiness')
            RealLabels(ii,:)=[0;0;0;0;0;0;1;0];
        end
        if strfind(Labels{ii,1}, 'Sadness')
            RealLabels(ii,:)=[0;0;0;0;0;1;0;0];
        end
        if strfind(Labels{ii,1}, 'Neutral')
            RealLabels(ii,:)=[0;0;0;0;1;0;0;0];
        end
        if strfind(Labels{ii,1}, 'Fear')
            RealLabels(ii,:)=[0;0;1;0;0;0;0;0];
        end
    end
end
```

```

if strfind(Labels{ii,1}, 'Disgust')
    RealLabels(ii,:)=[0;1;0;0;0;0;0;0];
end
if strfind(Labels{ii,1}, 'Anger')
    RealLabels(ii,:)=[1;0;0;0;0;0;0;0];
end

if size(BB,1) == 1
    FDetect=imcrop(A,BB);
    X = imresize(FDetect, 'OutputSize', [300 300]);

    gaborArray = gabor([4 6 8],[0 45 90]);
    gaborMag = imgaborfilt(X,gaborArray);

    Face=reshape(imresize(gaborMag(:,:,1), 'OutputSize', [100 100]), [],1);
    Y1=size(gaborMag);
    for j=2:Y1(3)
        A1=reshape(imresize(gaborMag(:,:,1), 'OutputSize', [100 100]), [],1);
        Face=[Face;A1];
    end
    F(:,ii)=Face;
    ii=ii+1;
    fprintf('        Processing image file %s\n', fullFileName);
else
    ii=ii+1;
    fprintf('        Multiple detections on file %s\n', fullFileName);
    delete(fullFileName);
end
end
else
    fprintf('        Folder %s has no image files in it.\n', thisFolder);
end
end

```

Ο κώδικας υλοποίησης του νευρωνικού δικτύου είναι ο ίδιος με αυτόν της πρώτης υλοποίησης. Για λόγους οργάνωσης των φακέλων της πτυχιακής εργασίας, ώστε να μην μπερδεύονται τα αρχεία μεταξύ τους, έχει δημιουργηθεί άλλο ένα αρχείο με όνομα **Method2_NeuralNetwork.m** που είναι το ίδιο με το **Method1_NeuralNetwork.m**.

5.3.3 Κώδικας τρίτης υλοποίησης

Για την τρίτη υλοποίηση που έγινε αφαίρεση της ουδέτερης έκφρασης σαν φίλτρο, έπρεπε να γίνει κατηγοριοποίηση των ατόμων σε ξεχωριστούς φακέλους ανά άτομο και να φορτωθούν οι εικόνες ουδέτερης έκφρασης σε ξεχωριστό κελί από τις υπόλοιπες. Για κάθε άτομο οι εικόνες με όλες τις εκφράσεις του έχουν οργανωθεί σε ένα φάκελο και σε ένα ξεχωριστό φάκελο με όνομα «Neutral» μία εικόνα με την ουδέτερη έκφραση του. Σε έναν βρόγχο συγκρίνεται με τη συνάρτηση `strfind` το κωδικό όνομα του φακέλου κάθε ατόμου με την διαδρομή στην οποία βρίσκεται εκείνη την στιγμή η επανάληψη. Στην περίπτωση που η `if` επιστρέψει αποτέλεσμα 1 ο κώδικας μπαίνει σε μία δεύτερο βρόγχο που συγκρίνει πάλι την τρέχουσα διαδρομή με την συμβολοσειρά `Neutral` και αν επιστρέψει ξανά αποτέλεσμα 1 βάζει την εικόνα στον κελί `NeutralCell`, διαφορετικά βάζει την εικόνα στην κατάλληλη θέση στο κελί `ImgCell1`. Αυτό γίνεται για όλες τις εικόνες μέχρι να αποθηκευτούν στο κατάλληλο κελί. Ο κώδικας που εκτελεί αυτή τη λειτουργία λέγεται **Method3_Sort.m** και δεν θα παρατεθεί παρακάτω γιατί είναι 850 γραμμές.

Η τελευταία γραμμή κώδικα του **Method3_Sort.m** καλεί τον κώδικα **Method3_filter.m** όπου γίνεται και η αφαίρεση την ουδέτερης έκφρασης από τις υπόλοιπες. Η συνάρτηση `imwrite` αποθηκεύει την φιλτραρισμένη εικόνα `D1` με το όνομα που της έχει δοθεί από την συμβολοσειρά `filename1`. Η συμβολοσειρά αυτή παίρνει την τιμή της με την χρήση της συνάρτησης `sprintf('file%d-%d.png', i, j)`; Ο κώδικας δίνεται παρακάτω.

```
h=fspecial('unsharp');
for i=1:95
    for j=1:35
        if isempty(ImgCell1{j,i})
            break;
        elseif isempty(NeutralCell{i})
            break;

        else
            L=ImgCell1{j,i};
            K=NeutralCell{i};
            D=L-K;
            D1=imfilter(D,h);
            ImgCell{j,i}=D1;
            filename1=sprintf('file%d-%d.png', i, j);
            imwrite(D1,filename1);
        end
    end
end
```

Αφού όλες οι φιλτραρισμένες εικόνες οργανωθούν ξανά ανά έκφραση και προσδιοριστεί ένα ποσοστό για επικύρωση και δοκιμή του δικτύου, οι εικόνες γίνονται στήλες (`reshape`) και αποθηκεύονται σε ένα πίνακα τον `Face`, ενώ οι έξι (πλέον) ετικέτες εκφράσεων αποθηκεύονται στον πίνακα `RealLabels`. Ο αριθμός των εικόνων, επειδή δεν υπάρχει η ουδέτερη έκφραση πλέον, μειώθηκε στις 1828, όπου 997 χρησιμοποιούνται για την εκπαίδευση, 470 για την επικύρωση και 359 για την δοκιμή της ακρίβειας του νευρωνικού δικτύου. Το στρώμα εξόδου του δικτύου θα έχει έξι εξόδους άρα και οι ετικέτες θα αποτελούνται από έξι “bit”. Για την έκφραση της έκπληξης το νευρωνικό δίκτυο πρέπει να έχει έξοδο ‘000001’, για την έκφραση της ευτυχίας ‘000010’, για την έκφραση της λύπης ‘000100’, για την έκφραση του φόβου ‘001000’, για την έκφραση της αηδίας ‘010000’ και για την έκφραση του θυμού ‘100000’. Το αρχείο **Method3_NeuralNetworkSet.m** περιέχει τον κώδικα που ορίζει τις νέες ετικέτες και εισάγει τις εικόνες στον πίνακα `Face`. Το αρχείο `Method3_NeuralNetwork.m` περιέχει τον κώδικα για την δημιουργία του δικτύου και το μόνο που αλλάζει είναι οι παράμετροι `trainInd`, `valInd` και `testInd` που προσαρμόζονται στον νέο αριθμό εικόνων.

5.4 Αποτελέσματα συστημάτων – Συζήτηση

5.4.1 Αποτελέσματα πρώτης υλοποίησης

Τα αποτελέσματα της πρώτης υλοποίησης είναι ικανοποιητικά εφόσον οι εικόνες που εισάγονται στο νευρωνικό δίκτυο δεν έχουν υποστεί κάποια επιπλέον επεξεργασία έκτος από την μετατροπή τους σε στήλες για να μπορούν να εισαχθούν στο δίκτυο. Το καλύτερο αποτέλεσμα ήταν 74.8% ποσοστό επιτυχίας αναγνώρισης των δοκιμαστικών εικόνων από συνολικά 496 εικόνες. Δηλαδή 371 εικόνες από τις 496 αναγνωρίστηκαν επιτυχώς από το δίκτυο. Παρακάτω δίνονται κάποιο πίνακες (Πίνακες 16-18) με αποτελέσματα δοκιμών που

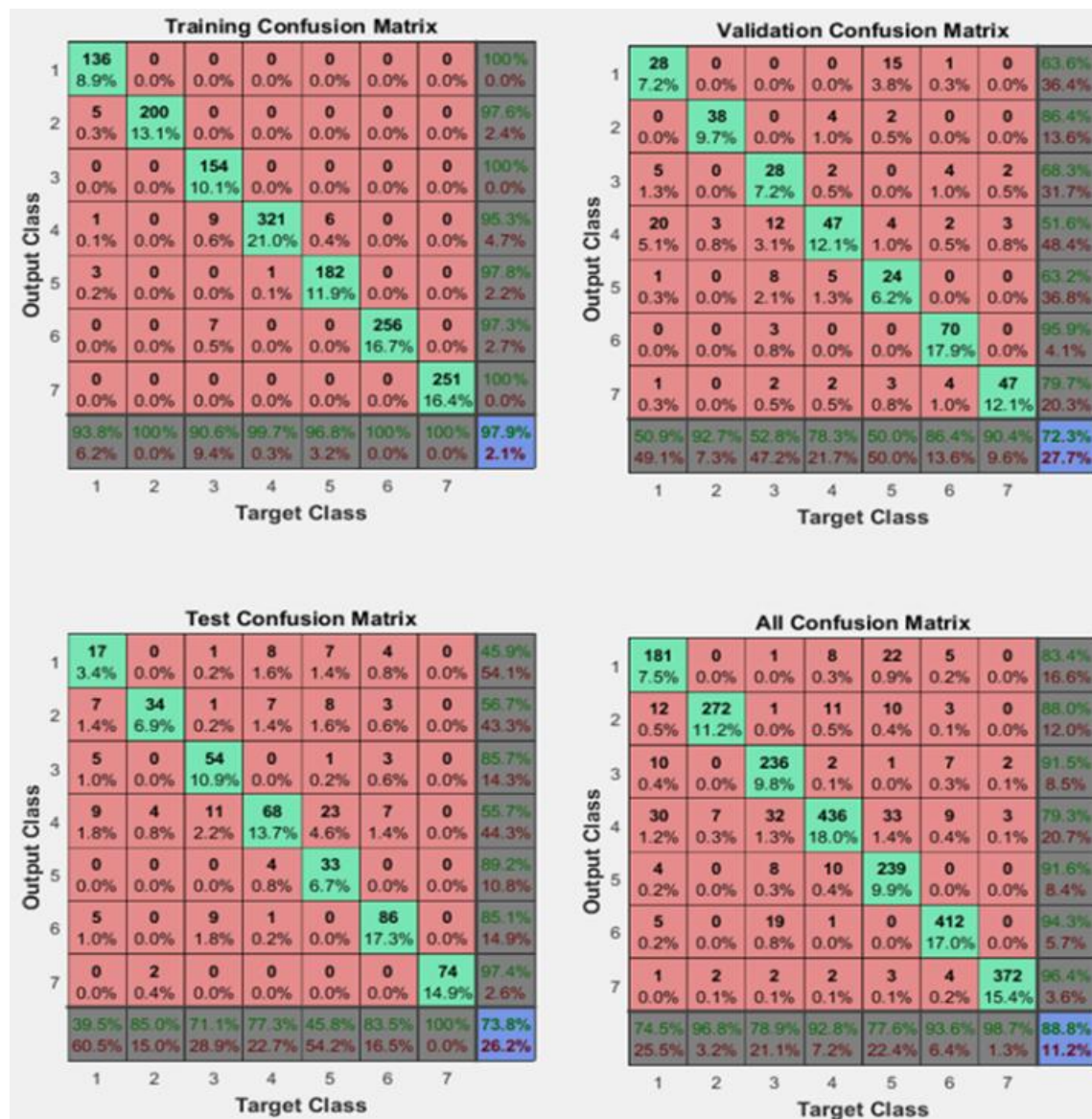
έγιναν για αυτή την μέθοδο, αλλάζοντας τον αριθμό των νευρώνων των κρυφών στρώματων και με την χρήση διαφορετικών συναρτήσεων για την εκπαίδευση του δικτύου. Στους πίνακες σύγκρισης (Εικόνες 9-14) αποτυπώνεται το ποσοστό των εικόνων που γίνεται λανθασμένη κατηγοριοποίηση. Ακόμα και αν η εκπαίδευση είναι επιτυχής με ένα ποσοστό κοντά ή ίσο με 100%, τα ποσοστά για τις δοκιμαστικές εικόνες και τις εικόνες επικύρωσης συνήθως είναι κοντά το ένα με το άλλο, με τις εικόνες επικύρωσης να έχουν ένα υψηλότερο ποσοστό αναγνώρισης.

Αυτό που παρατηρείται στους πίνακες σύγκρισης (Εικόνες 9-14) από τα αποτελέσματα των δοκιμαστικών εικόνων (και κατ' επέκταση των εικόνων επικύρωσης) είναι ότι εικόνες από όλες τις κατηγορίες εκφράσεων συγχύζονται με την τέταρτη κατηγορία που εκφράζει την ουδέτερη έκφραση. Η κατηγορία με το χαμηλότερο ποσοστό αναγνώρισης είναι η πρώτη που εκφράζει το συναίσθημα του θυμού, που συγχέεται με τις κατηγορίες δύο και τέσσερα και σε μικρότερο ποσοστό με την τρία, δηλαδή με την έκφραση της αηδίας, την ουδέτερη έκφραση και την έκφραση του φόβου αντίστοιχα. Οι κατηγορίες δύο, έξι και επτά που εκφράζουν τα συναισθήματα της αηδίας, της ευτυχίας και της έκπληξης δίνουν πάντα το πιο υψηλό ποσοστό αναγνώρισης με την έκπληξη να είναι συνήθως πολύ κοντά ή ίση στο 100%.

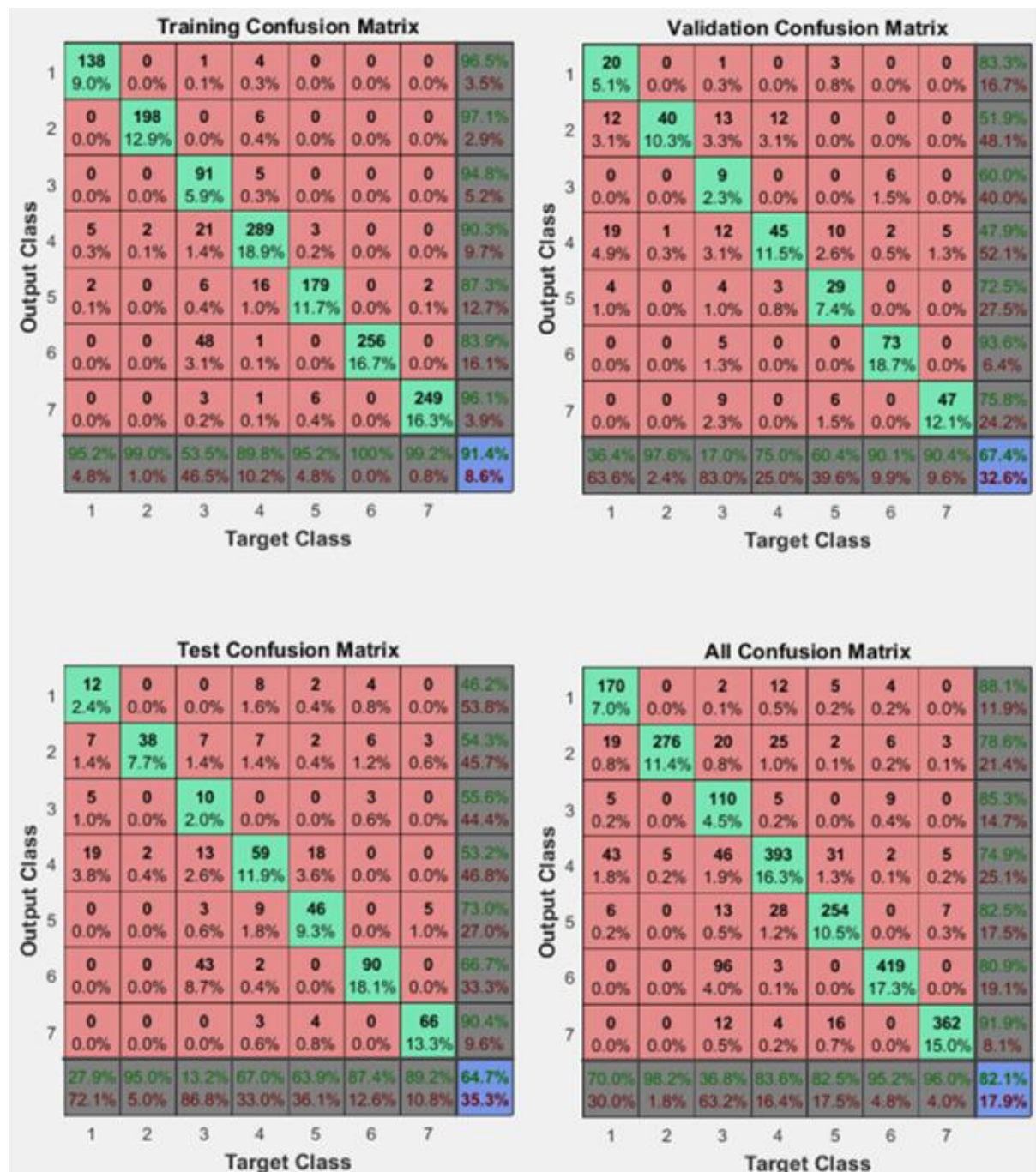
Κρυφά Στρώματα	[20 20]	[20 20]	[20 20]	[20 20]	[50 50]	[50 50]	[50 50]	[50 50]
1:Θυμός	93.8	50.9	39.5	74.5	95.2	36.4	27.9	70.0
2:Αηδία	100	92.7	85.0	96.8	99.0	97.6	95.0	98.2
3:Φόβος	90.6	52.8	71.1	78.9	53.5	17.0	13.2	36.8
4:Ουδέτερη	99.7	78.3	77.3	92.8	89.8	75.0	67.0	83.6
5:Αύπη	96.8	50.0	45.8	77.6	95.2	60.4	63.9	82.5
6:Ευτυχία	100	86.4	83.5	93.6	100	90.1	87.4	95.2
7:Έκπληξη	100	90.4	100	98.7	99.2	90.4	89.2	96.0
Σύνολο	97.9	72.3	73.8	88.8	91.4	67.4	64.7	82.1
Εποχές	96	96	96	96	55	55	55	55
	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο

Πίνακας 16: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [20 20] στην πρώτη και [50 50] στην δεύτερη.

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου



Εικόνα 11: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [20 20].

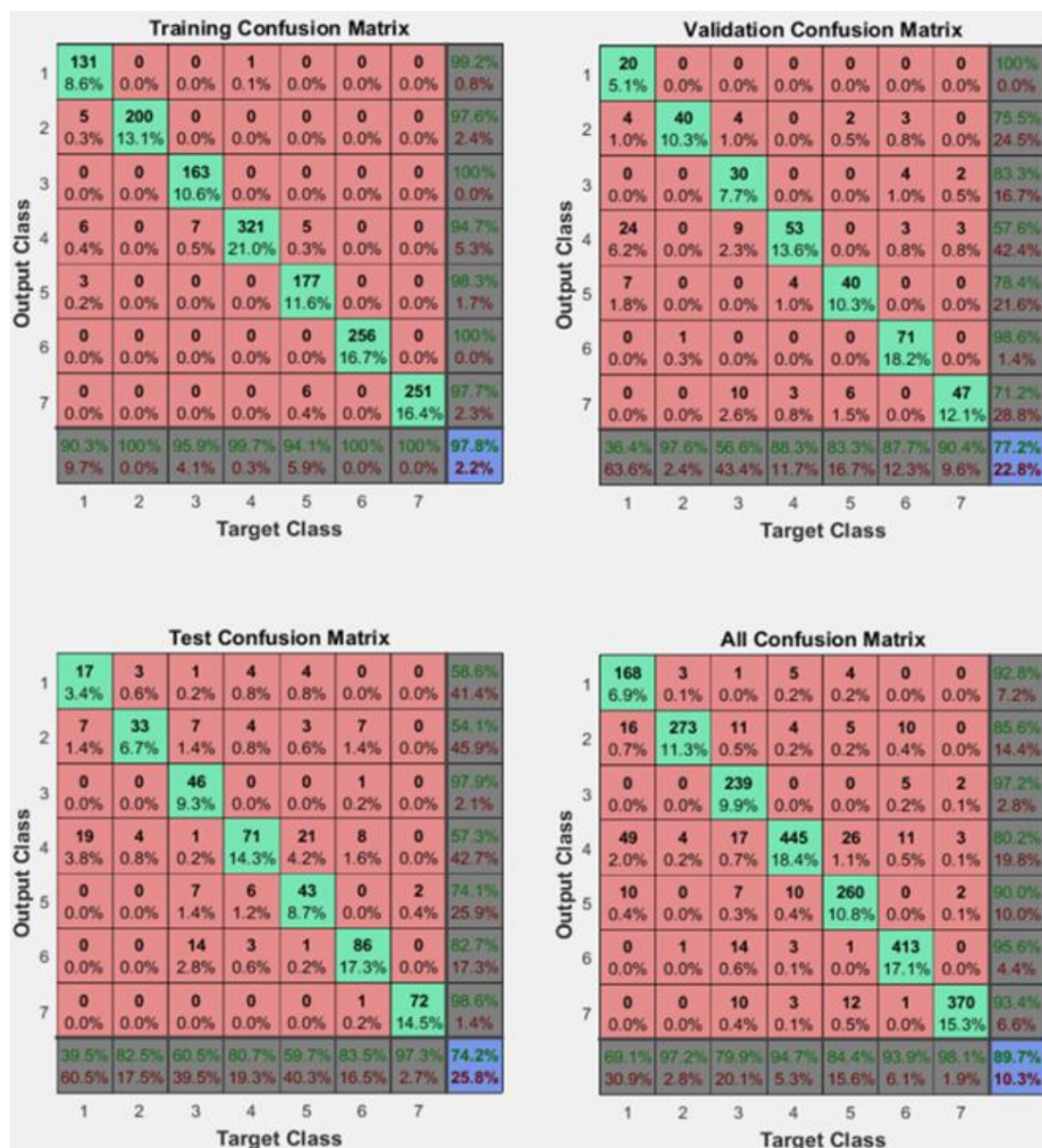


Εικόνα 12: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [50 50].

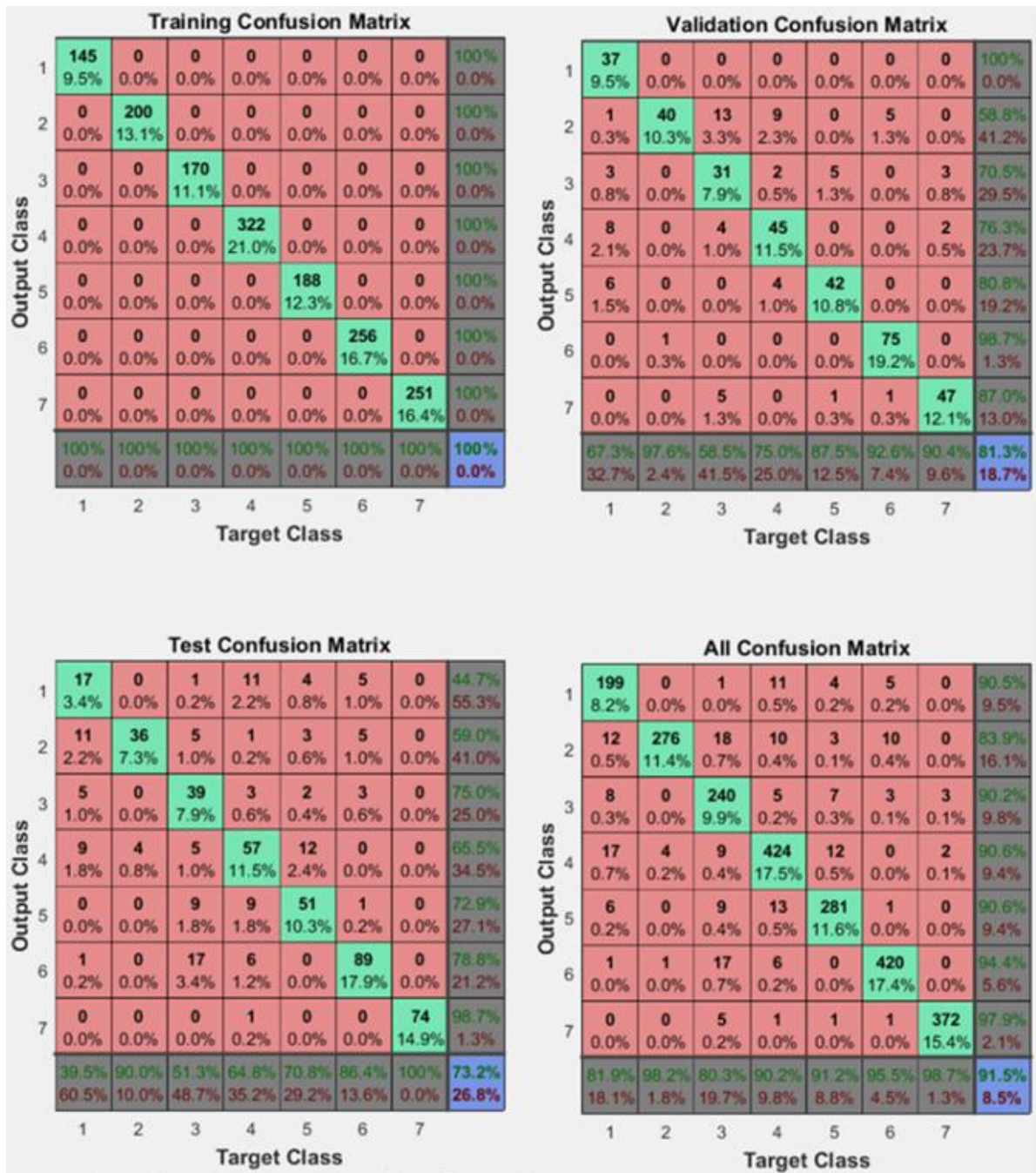
Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

Κρυφά Στρώματα	[60 40]	[60 40]	[60 40]	[60 40]	[20 10]	[20 10]	[20 10]	[20 10]
1:Θυμός	90.3	36.4	39.5	69.1	100	67.3	39.5	81.9
2:Αηδία	100	97.6	82.5	97.2	100	97.6	90.0	98.2
3:Φόβος	95.9	56.5	60.5	79.9	100	58.5	51.3	80.3
4:Ουδέτερη	99.7	88.3	80.7	94.7	100	75.0	64.8	90.2
5:Λύπη	94.1	83.3	59.7	84.4	100	87.5	70.8	91.2
6:Ευτυχία	100	87.7	83.5	93.9	100	92.6	86.4	95.5
7:Έκπληξη	100	90.4	97.3	98.1	100	90.4	100	98.7
Σύνολο	97.8	77.2	74.2	89.7	100	81.3	73.2	91.5
Εποχές	100	100	100	100	153	153	153	153
	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο

Πίνακας 17: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [60 40] στην πρώτη και [20 10] στην δεύτερη.



Εικόνα 13: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [60 40].



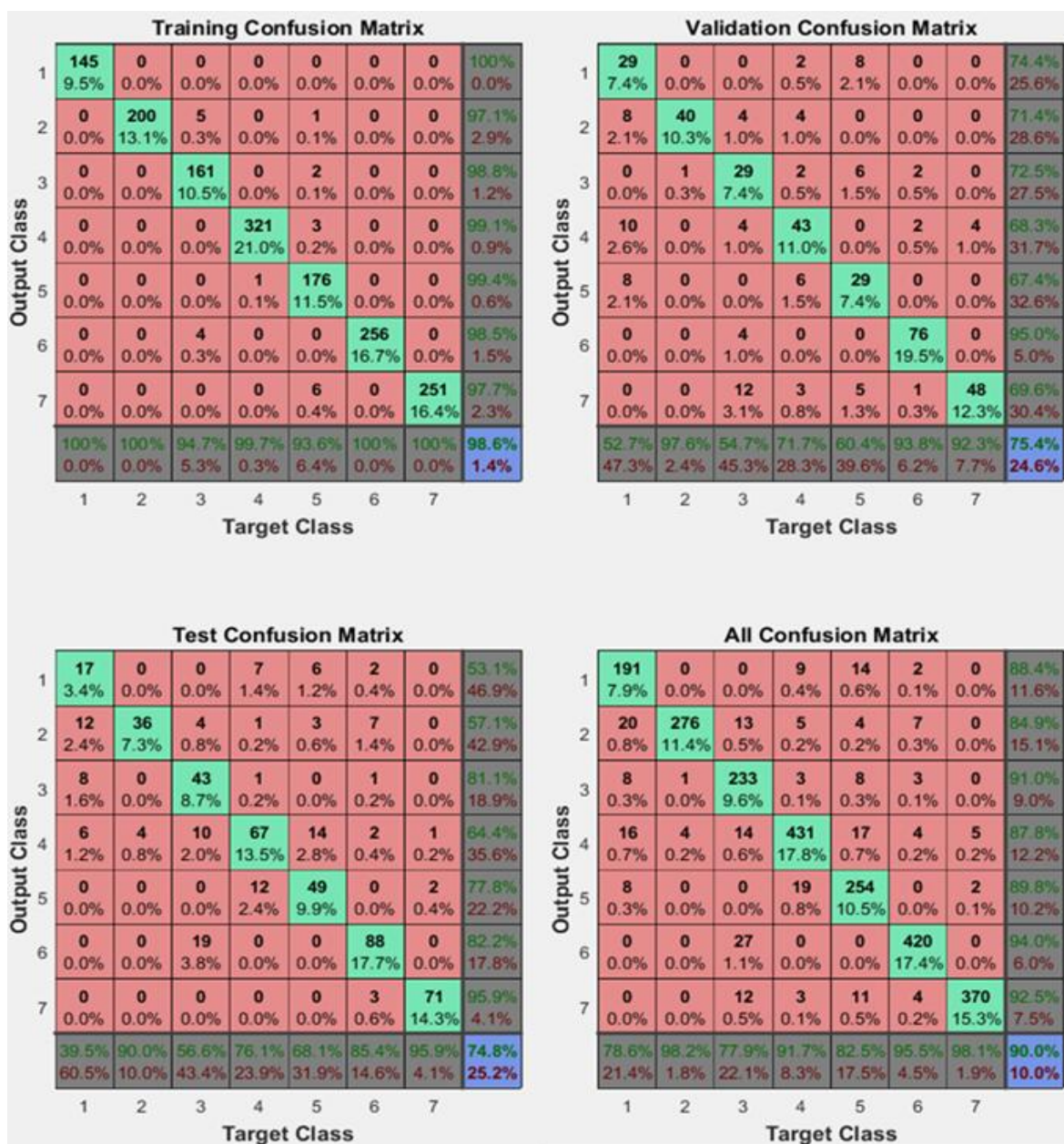
Εικόνα 14: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [20 10].

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

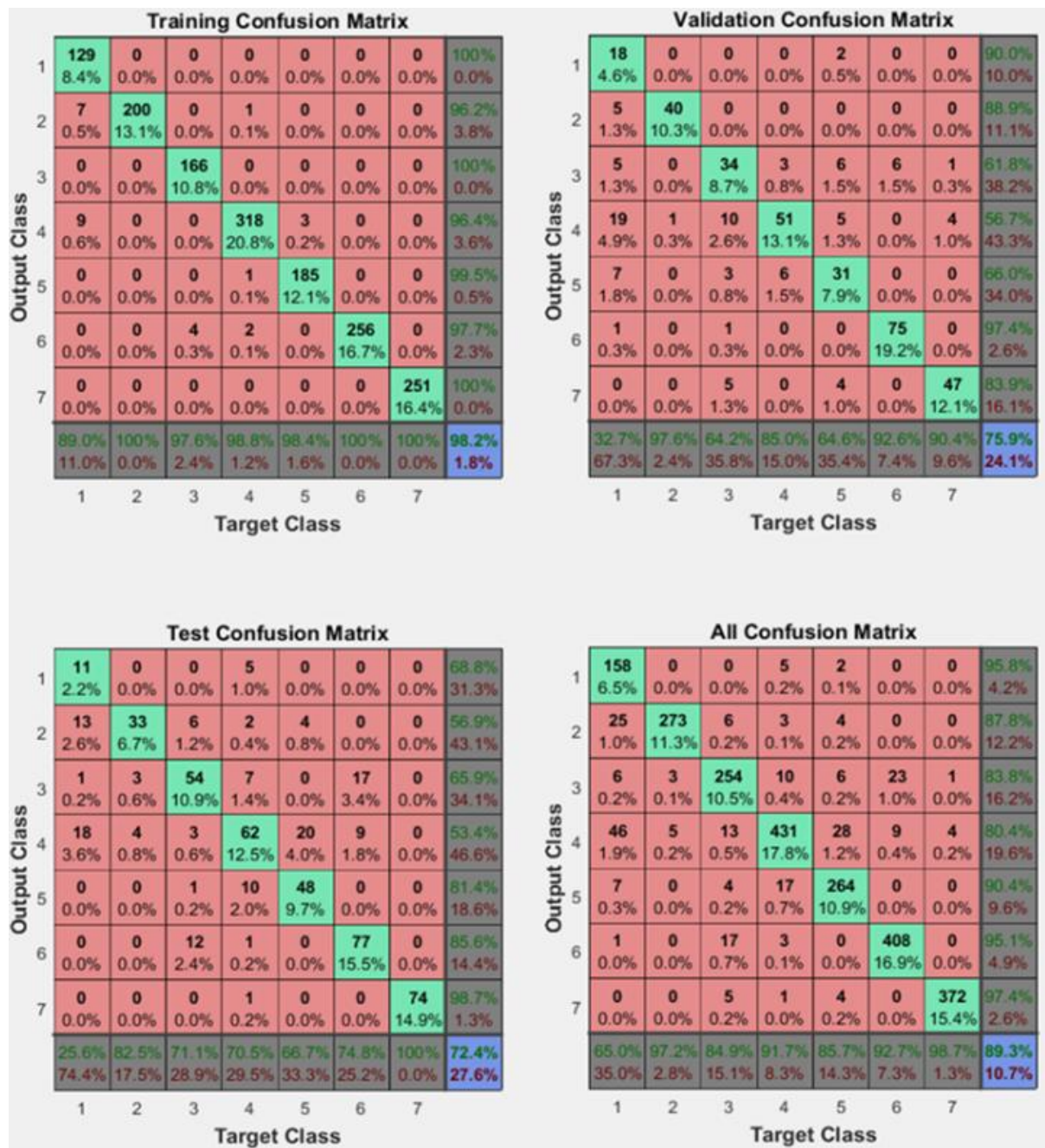
Κρυφά Στρώματα	[25 50]	[25 50]	[25 50]	[25 50]	[30 45]	[30 45]	[30 45]	[30 45]
1:Θυμός	100	52.7	39.5	78.6	89.0	32.7	25.6	65.0
2:Αηδία	100	97.9	90.0	98.2	100	97.6	82.5	97.2
3:Φόβος	94.7	54.7	56.6	77.9	97.6	64.2	71.1	84.9
4:Ουδέτερη	99.7	71.7	76.1	91.7	98.8	85.0	70.5	91.7
5:Λύπη	93.6	60.4	68.1	82.5	98.4	64.6	66.7	85.7
6:Ευτυχία	100	93.8	85.4	95.5	100	92.6	74.8	92.7
7:Έκπληξη	100	92.3	95.9	98.1	100	90.4	100	98.7
Σύνολο	98.6	75.4	74.8	90.0	98.2	75.9	72.4	89.3
Εποχές	99	99	99	99	78	78	78	78

Εκπαίδευση Επικύρωση Δοκιμή Σύνολο Εκπαίδευση Επικύρωση Δοκιμή Σύνολο

Πίνακας 18: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [25 50] στην πρώτη και [30 45] στην δεύτερη.



Εικόνα 15: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50].



Εικόνα 16: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [35 45].

5.4.2 Αποτελέσματα δεύτερης υλοποίησης

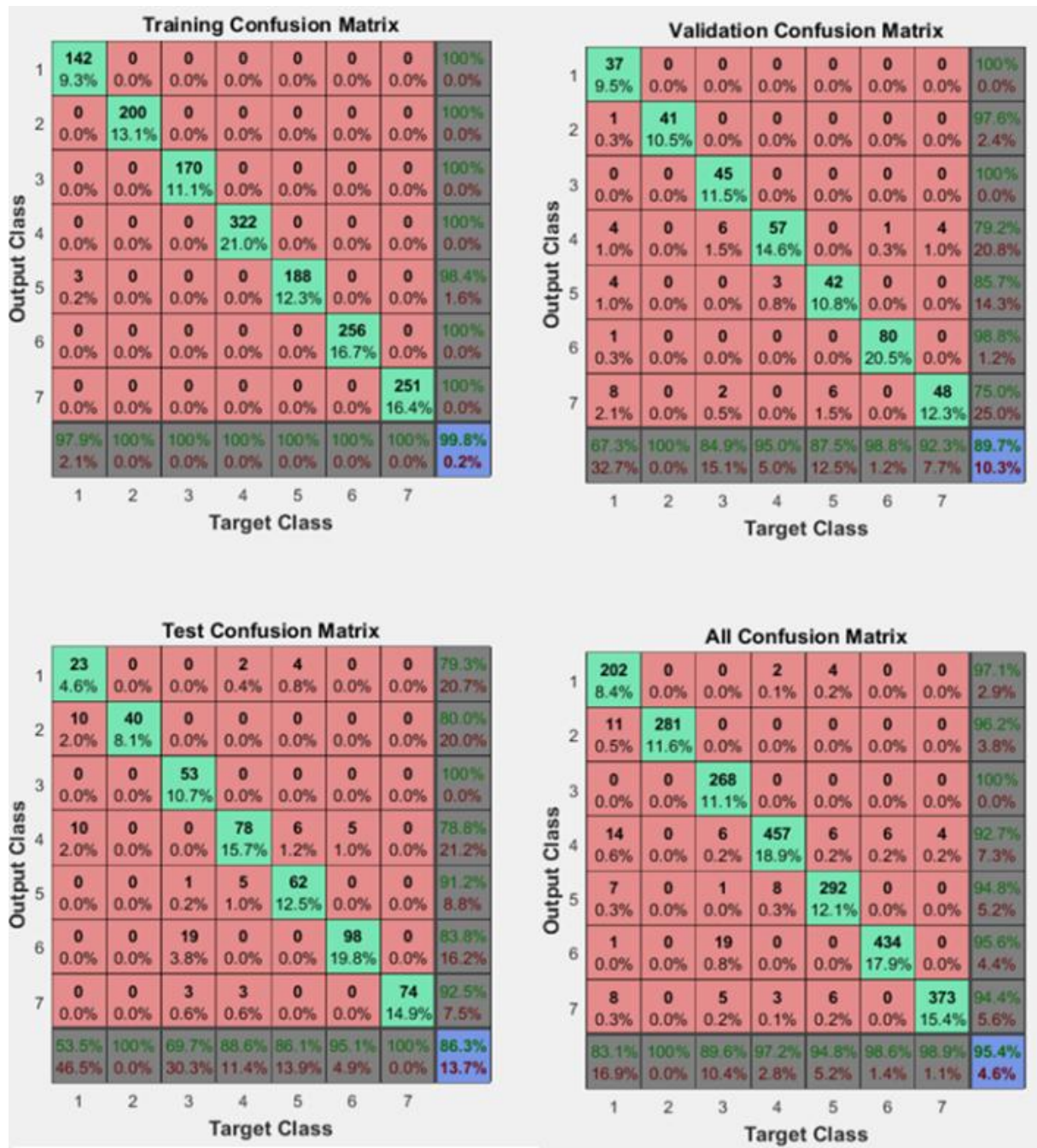
Για την δεύτερη υλοποίηση που εφαρμόστηκαν φίλτρα Gabor σε όλες τις εικόνες, βελτιώθηκαν σημαντικά τα αποτελέσματα στην επικύρωση και στην αναγνώριση των δοκιμαστικών εικόνων σε σχέση με την πρώτη μέθοδο. Εκτός από τους διαφορετικούς τρόπους εκπαίδευσης του δικτύου, με διαφορετικές συναρτήσεις εκπαίδευσης και διαφορετικό αριθμό νευρώνων ανά κρυφό στρώμα και την χρήση διαφορετικών συναρτήσεων επίδοσης, δοκιμάστηκαν πολλοί διαφορετικοί συνδυασμοί φίλτρων Gabor, όσο αναφορά τον αριθμό και το μέγεθος των φίλτρων που χρησιμοποιήθηκαν και τους διαφορετικούς συνδυασμούς προσανατολισμών και μήκους κυμάτων.

Η βελτίωση της αναγνώρισης των δοκιμαστικών εικόνων ξεπέρασε το 11% και η συνολική επίδοση του δικτύου έφτασε το 95.4%. Ο λόγος που οι φιλτραρισμένες εικόνες μειώθηκαν σε μέγεθος από 250x250 σε 100x100 ήταν για να μειωθεί το μέγεθος του τελικού πίνακα και να μπορεί να είναι λειτουργικό το δίκτυο. Στους πίνακες (Πίνακες 19-20) παρακάτω δίνονται τα αποτελέσματα κάποιων δοκιμών που έδωσαν την καλύτερη αναγνώριση στις δοκιμαστικές εικόνες. Όπου υπάρχει το +1 στον αριθμό των φίλτρων σημαίνει ότι εκτός από τα φίλτρα σαν είσοδος στο νευρωνικό δίκτυο δόθηκε και η αρχική εικόνα στο ίδιο μέγεθος με το μέγεθος των υπολοίπων φίλτρων. Η αρχική εικόνα δίνεται σαν είσοδος μαζί με τις φιλτραρισμένες εικόνες γιατί παρατηρήθηκε αύξηση στην επιτυχή αναγνώριση, σε σχέση με τις δοκιμές που χρησιμοποιήθηκαν μόνο οι φιλτραρισμένες εικόνες.

Το μεγαλύτερο ποσοστό επιτυχούς αναγνώρισης των δοκιμαστικών εικόνων 86.3% εμφανίζεται σε ένα δίκτυο με δύο κρυφά στρώματα, με αριθμό νευρώνων 25 και 50, με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient και συνάρτηση απόδοσης Cross-Entropy, με δεδομένα εισόδου τις φιλτραρισμένες εικόνες που δίνονται από έξι φίλτρα Gabor, προσανατολισμού [0 45 90] και μήκους κύματος [4 8] και την αρχική εικόνα μεγέθους 100x100. Ο προσανατολισμός έχει μονάδα μέτρησης σε μοίρες (παίρνει τιμές από 0 έως 360) και το μήκος κύματος έχει μονάδα μέτρησης σε pixels ανά κύκλο (παίρνει τιμές από 2 έως άπειρο).

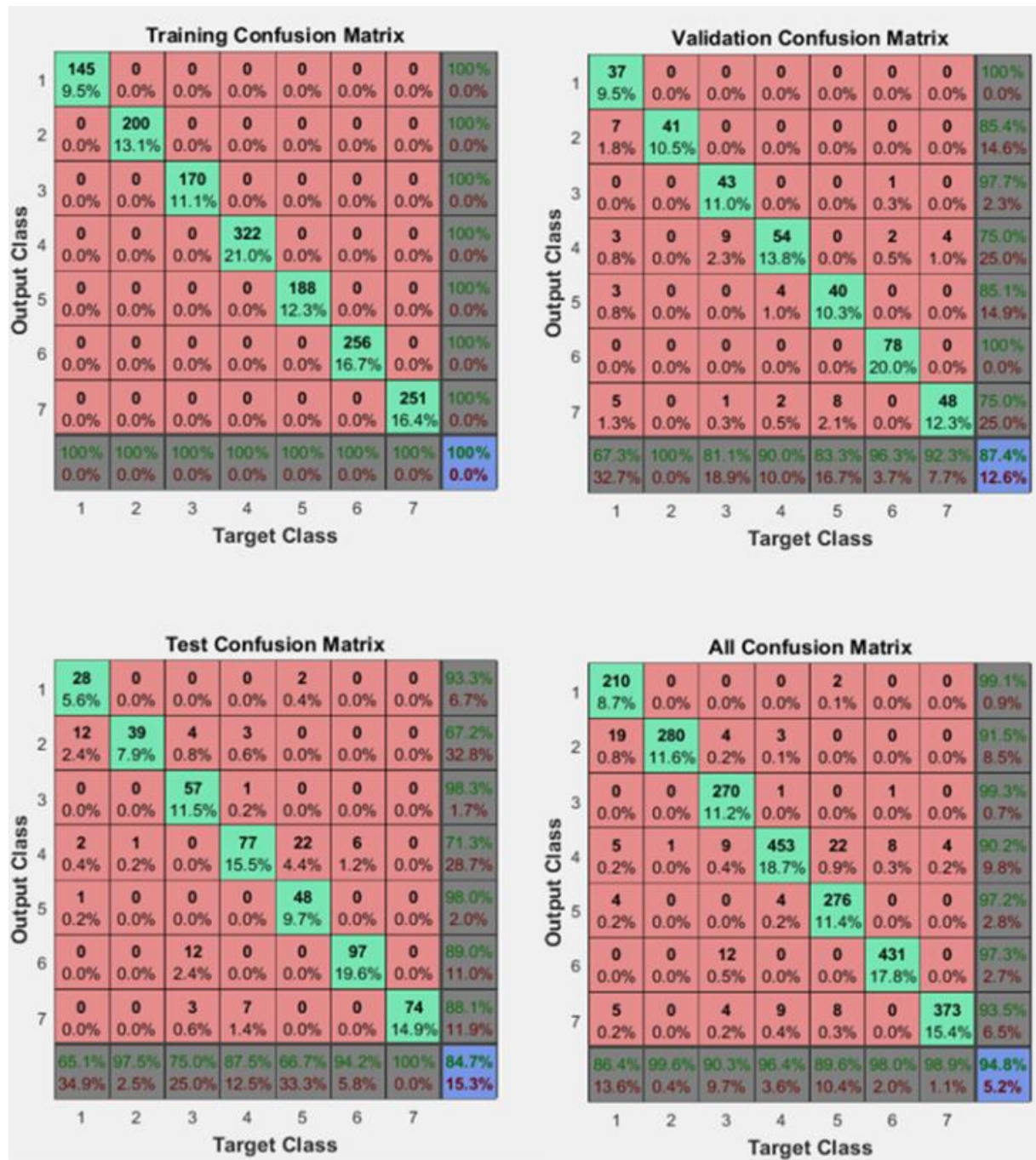
Κρυφά Στρώματα	[25 50]	[25 50]	[25 50]	[25 50]	[25 50]	[25 50]	[25 50]	[25 50]
Μέγεθος φίλτρων	100x100	100x100	100 x100	100x100	100x100	100x100	100x100	100x100
Αριθμός Φίλτρων	6+1	6+1	6+1	6+1	8+1	8+1	8+1	8+1
Προσανατολισμός	[0 45 90]	[0 45 90]	[0 45 90]	[0 45 90]	[0 45 90 135]	[0 45 90 135]	[0 45 90 135]	[0 45 90 135]
Μήκος Κύματος	[4 8]	[4 8]	[4 8]	[4 8]	[4 8]	[4 8]	[4 8]	[4 8]
1:Θυμός	97.9	67.3	53.5	83.1	100	67.3	65.1	86.4
2:Αηδία	100	100	100	100	100	100	97.5	99.6
3:Φόβος	100	84.9	69.7	89.6	100	81.1	75.0	90.3
4:Ουδέτερη	100	95.0	88.6	97.2	100	90.0	87.5	96.4
5:Λύπη	100	87.5	86.1	94.8	100	83.3	66.7	89.6
6:Ευτυχία	100	98.8	95.1	98.6	100	96.3	94.2	98.0
7:Εκπληξη	100	92.3	100	98.9	100	92.3	100	98.9
Σύνολο	99.8	89.7	86.3	95.4	100	87.4	84.7	94.8
Εποχές	148	148	148	148	230	230	230	230
	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο

Πίνακας 19: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και αριθμό νευρώνων [25 50] και στις δύο.



Εικόνα 17: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50], 6+1 φίλτρα Gabor 100x100 με προσανατολισμούς [0 45 90] και μήκη κύματος [4 8].

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου

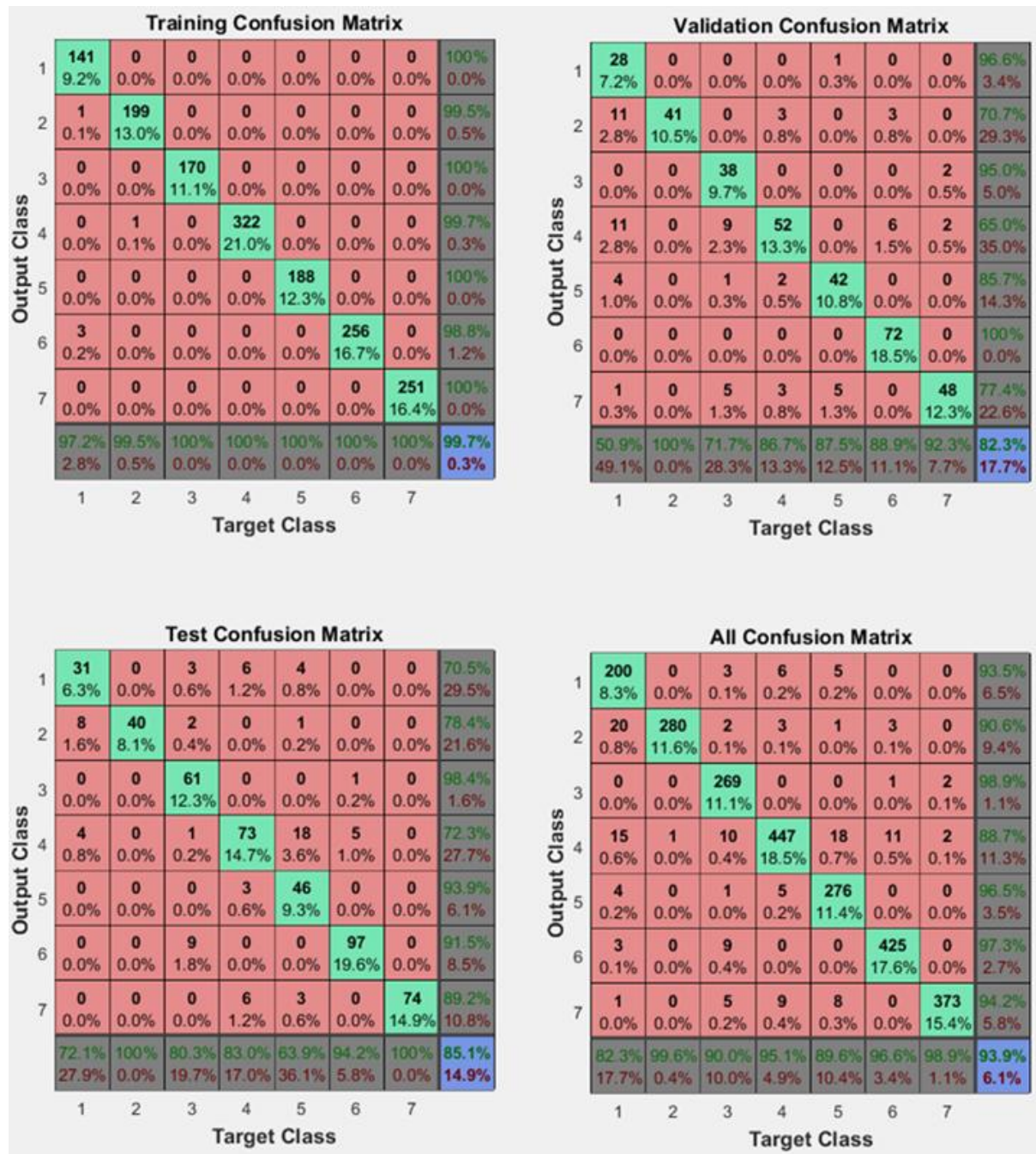


Εικόνα 18: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50], 8+1 φίλτρα Gabor 100x100 με προσανατολισμούς [0 45 90 135] και μήκη κύματος [4 8].

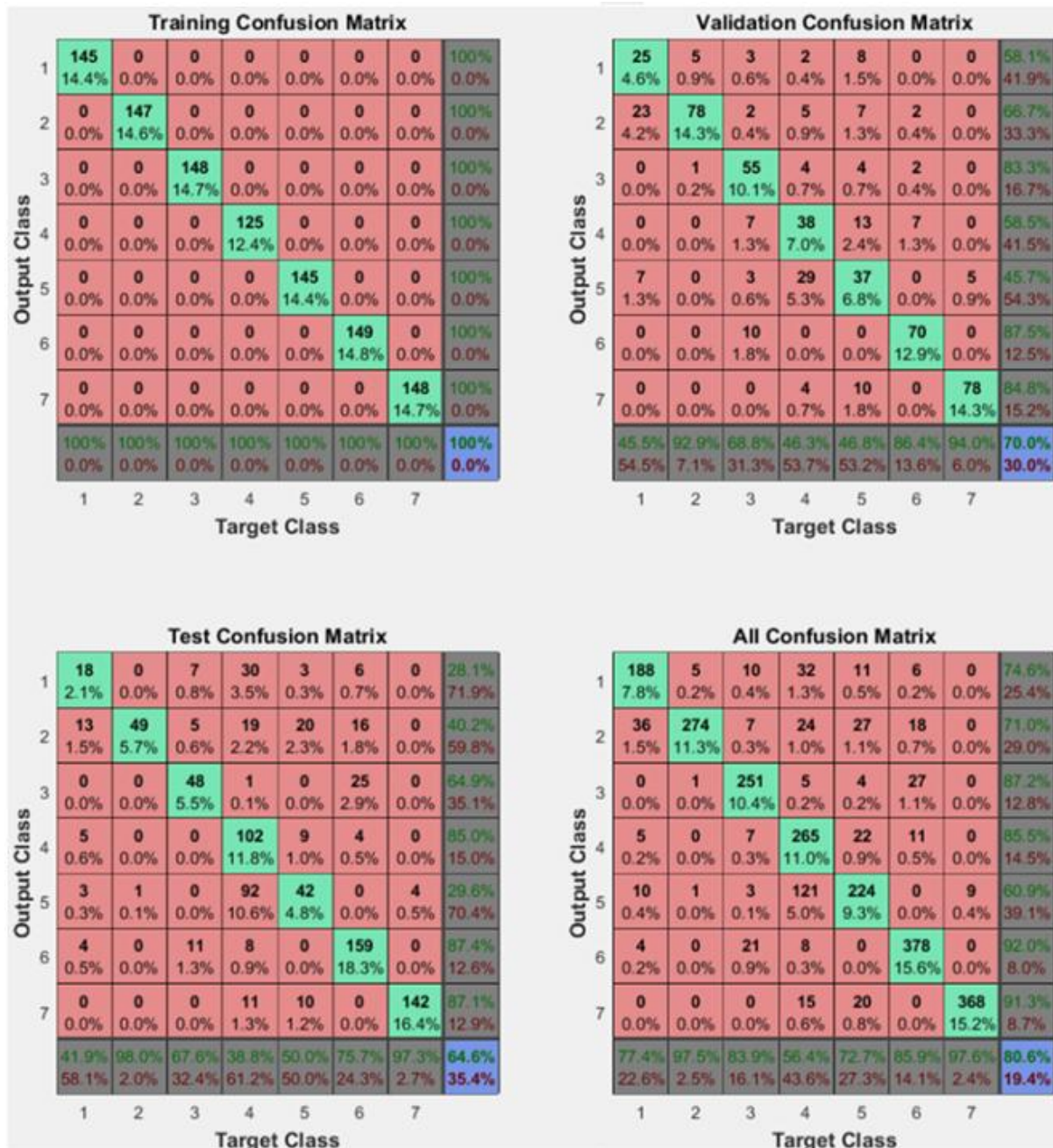
Κρυφά Στρώματα	[25 50]	[25 50]	[25 50]	[25 50]	[20 20]	[20 20]	[20 20]	[20 20]
Μέγεθος φίλτρων	100x100	100x100	100x100	100x100	50x50	50x50	50x50	50x50
Αριθμός Φίλτρων	6+1	6+1	6+1	6+1	40	40	40	40
Προσανατολισμός	[0 45 90]	[0 45 90]	[0 45 90]	[0 45 90]	[0 π/8 2π/8 3π/8 4π/8 5π/8 6π/8 7π/8]	[0 π/8 2π/8 3π/8 4π/8 5π/8 6π/8 7π/8]	[0 π/8 2π/8 3π/8 4π/8 5π/8 6π/8 7π/8]	[0 π/8 2π/8 3π/8 4π/8 5π/8 6π/8 7π/8]
Μήκος Κόματος	[2 4]	[2 4]	[2 4]	[2 4]	[2 4 6 8 10]	[2 4 6 8 10]	[2 4 6 8 10]	[2 4 6 8 10]
1:Θυμός	97.2	50.9	72.1	82.3	100	45.5	41.9	77.4
2:Αηδία	99.5	100	100	99.6	100	92.9	98.0	97.5
3:Φόβος	100	71.7	80.3	90.0	100	68.8	67.6	83.9
4:Ουδέτερη	100	86.7	83.0	95.1	100	46.3	38.8	56.4
5:Λύπη	100	87.5	63.9	89.6	100	46.8	50.0	72.7
6:Ευτυχία	100	88.9	94.2	96.6	100	86.4	75.7	85.9
7:Έκπληξη	100	92.3	100	98.9	100	94.0	97.3	97.6
Σύνολο	99.7	82.3	85.1	93.9	100	70.0	64.6	80.6
Εποχές	164	164	164	164	182	182	182	182
	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο

Πίνακας 20: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [25 50] και συνάρτηση απόδοσης Mean Squared Error στην πρώτη και αριθμό νευρώνων [30 45] και συνάρτηση απόδοσης Cross-Entropy στην δεύτερη.

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου



Εικόνα 19: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50], 8+1 φίλτρα Gabor 100x100 με προσανατολισμούς [0 45 90 135] και μήκη κύματος [4 8].



Εικόνα 20: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [20 30], 40 φίλτρα Gabor 50x50 με προσανατολισμούς [2 4 6 8 10] και μήκη κύματος [0 π/8 2π/8 3π/8 4π/8 5π/8 6π/8 7π/8].

5.4.3 Αποτελέσματα τρίτης υλοποίησης

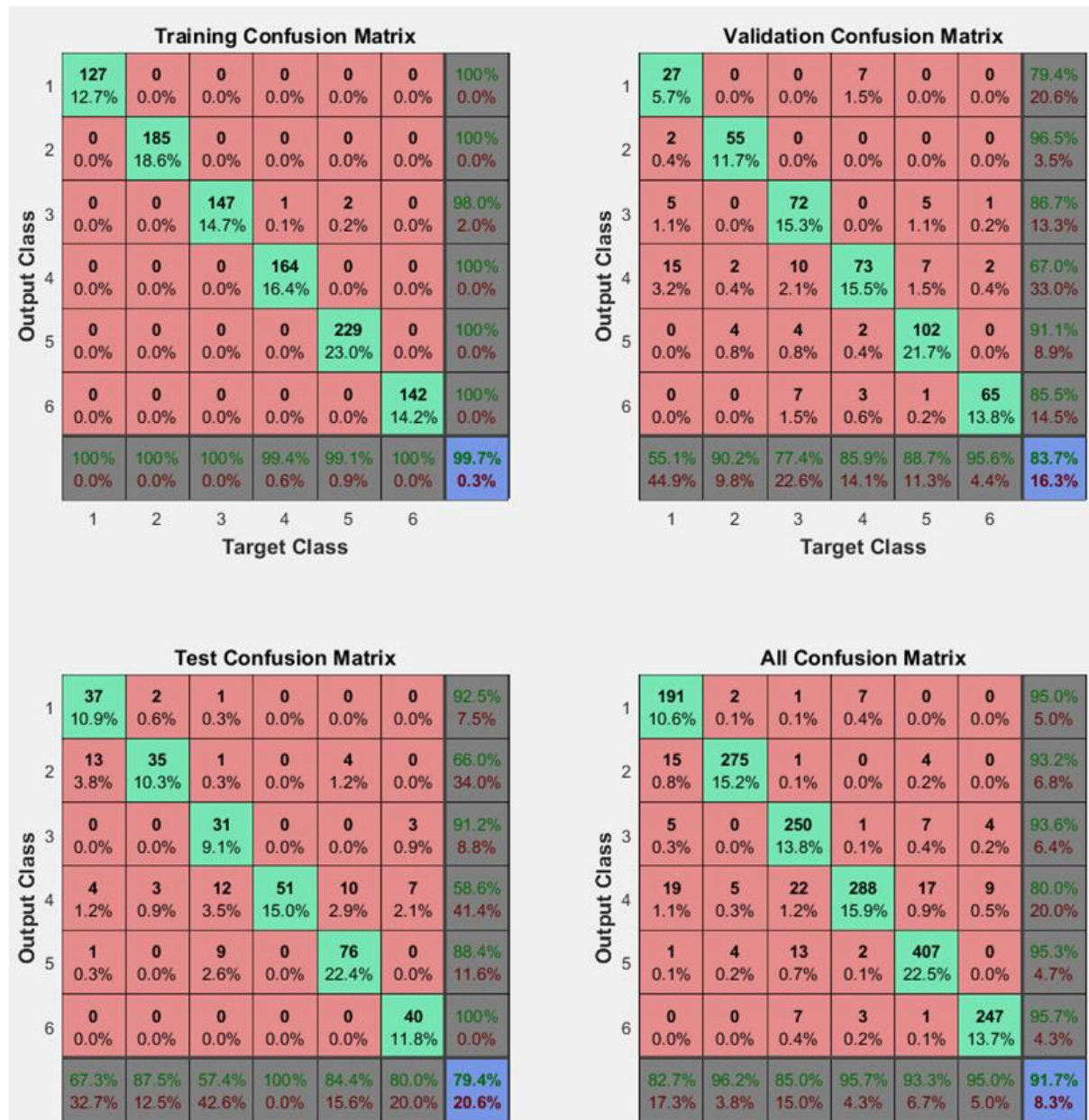
Τα αποτελέσματα της τρίτης υλοποίησης με την αφαίρεση της ουδέτερης έκφρασης σαν φίλτρο είχαν μία μικρή αύξηση στην αναγνώριση σε σχέση με την πρώτη υλοποίηση αλλά μία μείωση περισσότερη από 6% σε σχέση με την δεύτερη υλοποίηση. Πιστεύουμε ότι αυτό είναι λογικό καθώς τα φίλτρα Gabor είναι περισσότερα σε αριθμό και δίνουν μεγαλύτερο σύνολο χαρακτηριστικών που έχει να κάνει με προσανατολισμούς και συχνότητα στην εικόνα σε σχέση με το σύνολο χαρακτηριστικών που δίνει η αφαίρεση της ουδέτερης έκφραση. Οι εικόνες που χρησιμοποιήθηκαν για την εκπαίδευση, την δοκιμή και την επικύρωση του δικτύου ήταν λιγότερες αλλά και τα συναισθήματα στα οποία δοκιμάστηκε το δίκτυο ήταν επίσης λιγότερα. Εφόσον η ουδέτερη έκφραση χρησιμοποιείται σαν φίλτρο οι

εκφράσεις γίνονται έξι από επτά που έχει σαν αποτέλεσμα οι 470 εικόνες της ουδέτερης έκφρασης να μην περιλαμβάνονται στην εκπαίδευση, την επικύρωση και την δοκιμή του δικτύου.

Η καλύτερη δοκιμή σε αυτό το δίκτυο έβγαλε ποσοστό επιτυχούς αναγνώρισης 79.4% για τις δοκιμαστικές εικόνες και 83.7% στις εικόνες για την επικύρωση του δικτύου, με σχεδόν τέλεια εκπαίδευση και ποσοστό επιτυχούς εκπαίδευσης 99.7% μετά από 178 εποχές. Η διαμόρφωση του δικτύου ήταν με δύο κρυφά στρώματα 24 νευρώνων το καθένα, με Scaled Conjugate Gradient για την εκπαίδευση του δικτύου και Mean Squared Error για την απόδοση. Παρακάτω δίνονται κάποιοι πίνακες (Πίνακες 21 και 22) με τα καλύτερα αποτελέσματα καθώς και οι πίνακες σύγκυσης (Εικόνες 19-22) για κάθε μία από τις δοκιμές αυτές. Όπως στις άλλες δύο υλοποιήσεις έτσι και εδώ το συναίσθημα του θυμού συγχέεται με το συναίσθημα της αηδίας σε μεγάλο ποσοστό. Επίσης ένα μεγάλο ποσοστό εικόνων από το συναίσθημα του φόβου ταυτίζεται με το συναίσθημα της λύπης με αποτέλεσμα το ποσοστό αναγνώρισης να μην ξεπερνά το 80% σε όλες τις δοκιμές που έγιναν.

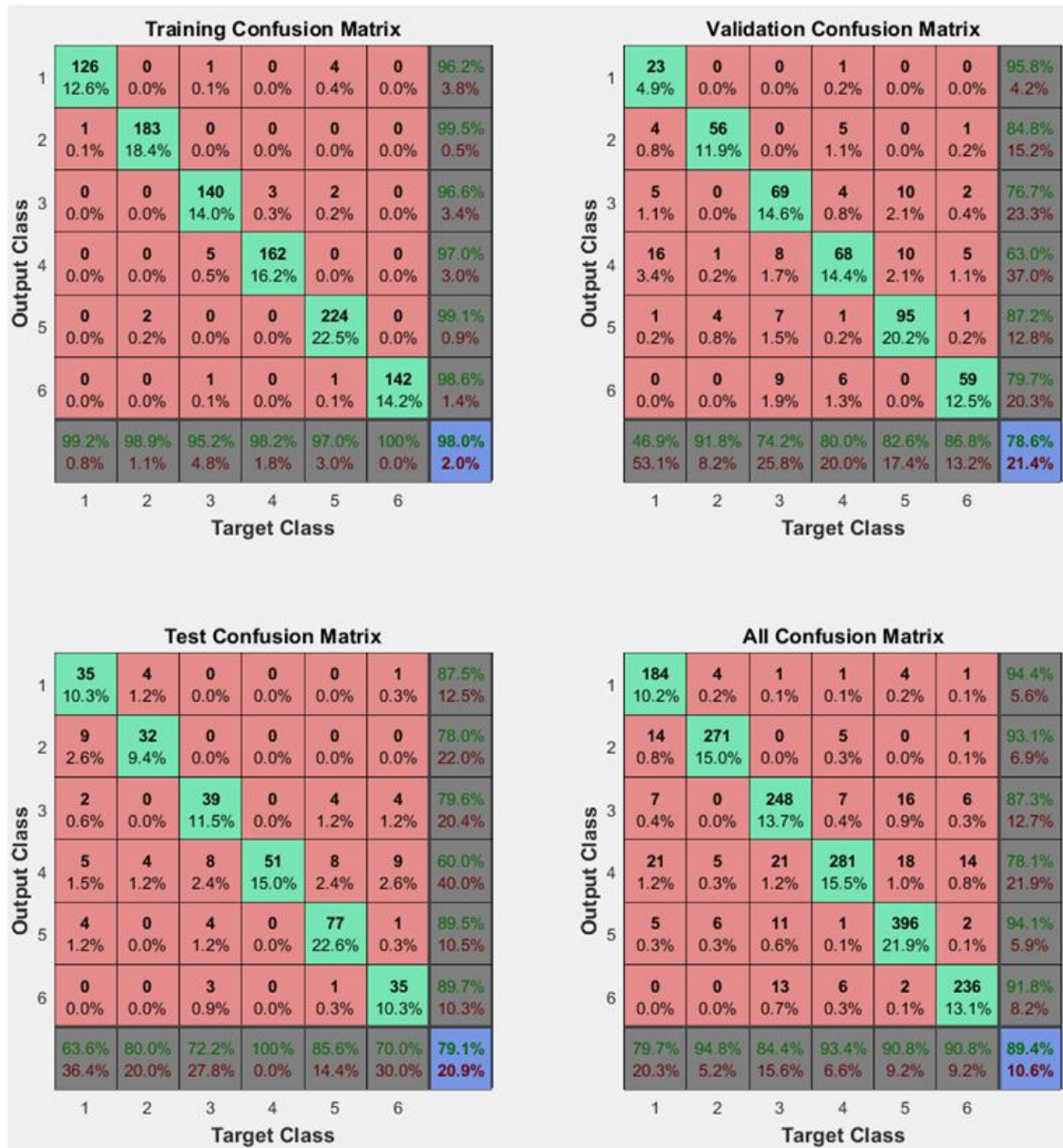
Κρυφά Στρώματα	[24 24]	[24 24]	[24 24]	[24 24]	[36 36]	[36 36]	[36 36]	[36 36]
1:Θυμός	100	55.1	67.3	82.7	99.2	46.9	63.6	79.7
2:Αηδία	100	90.2	87.5	96.2	98.9	91.8	80.0	94.8
3:Φόβος	100	77.4	57.4	85.0	95.2	74.2	72.2	84.4
4:Λύπη	99.4	85.9	100	95.7	98.2	80.0	100	93.4
5:Ευτυχία	99.1	88.7	84.4	93.3	97.0	82.6	85.6	90.8
6:Εκπληξη	100	95.6	80.0	95.0	100	86.8	70.0	90.8
Σύνολο	99.7	83.7	79.4	91.7	98.0	78.6	79.1	89.4
Εποχές	178	178	178	178	102	102	102	102
	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο

Πίνακας 21: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [24 24] και συνάρτηση απόδοσης Mean Squared Error στην πρώτη και αριθμό νευρώνων [36 36] και συνάρτηση απόδοσης Cross-Entropy στην δεύτερη.



Εικόνα 21: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [24 24].

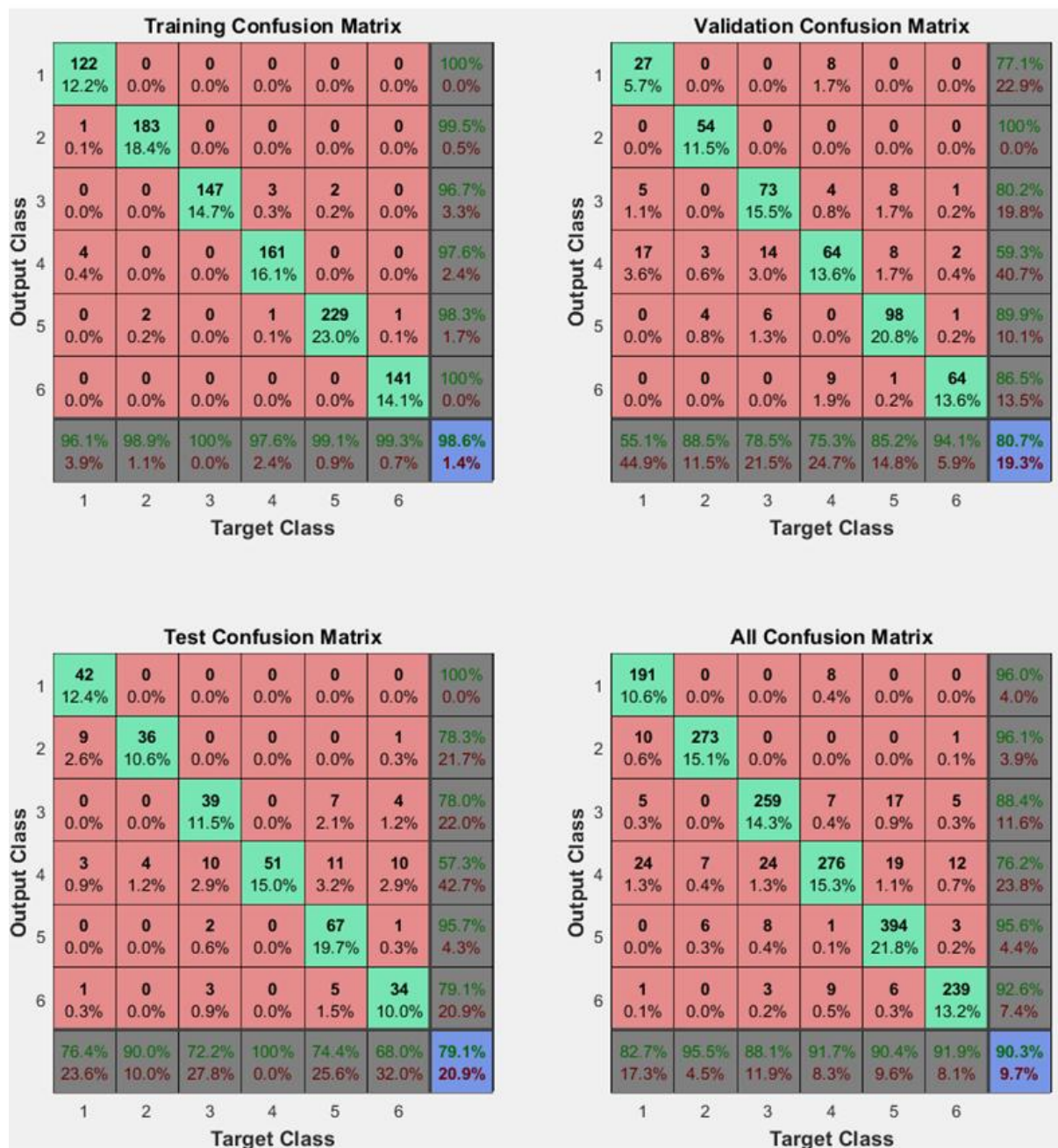
Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου



Εικόνα 22: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [36 36].

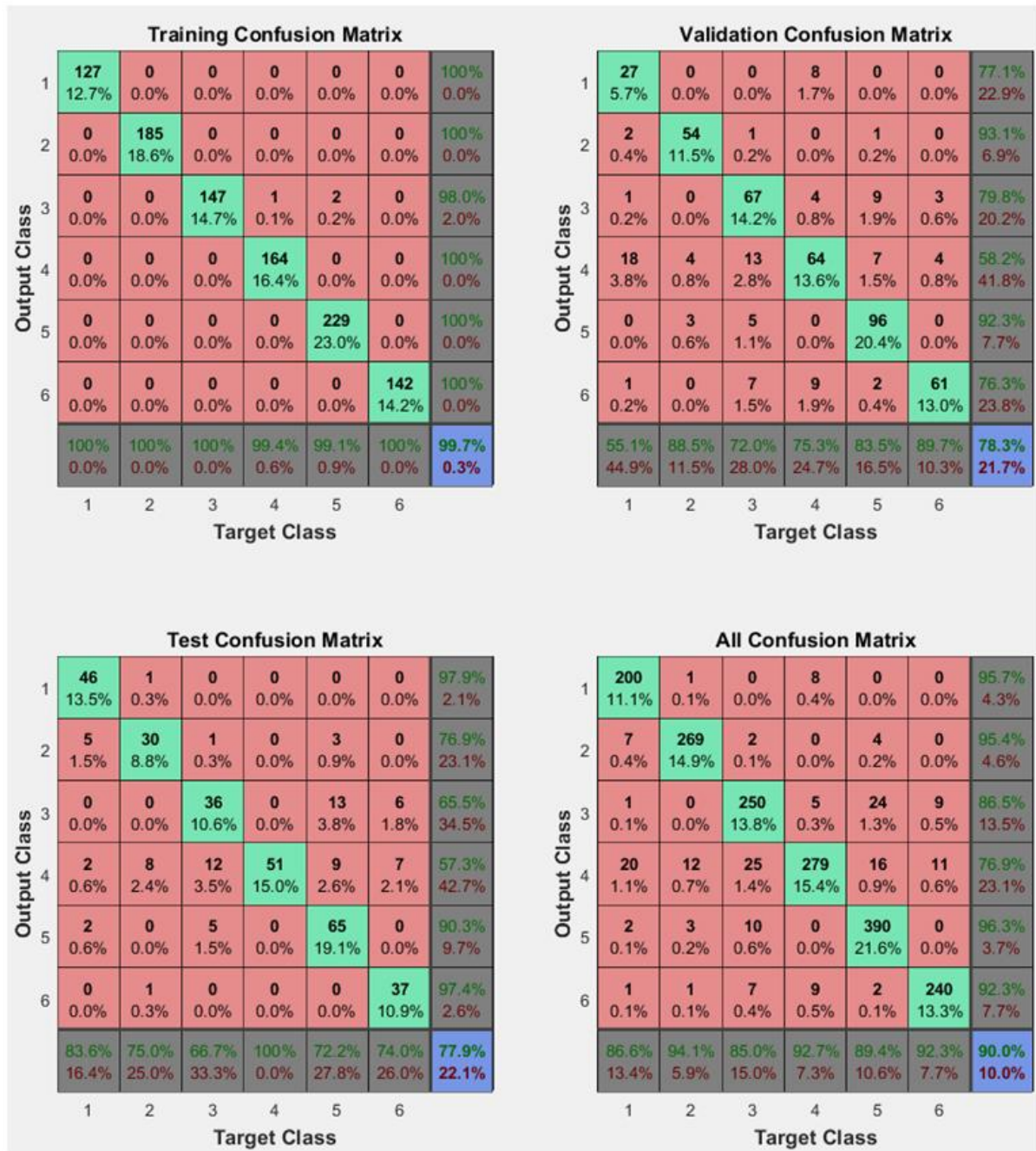
Κρυφά Στρώματα	[100 100]	[100 100]	[100 100]	[100 100]	[25 50]	[25 50]	[25 50]	[25 50]
1:Θυμός	96.1	55.1	76.4	82.7	100	55.1	83.6	86.6
2:Αηδία	98.9	88.5	90.0	95.5	100	88.5	75.0	94.1
3:Φόβος	100	78.5	72.2	88.1	100	72.0	66.7	85.0
4:Λύπη	97.6	75.3	100	91.7	99.4	75.3	100	92.7
5:Ευτυχία	99.1	85.2	74.4	90.4	99.1	83.5	72.2	89.4
6:Έκπληξη	99.3	94.1	68.0	91.9	100	89.7	74.0	92.3
Σύνολο	98.6	80.7	79.1	90.3	99.7	78.3	77.9	90.0
Εποχές	112	112	112	112	132	132	132	132
	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο	Εκπαίδευση	Επικύρωση	Δοκιμή	Σύνολο

Πίνακας 22: Ποσοστό επιτυχίας (σε %) για δύο δοκιμές του δικτύου με συνάρτηση εκπαίδευσης Scaled Conjugate Gradient, συνάρτηση απόδοσης Cross-Entropy, με 2 κρυφά στρώματα και στις δύο, με αριθμό νευρώνων [100 100] στην πρώτη και αριθμό νευρώνων [25 50] στην δεύτερη.



Εικόνα 23: Οι πίνακες σύγκρισης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [100 100].

Αναγνώριση συναισθημάτων από εκφράσεις του προσώπου



Εικόνα 24: Οι πίνακες σύγχυσης για την εκπαίδευση, επικύρωση, δοκιμή και το σύνολο του δικτύου με κρυφά στρώματα [25 50].

6. Συμπεράσματα

Κατά την διάρκεια υλοποίησης της Πτυχιακής Εργασίας δοκιμάστηκαν δύο μέθοδοι για την κατασκευή ενός αυτόματου συστήματος αναγνώρισης συναισθημάτων. Η πρώτη μέθοδος, με δύο διαφορετικές υλοποιήσεις, έκανε χρήση ιδιοδιανυσμάτων και ιδιοτιμών για την εξαγωγή χαρακτηριστικών και χρήση της Ευκλείδειας απόστασης σαν ταξινομητή για την εξαγωγή αποτελεσμάτων. Η δεύτερη μέθοδος έκανε γενική χρήση των τεχνητών νευρωνικών δικτύων που έχουν πολλές εφαρμογές και θετικά αποτελέσματα σε επεξεργασία εικόνας και ειδικότερα σε αυτόματη αναγνώριση συναισθημάτων ή εκφράσεων του προσώπου. Για την δεύτερη μέθοδο δοκιμάστηκαν τρεις διαφορετικές υλοποιήσεις.

Μετά από τις πολλές δοκιμές που έγιναν και για τις δύο μεθόδους υλοποίησης του αυτόματου συστήματος αναγνώρισης το εξής συμπέρασμα μπορεί να εξαχθεί. Τα τεχνητά νευρωνικά δίκτυα μπορούν να δώσουν αξιοπρεπή αποτελέσματα με ελάχιστη επεξεργασία στις εικόνες, λίγο καλύτερα με την χρήση αφαίρεσης της ουδέτερης έκφρασης σαν φίλτρο (~80%) και πάρα πολύ καλά (85%>) με την χρήση των φίλτρων Gabor. Εν αντιθέσει η χρήση ιδιοδιανυσμάτων και Ευκλείδειας απόστασης δεν έδωσε τα αποτελέσματα που περιμέναμε αλλά πολύ χαμηλότερα. Αυτό μπορεί να οφείλεται στην διαφορετική υλοποίηση του συστήματος, στο μεγαλύτερο μέγεθος και διαφορετικό τύπο της βάσης δεδομένων που δοκιμάσαμε σε σχέση με υλοποίηση των J. Kalita και K. Das [2]. Σε κάθε περίπτωση μπορούμε να πούμε ότι τα τεχνητά νευρωνικά δίκτυα στην ίδια βάση δεδομένων, κάνοντας χρήση του ίδιου ποσοστού εικόνων για εκπαίδευση και δοκιμή, έχουν ανώτερη προβλεπτική ικανότητα σε σχέση με την μέθοδο των ιδιοδιανυσμάτων και της Ευκλείδειας απόστασης.

Όσο αναφορά την βελτίωση των αποτελεσμάτων στο τεχνητό νευρωνικό δίκτυο θα μπορούσε να εκτελεστεί ένα είδος ανάλυσης κύριων συνιστωσών (principal component analysis) ώστε να μειωθεί το μέγεθος των χαρακτηριστικών που έδωσαν τα φίλτρα Gabor χωρίς να χαθεί μεγάλο μέρος πληροφορίας με αποτέλεσμα την μείωση των εισόδων του νευρωνικού δικτύου και πιθανώς την αύξηση της προβλεπτικής ικανότητας του δικτύου. Επίσης με περεταίρω δοκιμές στα φίλτρα Gabor για την εύρεση των βέλτιστων μηκών κύματος και προσανατολισμών θεωρούμε ότι τα αποτελέσματα έχουν χώρο βελτίωσης. Αλλαγές στον αριθμό νευρώνων ή τον αριθμό των στρωμάτων, καθώς και προσπάθεια για να μειωθούν οι εισοδοί παίζουν σημαντικό ρόλο στην αυξομείωση της προβλεπτικής ικανότητας του δικτύου.

Τέλος για την βελτίωση των αποτελεσμάτων με την χρήση της ουδέτερης έκφρασης σαν φίλτρο ένας αλγόριθμος για την τέλεια (ή όσο το δυνατόν καλύτερη) ευθυγράμμιση των δύο εικόνων (ουδέτερης εικόνας και εικόνας κάποιας έκφρασης) θα μπορούσε να δώσει μία καλύτερα φιλτραρισμένη τελική εικόνα που θα μπορούσε να αυξήσει τα αποτελέσματα. Η βελτίωση που είχε σαν αποτέλεσμα το δίκτυο της τρίτης υλοποίησης με νευρωνικά δίκτυα δεν είναι σίγουρο αν οφείλεται στο γεγονός ότι αντί για επτά συναισθήματα δοκιμάστηκε μόνο έξι ή αν η βελτίωση σε σχέση με την πρώτη υλοποίηση με νευρωνικά δίκτυα οφείλεται στο φίλτρο αφαίρεσης της ουδέτερης έκφρασης. Για να δοθεί μία τελική απάντηση στο ερώτημα αυτό θα πρέπει να γίνουν περαιτέρω δοκιμές με έξι εκφράσεις αντί για επτά και στις άλλες υλοποιήσεις με την χρήση τεχνητών νευρωνικών δικτύων.

Βιβλιογραφία

- [1] «Cohn-Kanade AU-Coded Expression Database,» 2017 03 2017. [Ηλεκτρονικό]. Available: <http://www.pitt.edu/~emotion/ck-spread.htm>.
- [2] J. Kalita και K. Das, «Recognition of Facial Expression Using Eigenvector Based Distributed Features and Euclidean Distance Based Decision Making Technique,» *IJACSA*, τόμ. 4, αρ. 2, 2013.
- [3] P. Ekman, *Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions*, Los Altos, CA 94023: MALOR BOOKS.
- [4] N. Ανδρέας, «Defining Emotion,» [Ηλεκτρονικό]. Available: <http://socialpolicy.gr/2013/12/%CF%80%CF%81%CE%BF%CF%83%CE%B4%CE%B9%CE%BF%CF%81%CE%AF%CE%B6%CE%BF%CE%BD%CF%84%CE%B1%CF%82-%CF%84%CE%BF-%CF%83%CF%85%CE%BD%CE%B1%CE%AF%CF%83%CE%B8%CE%B7%CE%BC%CE%B1.html>.
- [5] P. Ekman, “Facial Expressions” *Handbook of Cognition and Emotion*, New York: John Wiley & Sons Ltd., 1999.
- [6] S. V. e. a. Ioannou, *Emotion recognition through facial expression analysis based on a neurofuzzy network*, National Technical University of Athens, 27 March 2005.
- [7] I. Cohen, N. Sebe, A. Garg, L. S. Chen και T. S. Huang, «Facial Expression Recognition from Video Sequences: Temporal and Static Modeling,» *Computer Vision and Image Understanding*, αρ. 91, pp. 160-187, 2003.
- [8] R. Birdwhistell, «Kinesics and Context,» *Philadelphia, PA: University of Pennsylvania Press.*, 1970.
- [9] M. Suwa, N. Sugie και K. Fujimora, «A Preliminary Note on Pattern Recognition of Human Emotional Expression,» *Proc. Int’l Joint Conf. Pattern Recognition*, pp. 408-410, 1978..
- [10] K. Mase, «Recognition of Facial Expression from Optical Flow,» *IEICE Trans*, τόμ. 74, αρ. 10, pp. 3474-3483, 1991.
- [11] Z. Zeng, M. Pantic, G. I. Roisman και T. S. Huang, «A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions,» *IEEE Transactoins on Pattern Analysis and Machine Intelligence*, τόμ. 31, αρ. 1, January 2009.
- [12] M. S. M. R. a. J. H. S. Marryan Murtaza, «Analysis of Face Recognition under Varying Facial Expression: A Survey,» *Department of Computer Sciences, COMSATS Pakistan*, August 2012.

- [13] P. Ekman και et al., *Facial Action Coding System, A Human Face*, 2002.
- [14] C. Sumathi και et al., «Automatic Facial Expression Analysis: A Survey,» *IJCSES*, τόμ. 3, αρ. 6, December 2012.
- [15] E. Owusu και et al., «A neural-AdaBoost based facial expression recognition system,» *Expert Systems with Applications*, αρ. 41, pp. 3383-3390, 2014.
- [16] M. Mohammadi, E. Fatemizadeh και M. Mahoor, «PCA-based dictionary building for accurate facial expression recognition via sparse representation,» *Vis. Commun. Image R.*, αρ. 25, p. 1082–1092, 2014.
- [17] «MMI facial expression database,» [Ηλεκτρονικό]. Available: <http://mmifacedb.eu/>. [Πρόσβαση 03 03 2017].
- [18] «Bosphorus 3D face database,» [Ηλεκτρονικό]. Available: <http://bosporus.ee.boun.edu.tr/default.aspx>. [Πρόσβαση 03 03 2017].
- [19] G. Manisha Satone, «Selection of Eigenvectors for Face Recognition,» *IJACSA*, τόμ. 4, αρ. 3, 2013.
- [20] «GTAV face database,» [Ηλεκτρονικό]. Available: <https://gtav.upc.edu/en/research-areas/face-database>. [Πρόσβαση 03 03 2017].
- [21] «Deep learning, Convolution Neural Networks,» 19 02 2017. [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Deep_learning#Convolutional_neural_networks.
- [22] S. Alizadeh και A. Fazel, *Convolutional Neural Networks for Facial Expression Recognition*.
- [23] «Batch Normalization,» 06 03 2017. [Ηλεκτρονικό]. Available: <https://kratzert.github.io/2016/02/12/understanding-the-gradient-flow-through-the-batch-normalization-layer.html>.
- [24] «Convolutional neural network,» 06 03 2017. [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Convolutional_neural_network#Dropout.
- [25] «Softmax Function,» 06 03 2017. [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Softmax_function.
- [26] D. Duncan και et al, *Facial Emotion Recognition in Real Time*, Stanford.
- [27] P. Burkert, F. Trier, M. Z. Afzal, A. Dengel και M. Liwicki, *DeXpression: Deep Convolutional Neural Network for Expression Recognition*, Kaiserslautern, Germany: German Research Center for Artificial Intelligence (DFKI), 17 Aug 2016.
- [28] G. Lavi και et al., *Emotion Recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns*, Israel: The Open University of Israel, 2015.

- [29] «MATLAB,» [Ηλεκτρονικό]. Available: <https://en.wikipedia.org/wiki/MATLAB>. [Πρόσβαση 22 02 2017].
- [30] MathWorks, *Image Processing Toolbox : User's Guide R2016b*.
- [31] H. N. Koivo, *Basics using MATLAB Neural Network Toolbox*, Feb.1 2008.
- [32] H. Demuth και M. Beale, *Neural Network Toolbox :For Use with MATLAB*, Mathworks.
- [33] M. T. H. H. B. D. Mark Hudson Beale, *Neural Network Toolbox: User's Guide*, Mathworks.
- [34] «Mean value and mean vector,» [Ηλεκτρονικό]. Available: https://el.wikipedia.org/wiki/%CE%9C%CE%AD%CF%83%CE%BF%CF%82_%CF%8C%CF%81%CE%BF%CF%82. [Πρόσβαση 02 02 2017].
- [35] «Covariance matrix,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Covariance_matrix. [Πρόσβαση 02 02 2017].
- [36] «Eigenvalues and Eigenvectors,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors. [Πρόσβαση 02 02 2017].
- [37] «Eigenvalues and Eigenvectors,» [Ηλεκτρονικό]. Available: https://el.wikipedia.org/wiki/%CE%99%CE%B4%CE%B9%CE%BF%CF%84%CE%B9%CE%BC%CE%AD%CF%82_%CE%BA%CE%B1%CE%B9_%CE%B9%CE%B4%CE%B9%CE%BF%CE%B4%CE%B9%CE%B1%CE%BD%CF%8D%CF%83%CE%BC%CE%B1%CF%84%CE%B1. [Πρόσβαση 02 02 2017].
- [38] A. P. Matthew Turk, *Eigenfaces for Recognition*, Massachusetts Institute of Technology, Vision and Modeling Group, The Media Laboratory. .
- [39] «Euclidean Distance,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Euclidean_distance. [Πρόσβαση 02 02 2017].
- [40] «Euclidean Space,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Euclidean_space. [Πρόσβαση 02 02 2017].
- [41] «Viola-Jones algorithm,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Viola%E2%80%93Jones_object_detection_framework. [Πρόσβαση 02 02 2017].
- [42] «Image Crop in matlab,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/images/ref/imcrop.html#buxpsc-3>. [Πρόσβαση 05 02 2017].
- [43] «Resize Image in matlab,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/images/ref/imresize.html>. [Πρόσβαση 05 02 2017].
- [44] «Average or mean value of array in matlab,» [Ηλεκτρονικό]. Available:

- <https://www.mathworks.com/help/matlab/ref/mean.html>. [Πρόσβαση 05 02 2017].
- [45] «Covariance in matlab,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/matlab/ref/cov.html>. [Πρόσβαση 05 02 2017].
- [46] «Subset of eigenvalues and eigenvectors,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/matlab/ref/eigs.html>. [Πρόσβαση 05 02 2017].
- [47] «Euclidean norm in matlab,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/matlab/ref/norm.html>. [Πρόσβαση 08 02 2017].
- [48] «Synapse,» [Ηλεκτρονικό]. Available: <https://en.wikipedia.org/wiki/Synapse>. [Πρόσβαση 06 03 2017].
- [49] G. Papadourakis, *Introduction to Neural Networks, PowerPoint Notes*, Heraklion: Technological Educational Institute of Crete, Neural Networks Laboratory , Read 06/03/2017.
- [50] «Neural Networks: feedback NN,» [Ηλεκτρονικό]. Available: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Feedback%20networks. [Πρόσβαση 06 03 2017].
- [51] «Gabor filter,» [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Gabor_filter. [Πρόσβαση 02 02 2017].
- [52] «Gabor filter and gabor filter bank in matlab,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/images/ref/gabor.html>. [Πρόσβαση 02 02 2017].
- [53] Mathworks, «Pattern recognition network,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/nnet/ref/patternnet.html>. [Πρόσβαση 06 02 2017].
- [54] Mathworks, «Neural Network Performance : Crossentropy,» [Ηλεκτρονικό]. Available: https://www.mathworks.com/help/nnet/ref/crossentropy.html?searchHighlight=crossentropy&s_tid=doc_srchtile. [Πρόσβαση 06 02 2017].
- [55] Mathworks, «Neural Network Performance: Mean squared normalized error performance function,» [Ηλεκτρονικό]. Available: <https://www.mathworks.com/help/nnet/ref/mse.html>. [Πρόσβαση 06 02 2017].
- [56] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz και J. Taylor, «Emotion Recognition in Human-Computer Interaction,» *IEEE Signal Processing Magazine*, Τόμ. %1 από %21053-5888, 01 January 2001.
- [57] R. Adolphs, *Recognizing Emotion from Facial Expression: Psychological and Neurological Mechanisms*, University of Iowa, College of Medicine , 2002.
- [58] M. S. Nixon και A. S. Aguado, *Feature Extraction and Image Processing*, 2002.
- [59] I. Cohen, N. Sebe, A. Garg, L. S. Chen και T. S. Huang, «Facial Expression Recognition from Video Sequences: Temporal and Static Modeling,» *Computer Vision and Image*

- Understanding*, αρ. no.91, pp. 160-187, 2003.
- [60] C. Busso, Z. deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann και S. Narayanan, *Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information*, 2004.
- [61] I. C. a. T. S. H. Nicu Sebe, *Multimodal Emotion Recognition*, June 18, 2004.
- [62] M. Wollmer, A. Metallinou, F. Eyben, B. Schuller και S. Narayanan, *Context-Sensitive Multimodal Emotion Recognition from Speech and Facial Expression using Bidirectional LSTM Modeling*, INTERSPEECH 2010.
- [63] V. K. Azhnty, S. D.Kreibig, J. J. Gross, W. T. Roth και F. H. Wilhelm, «An Affective Computing Approach to Physiological Emotion Specificity: Toward Subject-Independent Classification of film-induced Emotions,» *Psychophysiology*, αρ. 48, 2011.
- [64] H. Gunes, B. Schuller, M. Pantic και R. Cowie, *Emotion Representation, Analysis and Synthesis in Continuous Space: A Survey*, European Community's 7th Framework Programme, 2011.
- [65] M. Murtaza, M. Sharif, M. Raza και J. H. Shah, *Analysis of Face Recognition under Varying Facial Expression: A Survey*, Department of Computer Sciences, COMSATS Pakistan, August 2012.
- [66] K. A. Lindquist, *Emotions Emerge from More Basic Psychological Ingredients: A Modern Psychological Constructionist Model*, Sage Publications, Sept, 2013.
- [67] M. S. Sarfraz, O. Hellwich και Z. Riaz, *Feature Extraction and Representation for Face Recognition*.
- [68] F. Ringeval, B. Schuller, M. Valstar, S. Jaiswal, E. Marchi, D. Lalanne, R. Cowie και M. Pantic, *AV+EC 2015- The First Affect Recognition Challenge Bridging Across Audio, Video, and Physiological Data*.
- [69] J. F. Cohn, *Foundations Of Human Computing: Facial Expression and Emotion*, Department of Psychology, University of Pittsburgh.
- [70] R. C. Gonzalez, R. E. Woods και S. L. Eddins, *Digital Image Processing Using MATLAB Second Edition*, Gatesmark Publishing.
- [71] M. D. P. K. B. Asit Kumar Datta, *Face Detection and Recognition: Theory and Practice*, CRC Press.
- [72] J. Nazari και O. K. Ersoy, *Implementation of back-propagation neural networks with MatLab*, Purdue University.
- [73] S. Theodoridis και K. Koutroumbas, *An Introduction to Pattern Recognition: A MATLAB Approach*, Academic Press, an imprint of Elsevier.

- [74] M. Pantic και M. S. Barlett, «Machine Analysis of Facial Expressions Open Access Database,» [Ηλεκτρονικό]. Available: www.i-techonline.com. [Πρόσβαση 05 12 2016].
- [75] P. Ekman., Darwin, Deception, and Facial Expression, San Francisco: Department of Psychiatry, University of California.
- [76] P. Koopman, «“How to Write an Abstract” Carnegie Mellon University,» October 1997 . [Ηλεκτρονικό]. Available: <https://users.ece.cmu.edu/~koopman/essays/abstract.html>. [Πρόσβαση 03 12 2016].
- [77] «Emotion,» [Ηλεκτρονικό]. Available: <https://en.wikipedia.org/wiki/Emotion>. [Πρόσβαση 02 02 2017].
- [78] C. Lee και A. Elgammal, «Facial Expression Analysis Using Nonlinear Decomposable Generative Models,» *Proc. Second IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures (AMFG)*, 2005.
- [79] N. Sebe, M. Lew, I. Cohen, Y. Sun, T. Gevers και T. Huang, «Authentic Facial Expression Analysis,» *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR)*, 2004.
- [80] A. Ashraf, S. Lucey, J. Cohn, T. Chen, Z. Ambadar, K.Prkachin, P. Solomon και B. Theobald, «The Painful Face: Pain Expression Recognition Using Active Appearance Models,» *Proc.Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, pp. 9-14, 2007.
- [81] M. Bartlett, G. Littlewort, P. Braathen, T. Sejnowski και J. Movellan, «A Prototype for Automatic Recognition of Spontaneous Facial Actions,» *Advances in Neural Information Processing Systems*, τόμ. 15, pp. 1271-1278, 2003.
- [82] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel και J. Movellan, «Fully Automatic Facial Action Recognition in Spontaneous Behavior,» *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06)*, pp. 223-230, 2006.
- [83] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel και J. Movellan, «Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior,» *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '05)*, pp. 568-573, 2005.
- [84] J. Cohn, L. Reed, Z. Ambadar, J. Xiao και T. Moriyama, «Automatic Analysis and Recognition of Brow Actions and Head Motion in Spontaneous Facial Behavior,» *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '04)*, τόμ. 1, pp. 610-616, 2004.
- [85] B. Fasel, F. Monay και D. Gatica-Perez, «Latent Semantic Analysis of Facial Action Codes for Automatic Facial Expression Recognition,» *Proc. Sixth ACM Int'l Workshop Multimedia Information Retrieval (MIR '04)*, pp. 181-188, 2004.
- [86] H. Gunes και M. Piccardi, «Affect Recognition from Face and Body: Early Fusion versus Late Fusion,» *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '05)*, pp. 3437-3443, 2005.

- [87] Q. Ji, P. Lan και C. Looney, «A Probabilistic Framework for Modeling and Real-Time Monitoring Human Fatigue,» *IEEE Systems, Man, and Cybernetics Part A*, τόμ. 36, αρ. 5, pp. 862-875, 2006.
- [88] R. E. Kaliouby και P. Robinson, «Real-Time Inference of Complex Mental States from Facial Expression and Head Gestures,» *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '04)*, τόμ. 3, p. 154, 2004.
- [89] A. Kapoor, W. Bursleson και R. Picard, «Automatic Prediction of Frustration,» *Int'l J. Human-Computer Studies*, τόμ. 65, αρ. 8, pp. 724-736, 2007.
- [90] A. Kapoor και R. Picard, «Multimodal Affect Recognition in Learning Environment,» *Proc. 13th ACM Int'l Conf. Multimedia (Multimedia '05)*, pp. 677-682, 2005.
- [91] G. Littlewort, M. Bartlett και K. Lee, «Faces of Pain: Automated Measurement of Spontaneous Facial Expressions of Genuine and Posed Pain,» *Proc. Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, pp. 15-21, 2007.
- [92] S. Lucey, A. Ashraf και J. Cohn, «Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face,» *Face Recognition, K. Delac, and M. Grgic, eds, I-Tech Education and Publishing*, pp. 275-286,, 2007.
- [93] M. Pantic και L. Rothkrantz, «Case-Based Reasoning for User-Profiled Recognition of Emotions from Face Images,» *Proc. 12th ACM Int'l Conf. Multimedia (Multimedia '04)*, pp. 391-394, 2004.
- [94] M. Pantic και I. Patras, «Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments Form Face Profile Image Sequences,» *IEEE Trans. Systems, Man, and Cybernetics Part B*, τόμ. 36, αρ. 2, pp. 433-449, 2006.
- [95] M. Pantic και L. Rothkrantz, «Facial Action Recognition for Facial Expression Analysis from Static Face Images,» *IEEE Trans. Systems, Man, and Cybernetics Part B*, τόμ. 34, αρ. 3, pp. 1449-1461, 2004.
- [96] M. Pantic και L. J. Rothkrantz, «Automatic Analysis of Facial Expressions: The State of the Art,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, τόμ. 22, αρ. 12, Dec 2000.
- [97] R. Reisenzein, E. Hudlicka, M. dastani, J. Gratch, K. Hindriks, E. Lorini και J.-J. C. Meyer, «Computational Modeling of Emotion: Toward Improving the Inter- and Intradisciplinary Exchange,» *IEEE Transactions on Affective Computing*, τόμ. 4, αρ. 3, July-September 2013.
- [98] J.-H. Seo, Y.-B. Park και Y.-J. Park, «Reliable Facial Color Analysis using a digital camera and its relationship with pathological patterns: A pilot study,» *European Journal of Integrative Medicine*, τόμ. 6, pp. 322-327., 2014.
- [99] Y. Tong, W. Liao και Q. Ji, «Facial Action Unit Recognition by Exploiting Their Dynamics and Semantic Relationships,» *IEEE Trans. Pattern Analysis and Machine Intelligence*, τόμ.

29, αρ. 10, pp. 1683-1699, 2007.

- [100] M. Valstar, M. Pantic και I. Patras, «Motion History for Facial Action Detection from Face Video,» *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '04)*, τόμ. 1, pp. 635-640, 2004.
- [101] M. Valstar, M. Pantic, Z. Ambadar και J. Cohn, «Spontaneous versus Posed Facial Behavior: Automatic Analysis of Brow Actions,» *Proc. Eight Int'l Conf. Multimodal Interfaces (ICMI '06)*, pp. 162-170, 2006.
- [102] M. Valstar, H. Gunes και M. Pantic, «How to Distinguish Posed from Spontaneous Smiles Using Geometric Features,» *Proc. Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, pp. 38-45, 2007.
- [103] H. Wang και N. Ahuja, «Facial Expression Decomposition,» *Proc. Ninth IEEE Int'l Conf. Computer Vision (ICCV '03)*, p. 958, 2003.
- [104] J. Wang, L. Yin, X. Wei και Y. Sun, «3D Facial Expression Recognition Based on Primitive Surface Feature Distribution,» *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '06)*, τόμ. 2, pp. 1399-1406, 2006.
- [105] Z. Wen και T. Huang, «Capturing Subtle Facial Motions in 3D Face Tracking,» *Proc. Ninth IEEE Int'l Conf. Computer Vision (ICCV'03)*, pp. 1343-1350, 2003.
- [106] J. Whitehill και C. Omlin, «Haar Features for FACS AU Recognition,» *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06)*, pp. 217-222, 2006.
- [107] M. Yeasin, B. Bulot και R. Sharma, «Recognition of Facial Expressions and Measurement of Levels of Interest from Video,» *IEEE Trans. Multimedia*, τόμ. 8, αρ. 3, pp. 500-507, June 2006.
- [108] Z. Zeng, Y. Fu, G. Roisman, Z. Wen, Y. Hu και T. Huang, «Spontaneous Emotional Facial Expression Detection,» *J. Multimedia*, τόμ. 1, αρ. 5, pp. 1-8, 2006.