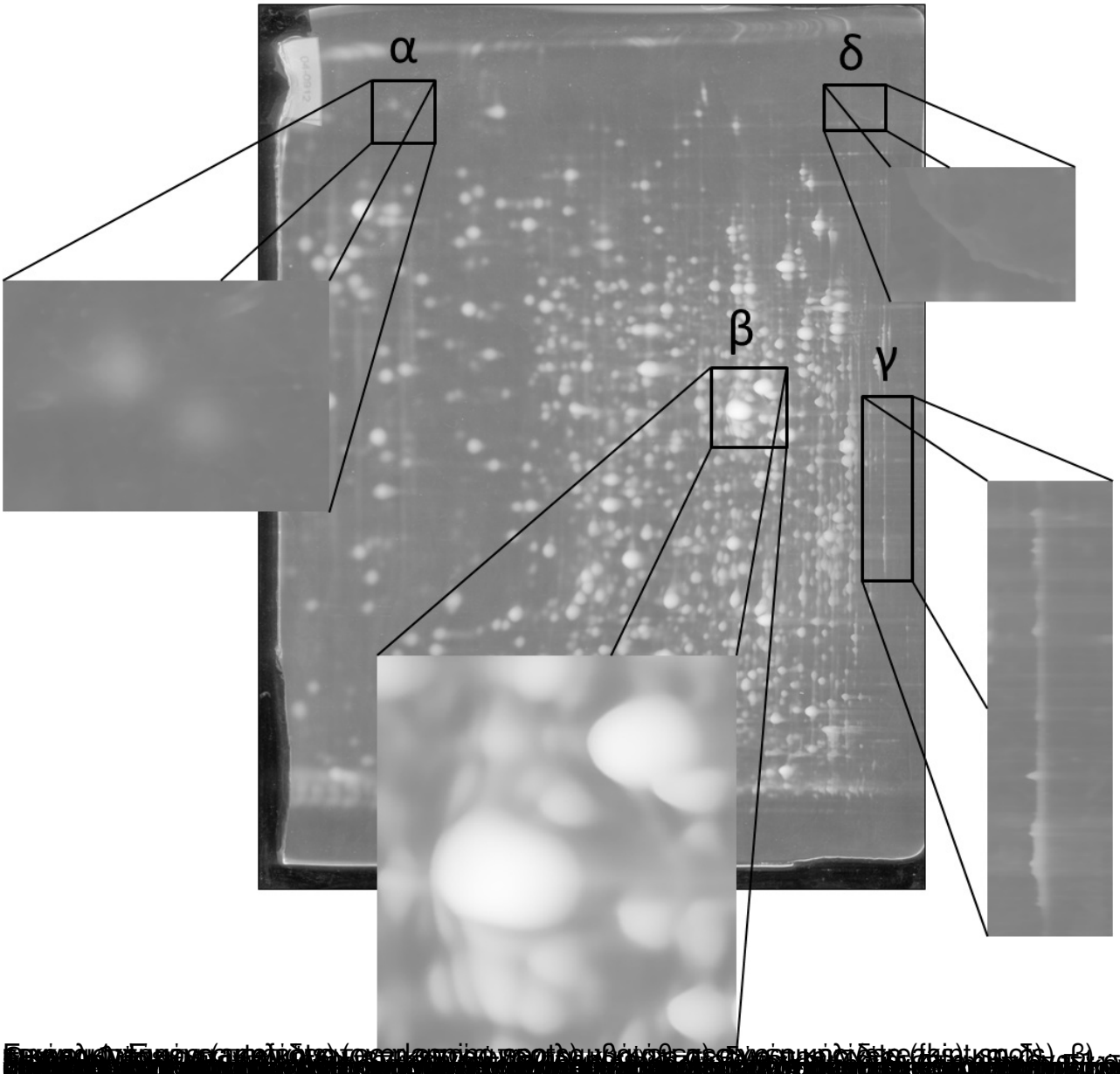
Έκθεση συγκριτικής μελέτης σύγχρονων μεθόδων επεξεργασίας εικόνων πηκτωμάτων

Στα πλαίσια της διδακτορικής έρευνας με Κ.Α. 70/3/11039 μελετήθηκαν και υλοποιήθηκαν σύγχρονες μέθοδοι επεξεργασίας εικόνων πηκτωμάτων. Οι μέθοδοι στοχεύουν στην ανίχνευση και κατάτμηση πρωτεϊνικών κηλίδων με σκοπό τον εντοπισμό εκείνων που έχουν υποστεί τροποποιήσεις στην έκφραση κατά την διάρκεια μίας νόσου. Η ποσοτική ανάλυση των εν λόγω πρωτεϊνών οδηγεί στην ανάπτυξη εφαρμογών σε επίπεδο διάγνωσης καθώς και σε επίπεδο θεραπείας ασθενειών με την παραγωγή νέων φαρμάκων. Οι εικόνες πηκτωμάτων προκύπτουν ως αποτέλεσμα μίας προηγμένης τεχνικής απεικόνισης που ονομάζεται δυσδιάστατη ηλεκτροφόρηση πηκτωμάτων [1]. Απεικονίζουν πρωτεΐνες οι οποίες εμφανίζονται ως λευκές κηλίδες σε σκοτεινό ανομοιογενές υπόβαθρο και αντίστροφα. Τα χρώματα του πηκτώματος μετατρέπονται σε έντασεις διαβαθμίσεων του γκρι (Εικόνα 1).



Μέθοδοι Ανίχνευσης

Τα σφάλματα που υπεισέρχονται στο στάδιο της ανίχνευσης περιλαμβάνουν κυρίως: α) την μη ανίχνευση πρωτεϊνικών κηλίδων όταν η ένταση της κηλίδας είναι χαμηλότερη από αυτήν του ανομοιογενούς υποβάθρου, β) την λανθασμένη ανίχνευση τεχνουργημάτων ή κορυφών του υποβάθρου ως κηλίδες και γ) την μη ανίχνευση πρωτεϊνικών κηλίδων όταν είναι επικαλυπτόμενες και το τοπικό μέγιστο της μίας επικαλύπτει το τοπικό μέγιστο της άλλης.

Η μέθοδος πυραμίδας [2] είναι μία σύγχρονη μέθοδος ανίχνευσης πρωτεϊνικών κηλίδων. Αρχικά, υπολογίζεται ο μέσος όρος των εντάσεων των εικονοστοιχείων στις εικόνες του δείγματος. Εκτελείται ανίχνευση των τοπικών μεγίστων της έντασης στην εικόνα που περιέχει τους μέσους όρους και οι συντεταγμένες τους αποτελούν τις θέσεις όπου

εφαρμόζονται οι πυραμίδες. Εάν δύο ή περισσότερες πυραμίδες βρεθούν στις ίδιες συντεταγμένες, τότε επικρατεί εκείνη με την υψηλότερη ένταση. Με τον τρόπο αυτό αφαιρούνται οι διπλές κορυφές και μειώνεται το πρόβλημα της επικάλυψης. Εντούτοις, πριν την ανίχνευση των τοπικών μεγίστων αφαιρούνται οι οριζόντιες και κάθετες γραμμές από την εικόνα γιατί οδηγούν σε λανθασμένη εκτίμηση των αποτελεσμάτων. Ακόμη, η παρουσία θορύβου δυσχεραίνει τον σωστό εντοπισμό των συντεταγμένων των κηλίδων αφού έχει ως αποτέλεσμα την λανθασμένη ανίχνευση των κορυφών του ανομοιογενούς υποβάθρου ως πρωτεΐνες.

Η μέθοδος των ενεργών περιγραμμάτων [3] παρουσιάζει αρκετά πλεονεκτήματα στην ανίχνευση καθώς και στην κατάτμηση εικόνων πηκτωμάτων. Αποτελείται από συνεχείς ανοιχτές ή κλειστές καμπύλες που εξελίσσονται σύμφωνα με την ελαχιστοποίηση μίας κατάλληλα επιλεγμένης συνάρτησης ενέργειας χωρίς να απαιτούν πρόσθετες λειτουργίες συνένωσης ακμών και επαρκή αντίθεση δομής ενδιαφέροντος σε σχέση με το υπόβαθρο. Επίσης, παραμένουν ανεπηρέαστα στην ανομοιογένεια του υποβάθρου και ελαττώνουν τις συνέπειες του θορύβου. Επιπλέον, υιοθετούν τοπολογική προσαρμοστικότητα, πράγμα που είναι σημαντικό προτέρημα διότι μία εικόνα πηκτώματος περιέχει περίπου χίλιες πρωτεϊνικές κηλίδες.

Στα πλαίσια της διδακτορικής έρευνας αναπτύχθηκε και υλοποιήθηκε μία πρωτότυπη μέθοδος ανίχνευσης κηλίδων και κυρίως επικαλυπτόμενων που καταλαμβάνουν ένα μεγάλο μέρος στις εικόνες πηκτωμάτων. Η μέθοδος στηρίζεται στο μοντέλο των ενεργών περιγραμμάτων το οποίο εφαρμόζεται στην αρχική εικόνα πηκτώματος και σε δεύτερο στάδιο, στις περιοχές που εσωκλείονται της αρχικής καμπύλης οι οποίες έχουν μεγάλη πιθανότητα να εμπεριέχουν επικαλυπτόμενες κηλίδες. Η μέθοδος κρίθηκε αξιόπιστη και επιτυχής σε σύγκριση με υπάρχον λογισμικό πακέτο [4] και υπάρχουσες μεθόδους εντοπισμού πρωτεϊνικών κηλίδων [3],[5] και δημοσιεύθηκε σε διεθνή έγκυρα επιστημονικά συνέδρια [6], [7].

Στην Εικόνα 2 απεικονίζονται τα αποτελέσματα ανίχνευσης τα οποία προέρχονται από: (α_1) -(δ_1) το μοντέλο των Chan-Vese, (α_2)-(δ_2) το λογισμικό πακέτο PDQuest 8.0.1 και (α_3)-(δ_3

) την πρωτότυπη τεχνική μας [6]. Είναι εμφανές ότι το μοντέλο των Chan-Vese εντοπίζει τα όρια μίας συστάδας πρωτεϊνών αλλά δεν είναι ικανό να ανιχνεύσει τις επικαλυπτόμενες κηλίδες που περιλαμβάνονται. Επίσης, το λογισμικό πακέτο PDQuest 8.0.1 παράγει ελλειπτικά περιγράμματα τα οποία εσωκλείουν εκτός από εικονοστοιχεία πρωτεΐνης και εικονοστοιχεία υποβάθρου. Ακόμη, οδηγεί σε λανθασμένα αποτελέσματα διότι: 1) δεν ανιχνεύει κάποιες κηλίδες (Εικόνα 2(β

2

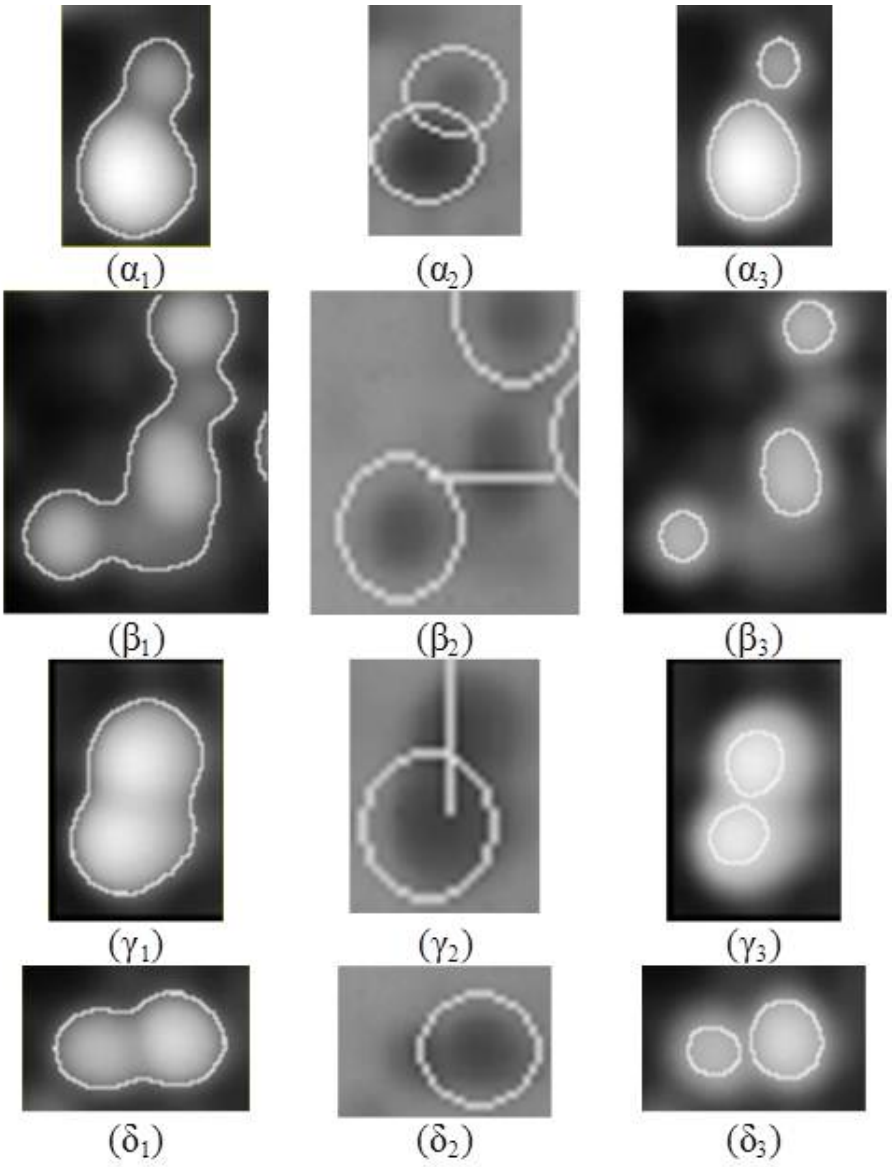
),(γ

2

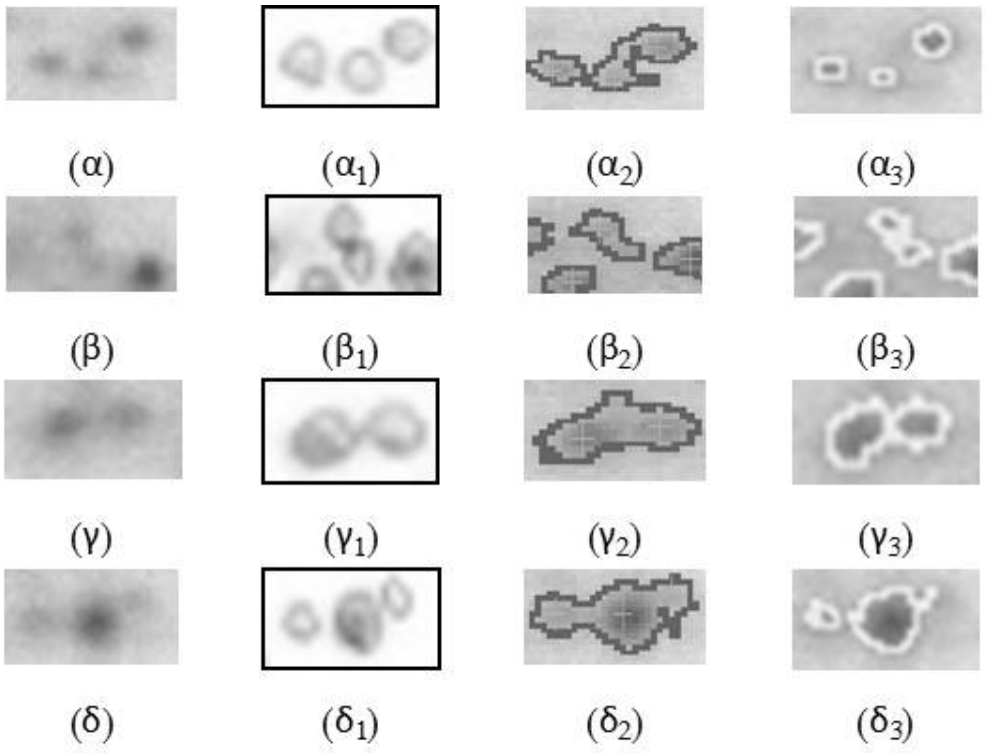
)) και 2) ανιχνεύει λανθασμένα δύο κηλίδες ως μία (Εικόνα 2(δ

2

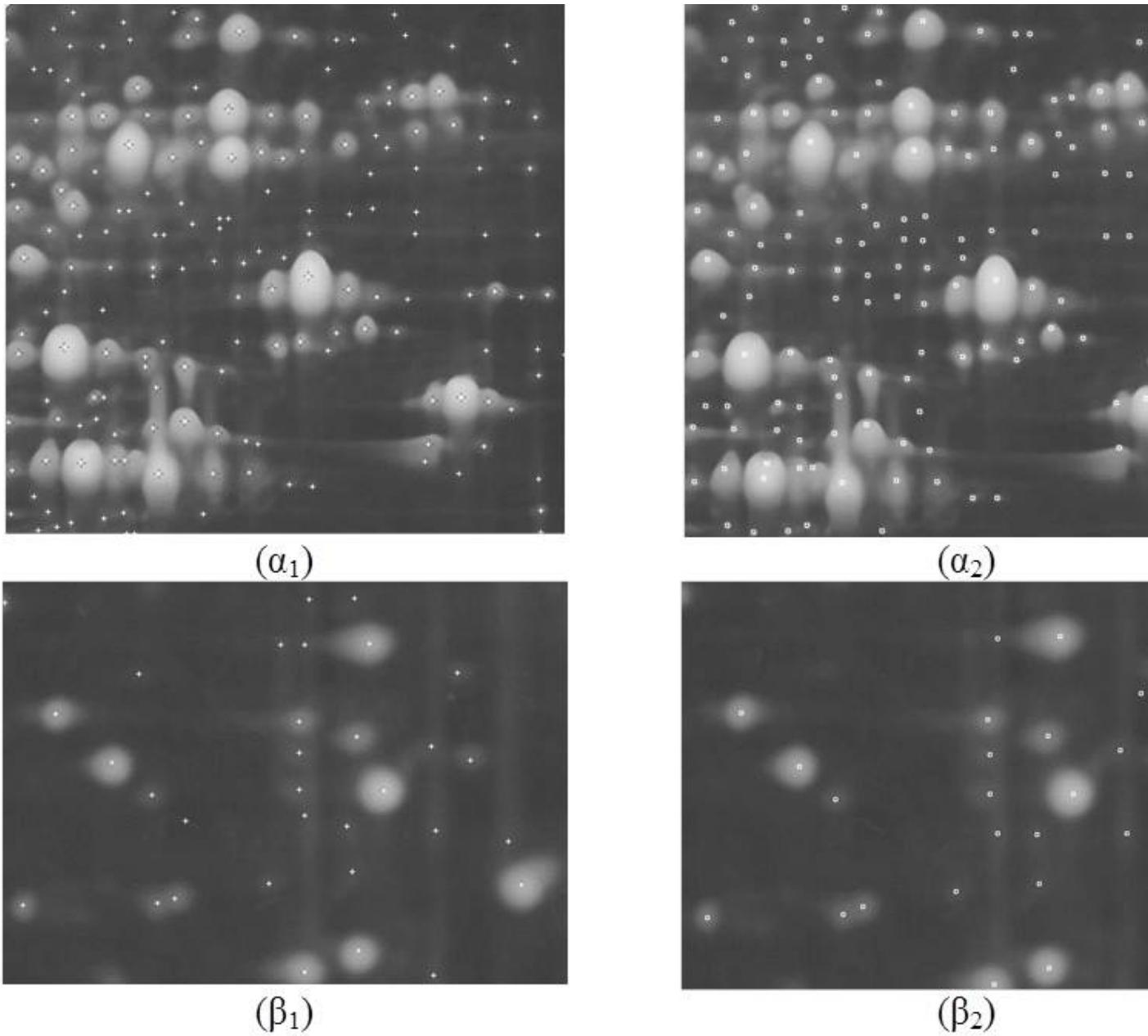
)). Ωστόσο, η πρωτότυπη τεχνική μας ανιχνεύει σωστά τις επικαλυπτόμενες κηλίδες και συνεπώς μπορεί να ενσωματωθεί σε ένα ολοκληρωμένο σύστημα ανίχνευσης εικόνων πρωτεομικής.



Εικόνα 2: Αποτελέσματα επεξεργασίας πυρηνικών κηλίδων τα οποία προέρχονται από: (α1)-



Εικόνα 8: (α)-(δ) Αρχικές εικόνες πηκτωμάτων, (α1)-(δ1) Συνδεδεμένα στοιχεία των εικόνων (α)-(δ),



Εικόνα 4: Αποτελέσματα ανίχνευσης πρωτεϊνικών κηλίδων τα οποία προέρχονται από: (α_1), (α_2), (β_1), (β_2)

Μέθοδοι Κατάτμησης

Κατά την διαδικασία της κατάτμησης η εικόνα χωρίζεται σε μη επικαλυπτόμενα τμήματα, κατηγοριοποιώντας τα εικονοστοιχεία σε αυτά που ανήκουν σε πρωτεϊνικές κηλίδες και σε αυτά που ανήκουν στο ανομοιογενές υπόβαθρο. Οι κηλίδες διαχωρίζονται με περιγράμματα στις ακμές τους έτσι ώστε να είναι εφικτός ο υπολογισμός της ποσότητας της κάθε πρωτεΐνης που περικλείεται στο περίγραμμα. Ιδανικά, κάθε περίγραμμα περικλείει μία και μοναδική πρωτεΐνη. Για την ποσοτικοποίηση της πρωτεΐνης απαιτείται ο υπολογισμός του όγκου της, δηλαδή το άθροισμα των εντάσεων όλων των εικονοστοιχείων που περιλαμβάνονται στο περίγραμμά της. Οι πιο διαδεδομένες μέθοδοι κατάτμησης είναι η μέθοδος κατωφλίωσης [11], ανίχνευσης ακμών [12] και υδροκριτικών γραμμών (watersheds) [13]. Οι μέθοδοι αυτές μελετήθηκαν και αξιολογήθηκαν με βάση τις αρχές λειτουργίας τους

καθώς και τα αποτελέσματά τους.

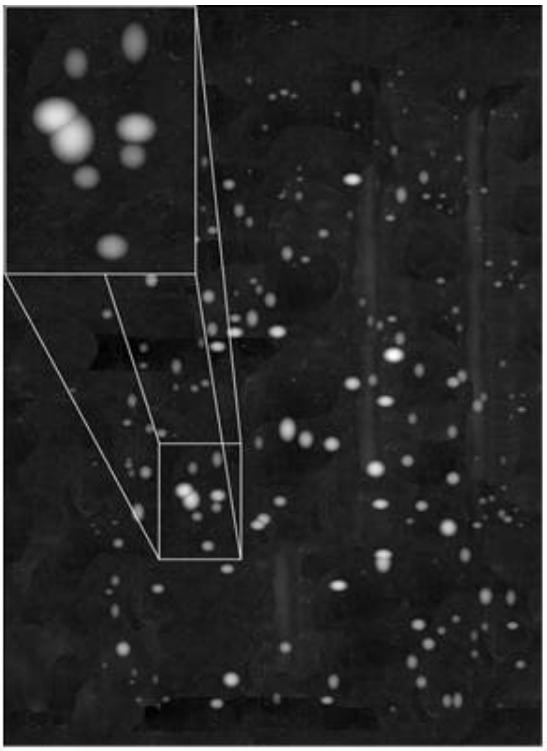
Σύμφωνα με την μέθοδο κατωφλίωσης, αρχικά εφαρμόζεται στην εικόνα ένα κατώφλι το οποίο σχετίζεται με μία πρωτεϊνική κηλίδα. Ο αλγόριθμος επαναλαμβάνεται με αυξανόμενο κατώφλι έτσι ώστε να διαχωριστούν όλες οι συνδεδεμένες περιοχές και τερματίζεται όταν επιπλέον διαχωρισμοί είναι ανέφικτοι. Ιδανικά, το αποτέλεσμα της κατάτμησης θα αποτελείται από συνδεδεμένες περιοχές οι οποίες θα αντιπροσωπεύουν μεμονωμένες κηλίδες. Η μέθοδος Otsu [14],[15] είναι μία από τις πιο διαδεδομένες μεθόδους κατωφλίωσης. Αρχικά, υπολογίζεται το ιστόγραμμα της εικόνας και γίνεται επιλογή του κατωφλίου για τον διαχωρισμό δύο κλάσεων. Ακολουθεί υπολογισμός των μέσων τιμών των κλάσεων καθώς και της διακύμανσης μεταξύ των κλάσεων. Το βέλτιστο κατώφλι είναι εκείνο για το οποίο η διακύμανση μεταξύ των κλάσεων μεγιστοποιείται. Η μέθοδος είναι ιδιαίτερα ευαίσθητη στην παρουσία θορύβου και μη πρωτεϊνικών αντικειμένων καθώς και στην ανομοιογένεια του υποβάθρου.

Η μέθοδος ανίχνευσης ακμών βασίζεται στο γεγονός ότι η μεταβολή της φωτεινότητας στις περιοχές των ακμών είναι πολύ πιο έντονη από ότι στην υπόλοιπη εικόνα. Ιδιαίτερα στις ακμές παρατηρούνται οι υψηλότερες συχνότητες. Αυτό σημαίνει ότι κατά μήκος των ακμών η παράγωγος της φωτεινότητας είναι αυξημένη σημαντικά. Κατά συνέπεια, υπολογίζοντας την παράγωγο σε όλα τα σημεία της εικόνας και κρατώντας εκείνα για τα οποία η παράγωγος παρουσιάζει μεγάλες τιμές, είναι δυνατός ο εντοπισμός της θέσης τους. Εντούτοις, η παρουσία θορύβου καθώς και διάφορες ανωμαλίες στις ακμές δυσχεραίνουν την εκτίμηση της παραγώγου. Για τον λόγο αυτό, συνήθως προηγείται της παραγώγισης ένα φιλτράρισμα του θορύβου.

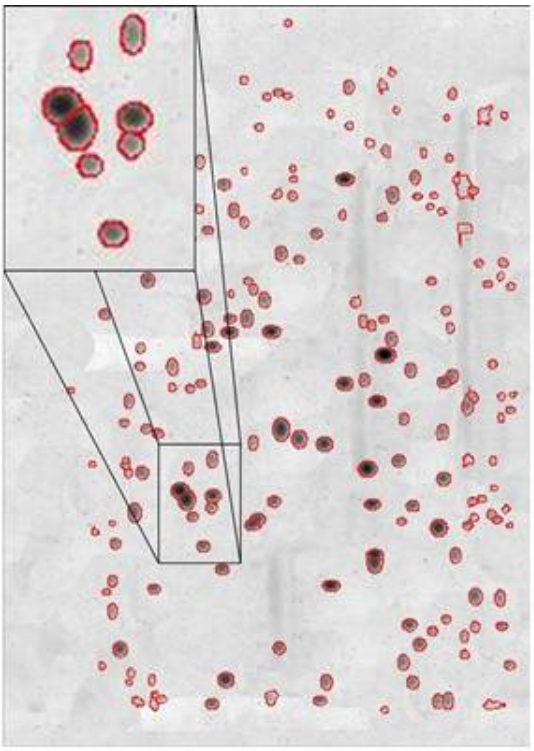
Η μέθοδος των υδροκριτικών γραμμών είναι ίσως η πιο διαδεδομένη μέθοδος στην κατάτμηση εικόνων πηκτωμάτων [16]. Σύμφωνα με την μέθοδο αυτή, η εικόνα θεωρείται ως τοπίο το οποίο περιέχει κοιλάδες και κάθε κοιλάδα αντιπροσωπεύει μία μεμονωμένη κηλίδα. Η εφαρμογή του αλγορίθμου έγκειται στον εντοπισμό των τοπικών ελαχίστων της εικόνας, δηλαδή των κέντρων των κοιλάδων του τοπίου. Εάν θεωρήσουμε ότι στο τοπίο αρχίζει να βρέχει, το νερό θα γεμίσει αργά όλες τις κοιλάδες. Τα φράγματα που τοποθετούνται ανάμεσα στις κοιλάδες, ούτως ώστε να αποφύγουμε την διαρροή, ονομάζονται υδροκριτικές γραμμές. Ο αλγόριθμος των υδροκριτικών γραμμών οδηγεί σε υπερ-κατάτμηση της εικόνας λόγω του εντοπισμού τοπικών ελάχιστων που σχετίζονται με τεχνουργήματα καθώς και με θόρυβο υποβάθρου. Συνεπώς, απαιτείται επεξεργασία με σημαδευτές οι οποίοι αρχικοποιούν τον αλγόριθμο σε συγκεκριμένα σημεία ενδιαφέροντος επάνω στην εικόνα [17]. Ωστόσο, σύμφωνα με την μέθοδο αυτή υπάρχει ανθρώπινη παρέμβαση η οποία οδηγεί σε μη αντικειμενικά και αμφιλεγόμενα αποτελέσματα.

Η μέθοδος των ενεργών περιγραμμάτων [3] παρουσιάζει αρκετά πλεονεκτήματα στην κατάτμηση εικόνων πηκτωμάτων. Στα πλαίσια της διδακτορικής έρευνας αναπτύχθηκε και υλοποιήθηκε μία πρωτότυπη μέθοδος που στοχεύει στην κατάτμηση κηλίδων αχνών και επικαλυπτόμενων ενσωματώνοντας την μέθοδο ενεργών περιγραμμάτων. Αρχικά, η εικόνα σαρώνεται σε διαφορετικές κατευθύνσεις έτσι ώστε να εντοπιστούν τα τοπικά ελάχιστα τα οποία σχετίζονται με όρια επικάλυψης σε περιοχές πρωτεϊνικών κηλίδων. Εφαρμόζονται κριτήρια επιλογής των τοπικών ελαχίστων για αποφυγή εκείνων που αφορούν ελάχιστα υποβάθρου και τα τελευταία σημαδεύονται. Στη συνέχεια, ακολουθεί προσαρμοσμένη εξισορρόπηση ιστογράμματος έτσι ώστε να ενισχυθεί η αντίθεση των αχνών κηλίδων σε σχέση με το υπόβαθρο. Τέλος, εφαρμόζεται το μοντέλο του ενεργού περιγράμματος το οποίο εξελίσσεται σύμφωνα με τις πληροφορίες έντασης της αρχικής εικόνας καθώς και εκείνης που ενσωματώνει τις πληροφορίες των διαχωριζόμενων επικαλυπτόμενων κηλίδων και των ενισχυμένων αχνών κηλίδων. Η μέθοδος κρίθηκε αξιόπιστη και επιτυχής σε σύγκριση με υπάρχοντα λογισμικά πακέτα [4],[9] και δημοσιεύτηκε σε διεθνές έγκυρο επιστημονικό συνέδριο με κριτές [18]. Επίσης, η μέθοδος αυτή αυτοματοποιήθηκε με την αρχικοποίηση του ενεργού περιγράμματος στα τοπικά μέγιστα τα οποία σχετίζονται με πρωτεϊνικές κηλίδες και δημοσιεύτηκε σε διεθνές έγκυρο επιστημονικό συνέδριο με κριτές [19].

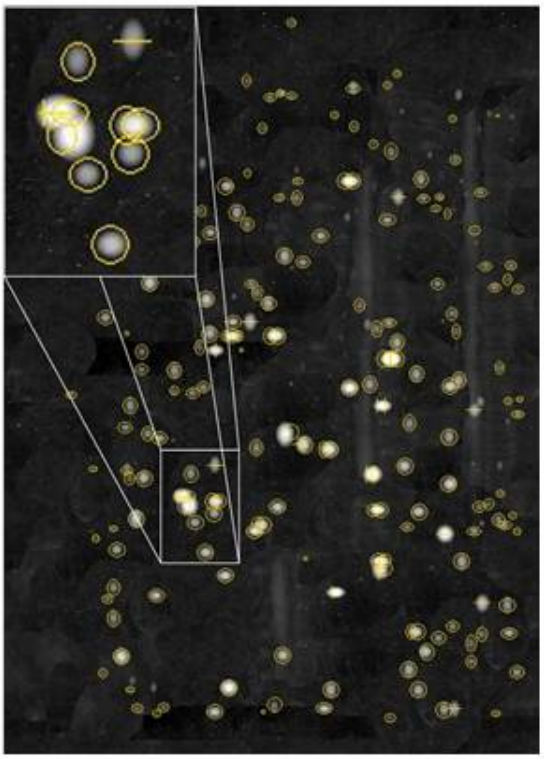
Στην Εικόνα 5(α) απεικονίζεται συνθετική εικόνα πρωτεομικής καθώς και αποτελέσματα κατάτμησης τα οποία προέρχονται από: (β) το λογισμικό πακέτο Melanie 7, (γ) το λογισμικό πακέτο PDQuest 8.0.1 και (δ) την πρωτότυπη μέθοδο που υλοποιήσαμε. Είναι εμφανές ότι το λογισμικό πακέτο Melanie 7 κατατμεί κορυφές του υποβάθρου ως κηλίδες ενώ το λογισμικό πακέτο PDQuest 8.0.1 παράγει ελλειπτικά περιγράμματα τα οποία περικλείουν εικονοστοιχεία υποβάθρου οδηγώντας έτσι σε λανθασμένη ποσοτική ανάλυση του όγκου της κάθε πρωτεΐνης.



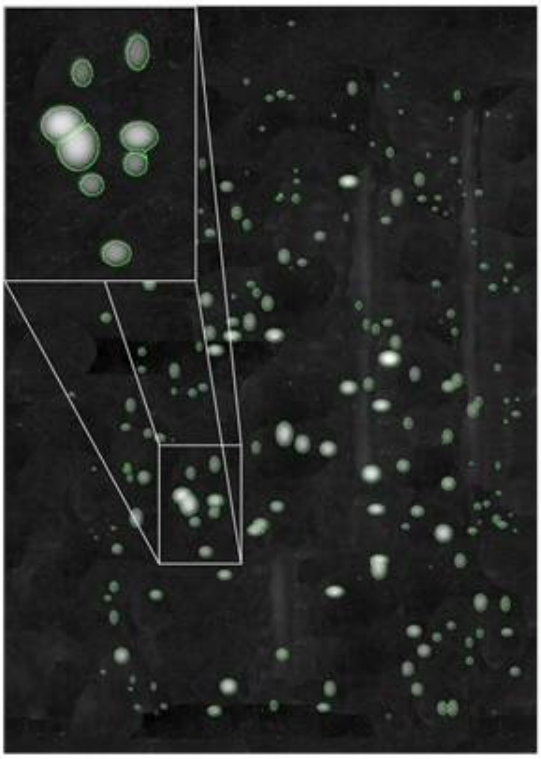
(α)



(β)



(γ)



(δ)

Εικόνα 5: (α) Σημάδι (β) τα μόρια που ανιχνεύθηκαν από το λογισμικό PDQuest

ΑΝΑΦΟΡΕΣ

1.

M. Berth, F.M. Moser, M. Kolbe, J. Bernhardt, “The State of the Art in the Analysis of Two-dimensional Gel Electrophoresis Images,” Appl. Microb. Biotechnol. vol. 76, no. 6, pp. 1223-1243, 2007.

2.

J.S. Morris, B.N. Lark, H.B. Gutstein, “Pinnacle: a Fast, Automatic and Accurate Method for Detecting and Quantifying Protein Spots in 2-dimensional Gel Electrophoresis Data,” Bioinformatics, vol. 4, no. 24, pp. 529-536, 2008.

3.

T.F. Chan, L.A. Vese, “Active Contour without Edges,” IEEE Trans. Im. Proc., vol. 10, no. 2, pp. 226-277, 2001.

4.

http://www.biorad.com

5.

P. Tsakanikas, E.S. Manolakos, "Active contours based segmentation of 2DGE proteomics images," in Proc. Euro-pean Signal Processing Conference (EUSIPCO), Lausanne, Switzerland, 2008.

6.

E. Mylona, M. Savelonas, D. Maroulis, A two-stage active contour-based scheme for spot detection in proteomics images, in Proc. IEEE International Conference on Information Technology Applications in Biomedicine (ITAB), Corfu, Greece, pp. 1-4, 2010.

7.

M. Savelonas, E. Mylona, D. Maroulis, A level set approach for proteomics image analysis, in Proc. European Signal Processing Conference (EUSIPCO), Aalborg, Denmark, pp. 1229-1233, 2010.

8.

P. Cutler, G. Heald, I.R. White, J. Ruan, “A Novel Approach to Spot Detection for Two-dimensional Gel Electrophoresis Images using Pixel Value Collection,” Proteomics, vol. 3, no. 4, p. 392-401, 2003.

9.

R.D. Appel, J.R. Vargas, P.M. Palagi, D. Walther, D.F. Hochstrasser, “Melanie II – A Third Generation Software Package for Analysis of Two-dimensional Electrophoresis Images: II. Algorithms,” Electrophoresis, vol. 8, no. 15, pp. 2735-2748, 1997.

10.

E.A. Mylona, M.A. Savelonas, D. Maroulis, S. Kossida, “A Computer-based Technique for Automated Spot Detection in Proteomics Images,” IEEE Trans. on Inf. Tech. in Biomed., vol. 15, no. 4, pp. 661-667, 2011.

11.

R.C. Gonzalez, R.E. Woods, Digital Image Processing, Pearson International Edition, USA, pp. 728-747, 2008.

12.

N. Khan, S. Rahman, “A New Approach to Detect Similar Proteins from 2D Gel Electrophoresis Images,” IEEE 3rd Symp. on Bioinf. and Bioeng., pp. 182-189, 2003.

13.

Y. Kim, J. Kim, Y. Won, Y. In, “Segmentation of Protein Spots in 2-D Gel Electrophoresis Images with Watershed using Hierarchical Threshold,” Lect. Notes in Comp. Sci., vol. 2869, pp. 389-396, 2003.

14.

M.H.M. Noor, Z. Hussain, K.A. Ahmad, A.R. Ainihayati, “Gel Electrophoresis Image Segmentation with Otsu Method Based on Particle Swarm Optimization,” IEEE 7th Int. Col. on Signal Proc. and its Appl. (CSPA), pp. 426-429, 2011.

15.

E. Zacharia, E. Kostopoulou, D. Maroulis, S. Kossida, “A Spot Segmentation Approach for 2D Gel Electrophoresis Images based on 2D Histograms,” IEEE Int. Conf. on Patt. Rec. (ICPR), Istanbul, Turkey, 2010.

16.

S. Rashwan, T. Faheem, A. Sarhan, A.B. Youssef, “A Fuzzy-Watershed based Algorithm for Protein Spot Detection in 2DGE Images,” Int. J. of Comp. Sci. and Network Secur., vol. 10, no. 5 pp. 254-260, 2010.

17.

C.-li Sun, X.-min Wang, “Spot Segmentation and Verification based on Improve Marker Controlled Watershed Transform,” IEEE 3rd Int. Conf. on Comp. Sci. and Inf. Tech. (ICCSIT), pp. 63-66, 2010.

18.

E.A. Mylona, M.A. Savelonas, D. Maroulis, M. Aivaliotis, “2D-GE Image Segmentation based on Level-Sets,” IEEE Int. Conf. on Im. Proc. (ICIP), Brussels, Belgium, 2011.

19.

M. Savelonas, E. Mylona, D. Maroulis, “An Automatically Initialized Level-Set Approach for the Segmentation of Proteomics Images,” IEEE Int. Workshop on Biomed. Eng., Kos, Greece, 2011.