השערות ומטרות העבודה

השערת המחקר

הדיווחים לעיל מדגימים יכולת טובה של זיהוי חיידקים בספקטרוסקופיית ראמאן ברזולוציה גבוהה וכן היתכנות של זיהוי באמצעות פלואורסנציה. אנו משערים ששימוש בכלים לכריית מידע (Data mining) יכול לאפשר גילוי וזיהוי של חיידקים הן באמצעות ספקטרוסקופיית ראמאן ברזולוציה נמוכה והן בפלואורסנציה, כאשר הכמות הגדולה של הנתונים תאפשר עיבוד מעמיק יותר.

מטרות העבודה

1. בחינת יכולת של ספקטרוסקופיית ראמאן ברזולוציה נמוכה ופלואורסנציה לגילוי וכימות חיידקים במים.
2. בחינת יכולת של ספקטרוסקופיית ראמאן ברזולוציה נמוכה ופלואורסנציה לזיהוי סוגי חיידקים במים.

תכנית המחקר

חלק I – ספקטרוסקופיית ראמאן

שלב 1 – אופטימיזציה

במהלך העבודה יבוצע תהליך אופטימיזציה למכשיר ספקטרוסקופיית הראמאן. שלב זה נדרש למרות שבעבודות קודמות הצליחו לזהות חיידקים בטכניקה דומה משום שהציוד ששימש בעבודה ב-2005 [30] אינו תקין, הוחלף ושודרג. כיוון שחלק מהמכשיר עצמו נבנה במכון להנדסה חקלאית, יש לנסות ולהגיע לשילוב הנכון לזיהוי חיידקים מבחינת אורך הגל של הלייזר, גובה המשדר, זמן החשיפה והקריאה, גודל הדוגמה, צורת התושבת (המשפיעה על צורתה הפיזית של הדוגמה), חומר התושבת (המשפיע על החזר הלייזר), שיטת הדגימה – בטבילה, דרך זכוכית או מהאוויר, עוצמת הלייזר, כמות החזרות להפחתת ה"רעש" ואופן הכנת הדוגמה לבדיקה. לשם כך נשתמש במשדר לייזר 785 nm וגלאי מסוג QE65, שניהם של חברת OceanOptics. במהלך האופטימיזציה ניצור מאגרי נתונים שעליהם תבוצע אנליזה סטטיסטית, מבוססת על מודל מתמטי PLS באמצעות תוכנות Matlab (MATLAB 2015a, Systematics, Natick, MA, ארה"ב) ו-JMP (JMP Pro 13, SAS Institute Inc., Cary, NC, ארה"ב). במידת הצורך יבחנו גם שיטות כימומטריות נוספות כגון רשתות עצביות (Neural Networks) או Support Vector Machine. לאורך העבודה הבדיקות תעשנה על חיידקי *E. coli* (זן DH5α). לאורך כל שלבי העבודה נשתמש בשיטות של Machine Learning כדי לאמן (Training) את מודל החיזוי שלנו על נתונים קיימים, לבצע תיקוף (Validation) של יכולת החיזוי ואף לבדוק את יכולת החיזוי בתנאי אמת (Prediction). זאת נעשה בעזרת שימוש מגוון במודלים מבוססי PLS, בצירוף עיבוד-מקדים (Preprocessing) מתמטי כמו גזירה (Derivitive).

נאסוף בין 200-50 סריקות של חיידקים במים מזוקקים על מנת לבנות מודל אמין ואיכותי לזיהוי סגולי של החיידקים במים. ככל הנראה שלב האופטימיזציה הוא הארוך ביותר ועלול לקחת מעל לשנת עבודה עד להגעה לתנאים המתאימים לזיהוי החיידקים. שיפור הפלטפורמה אמור להיות יחסית פשוט.

שלב 2 - בחינת סף הרגישות של המכשיר לחיידקי *E. coli*

לאחר שתימצא הפלטפורמה המתאימה (כלומר, מצב מסוים של כלל הפרמטרים הנ"ל), תיערכנה בדיקות לבחינת יכולת המכשיר לזהות חיידקים בריכוזים נמוכים עד להגעה למינימום רגישות המכשיר. באופן זה נוכל לפתח מודל לכימות של חיידקים מזן יחיד על רקע מים מזוקקים. בחינה זו תעשה על ידי מיהול וסריקת דוגמאות בריכוזים ידועים וניתוח התוצאות בעזרת מודל מבוסס PLS.

שלב 3 - בחינת סף הרגישות של המכשיר לחיידקי *Bacillus subtilis*

בהמשך לזיהוי חיידק המודל (*E. coli*) ננסה להבחין בזן אחר של חיידקים – *B. subtilis* (זן NCIB 3610). חיידק זה נבחר היות והוא חיידק מסוג גראם חיובי, השונה באופן מהותי מבחינת הרכב החלבונים, השומנים והדופן שלו. תיערכנה בדיקות לבחינת יכולת המכשיר לזהות חיידקים בריכוזים נמוכים עד להגעה למינימום רגישות המכשיר. באופן זה נוכל לפתח מודל לכימות של חיידקים אלה על רקע מים מזוקקים. בחינה זו תעשה על ידי מיהול וסריקת דוגמאות בריכוזים ידועים וניתוח התוצאות בעזרת מודל מבוסס PLS.

שלב 4 – בחינת יכולת אבחנה בין חיידקי *E. coli* ו-*B. subtilis*

במידה ויימצא שקיים הבדל בין טביעת האצבע הספקטראלית של *E. coli* לזו של *B. subtilis*, אנו ננסה לפתח מודל סטטיסטי לזיהוי של החיידקים השונים. לשם כך נשתמש בדוגמאות מעורבות בריכוזים שונים של חיידקים על רקע מים מזוקקים, ניצור מאגר נתונים רחב וננסה לפתח את המודל.

במידה ולא נמצא הבדל בין טביעת האצבע הספקטראלית של חיידקים אלה, נבצע שיפור של המודל הסטטיסטי לכימות כלל החיידקים בדוגמה, ונייצר למעשה מודל לספירה כללית של חיידקים במים באמצעות ספקטרוסקופיית ראמאן ברזולוציה נמוכה.

חלק II – ספקטרוסקופיית פלואורסנציה

עבודות קודמות במעבדה הראו כי ניתן לקשר בין ריכוז חיידקים במי שתייה ועוצמת פלורסנציה. עבודות אלה הראו יכולת לזהות חיידקים ידועים בריכוזים של עד 10^2 Cells ml^{-1}, והחוקרים קישרו בין התופעה לנוכחות נגזרות של טריפטופן. אנו ננסה להמשיך עבודות אלה כדי להראות שימוש בספקטרוסקופיית פלורסנציה ללזיהוי חיידקים בדוגמאות סביבתיות מקידוחי מים בצפון הארץ, ובנוסף ננסה להשוות בין יכולת הזיהוי באמצעות פלורסנציה למכשיר הראמאן שלנו.

הקידוחים נבחרו מכיוון שבעבר נצפו בהם זיהומים מיקרוביאליים ספורדיים. אנו ננסה לזהות זיהומים אלה באמצעות ספקטרוסקופיית פלואורסנציה מבוססת עירור-הארה ולקשר בין תגובות הארה לנוכחות מיקרואורגניזמים במים. את הדוגמאות נבדוק באמצעות מכשיר ספקטרוסקופיית פלואורסנציה (מדגם RF-5301PC של חברת Shimadzu, קיוטו, יפן). לאחר סריקת הדוגמאות ואיסוף נתוני EEMs נבצע אנליזה סטטיסטית ראשונית באזורים הידועים מעבודות קודמות (עירור-הארה 210/370, 225/336 ו-280/332-336). בהמשך ייתכן ונבצע אנליזה מבוססת על המודל המתמטי PARAFAC (Parallel Factor Analysis) או PLS. מודלים אלו מאפשרים ניתוח רב-גורמי על מנת לאתר את "אזורי העניין" במפות ה-EEM. נשתמש בתוכנת Matlab (MATLAB 2015a, Systematics, Natick, MA, ארה"ב) על מנת לבצע את האנליזה. נבחן את אזורי העירור-הארה שבין 210-400 nm.

בנוסף לכך, נבצע מדידות בריכוזי חיידקי *E. coli* ו-*B. subtilis* ידועים על מנת לכייל את המערכת ולהכין אוסף סריקות כדי שנוכל להשוות בין פלטפורמת הראמאן לפלורסנצייה.

<u>שלב 1 – איסוף נתוני פלורסנציה מדגימות קידוחי מים</u>

לאורך שנה אחת (החל באוגוסט 2017 ועד אוגוסט 2018) נסרוק דוגמאות מים מקידוחים שונים בצפון הארץ על מנת לייצר מפות EEM בטווח עירור-הארה שבין 210-400nm. כל דוגמה תבחן לפני ואחרי הכלרת המים ולפני ואחרי סינון המים בפילטר 0.45μm על מנת לבדוק את השפעתם של מיקרואורגניזמים על מטריצת העירור-הארה. במקביל, ייאספו נתוני מיקרוביולוגיה (ספירה כללית, נוכחות קוליפורמים, קוליפורמים צואתיים, וסטרפטווקים צואתיים) לפי השיטות הסטנדרטיות (ref). הנתונים ייאספו לניתוח בהמשך העבודה.

<u>שלב 2 – ניתוח ראשוני של הנתונים באורכי גל ידועים</u>

בהמשך לעבודתו של Simelane, נבחן התאמה בין נתוני ספירה כללית (או נתוני מיקרוביולוגיה אחרים, אם יהיו רלוונטיים) במים שנתקבלו מבדיקות סטנדרטיות לעוצמות הארה באורכי גל עירור/הארה: 210/370, 225/336 ו-280/332-336. שלב זה יבוצע כאשר מאגר הנתונים יהיה בעל מעל 50 דוגמאות שעליהן ניתן לבצע ניתוח. איסוף הנתונים יימשך במקביל.

<u>שלב 3 – ניתוח מתקדם של הנתונים</u>

בהעדר קורלציה חזקה בין נתוני הספירה הכללית (או נתוני מיקרוביולוגיה אחרים, אם יהיו רלוונטיים) לעוצמות הארה באורכי הגל הידועים, נשתמש בשיטות מתקדמות על מנת לייצר יכולת חיזוי טובה יותר ממטריצות הפלורסנציה המלאות (EEMs). נשתמש במודלים מבוססי PLS, כולל מניפולציות מתמטיות כגון גזירה, Centering, נרמול התוצאות ועוד כנדרש. אם יהיה צורך נוכל גם להשתמש במודלים מבוססי אנליזת PARAFAC, רשתות עצביות (Neural Network) או SVM. שלב זה יבוצע כאשר מאגר הנתונים יהיה בעל מעל 50 דוגמאות שעליהן ניתן לבצע ניתוח. איסוף הנתונים יימשך במקביל.

<u>שלב 4 – בחינת ספי רגישות לחיידקי *E. coli* ו-*B. subtilis*</u>

במקביל לעבודה עם דוגמאות המים מהקידוחים, על מנת שנוכל להשוות בין העבודה בראמאן ובפלורסנציה, ניצור מאגר נתונים של EEMs של חיידקי *E. coli* ו-*B. subtilis* בריכוזים ידועים בטווח שבין $1\text{-}10^8$ CFUs ml^{-1} במים מזוקקים. את ריכוז החיידקים ננסה לחזות בשיטות שתוארו בשלבים 2 ו-3.

<u>חלק III – השוואה בין ספקטרוסקופיות ראמאן ברזולוציה נמוכה לפלואורסנציה לזיהוי וכימות חיידקים במים</u>

באמצעות מודלים סטטיסטיים ידועים המשווים דיוק (Accuracy), חוסן (Robustness) והדירות (Repeatability) נשווה בין יכולת החיזוי של המודלים שנפתח לספקטרוסקופיית ראמאן ולפלואורסנציה. נשתמש במודל ההשוואה של Ignat [48] המתבסס על ניתוח root mean square error of calibration (RMSEC) ו-root mean square error of cross validation (RMSECV) כדי לייצר ערך Standerdized Weighted Sum (SWS) שמהווה מדד לאיכות החיזוי של המודלים. מערך זה נוכל להסיק מי מהשיטות מתאימה יותר לזיהוי וכימות חיידקים במים.

את כל עבודת המעבדה, כולל גידול החיידקים, הכנת הדוגמה ומדידה בפועל יבצע התלמיד (אמיר נקר) בשיתוף פעולה מלא עם החוקרים מהמכון להנדסה חקלאית (ד"ר זאב שמילוביץ' וד"ר תימאה איגנת) וחוקרים מהמעבדה של פרופ' שלמה סלע מהמכון לאיכות ובטיחות מזון של מרכז המחקר החקלאי. התלמיד יהיה שותף בתכנון ובניית פלטפורמת הראמאן (יחד עם עובדי המכון להנדסה חקלאית). בעבודה על ספקטרוסקופיית פלואורסנציה התלמיד יבצע את כל הסריקות והאנליזות בשיתוף פעולה עם חוקרים מהמעבדה של ד"ר מיכאל בוריסובר מהמכון למדעי הקרקע והמים של מרכז המחקר החקלאי. התלמיד יבצע את כל האנליזות הסטטיסטיות, פיתוח, שיפור והשוואת המודלים הסטטיסטיים לאחר הסריקות.

לוח הזמנים

תאריך התחלה משוער	תאריך סיום משוער	פעולה
11/2016	10/2017	חלק I, שלב 1 - אופטימיזציה של הפלטפורמה ושיטת הכנת הדוגמה לזיהוי חיידקי *E. coli* בעזרת מכשיר ראמאן.
8/2017	8/2018	חלק II, שלב 1 – איסוף נתוני פלורסנציה מדגימות קידוחי מים
10/2017	12/2017	חלק I, שלב 2 - בחינת סף הרגישות של פלטפורמת הראמאן לחיידקי *E. coli*
10/2017	12/2017	חלק I, שלב 3 - בחינת סף הרגישות של פלטפורמת הראמאן לחיידקי *B. subtilis*
12/2017	4/2018	חלק II, שלב 2,3 – ניתוח ראשוני ומתקדם של נתוני הפלורסנציה שנאספו עד כה
1/2018	4/2018	חלק I, שלב 4 - בחינת יכולת אבחנה סגולית בין חיידקי *E. coli* ו- *B. subtilis* בפלטפורמת ראמאן
1/2018	4/2018	חלק I, שלב 4 - בחינת יכולת אבחנה סגולית בין חיידקי *E. coli* ו- *B. subtilis* בספקטרוסקופיית פלואורסנציה
1/2018	4/2018	חלק II, שלב 4 - בחינת ספי רגישות לחיידקי *E. coli* ו-*B. subtilis*
4/2018	8/2018	חלק III – השוואה בין ספקטרוסקופיות ראמאן ברזולוציה נמוכה לפלואורסנציה לזיהוי וכימות חיידקים במים
8/2018	10/2018	כתיבת דוח העבודה והגשתו.