

VisualiseBGU

Video Analysis and Insights



אוראל קקון
אור סעדה
נתנאל קלונטר

תוכן עניינים

1. מבוא
2. תרחישי שימוש
3. דרישות פונקציונליות
4. דרישות לא פונקציונליות
5. ניהול סיכונים והערכת הרעיון
6. נספח 1 - דוח Face Recognition

מבוא

הפרויקט עוסק בקידום בניית יכולת אוטומטית של מחשב להבין את תוכנו של סרט קולנוע, כדוגמת זיהוי הדמויות הראשיות, טיב הקשרים ביניהן, מה מניע אותן, וכו'. אנו נבנה ספרייה של פונקציות בסיסיות, כדוגמת זיהוי הדמויות במהלך הסרט) למרות שהן לעתים עם הגב למצלמה (זיהוי מעשיהן של הדמויות, הרגשות שלהן וכו'. לצורך כך יהיה צורך להשתמש בספריות ניסיוניות קיימות אך פרימיטיביות של עיבוד וידאו, ניתוח וידאו, טקסט ואודיו. איכות היכולות שנבנה תיבדק בהשוואה למאגר של 45 סרטים שתויגו ידנית.

1.1 תחום הבעיה

ניתוח תוכן היא שיטת מחקר לניתוח של טקסטים במגוון פורמטים (טקסט כתוב, קול, תמונה, וידאו, דיגיטלי). שיטת ניתוח התוכן התפתחה במסגרת מדעי החברה האמפיריים והיא מגשרת בין הניתוח הכמותי הסטטיסטי לבין הבחינה האיכותנית של הפריטים הנבחרים. יש להבחין בין ניתוח תוכן לקיטלוג של טקסט, כאשר מטרתו של ניתוח התוכן היא להגיע למסקנות על-אודות דפוסים, קשרים או הבדלים בטקסטים הנבחרים. קיימות ספריות שונות לניתוח וידאו כיום, אך לא קיימת ספרייה שמצליחה להביא מידע רב באחוזי דיוק מספיק גבוהים.

1.2 הקשר

עושים שימוש כיום בניתוח וידיאו למטרות משפטיות, אבטחה ועוד. ע"י שיפור יכולות ניתוח הוידאו והסקת תובנות נוכל לקבל מידע איכותני אשר יאפשר להגיע למסקנות על-אודות דפוסים, קשרים או הבדלים.

1.3 מטרה

משתמשים יוכלו לטעון סרטי וידאו ולקבל ניתוח מלא על הדמויות המופיעות (מי מופיע, מתי הופיע), ניתוח רגשות של הדמויות אשר הובעו במהלך הסצנות וניתוח תובנות על הסיטואציה אשר קורית בכל סצנה.

1.4 בעלי העניין

המשתמשים הפוטנציאליים הם עורכי וידאו, מפרסמים ועוד המעוניינים לקבל תובנות על הידע. בנוסף, הפרויקט ישמש למחקר לגבי התובנות שניתן לחלץ מוידאו ואיכותן.

1.5 הקשר לתוכנה

כריית מידע או כריית נתונים היא הפעלת אלגוריתם או תוכנית מחשב לצורך גילוי מידע הטמון בבסיסי נתונים קיימים, והסקת מסקנות מהצלבתו. גילוי ידע בבסיסי נתונים הוא תהליך שנועד לחקור ולנתח כמות גדולה של מידע באמצעים אוטומטיים ככל שניתן כדי לגלות דפוסים תקפים חדשים, שימושיים (מקבלי החלטות יכולים להשתמש בדפוסים שהתגלו כדי לשפר תהליכים קיימים) ובעלי משמעות (מובנים למקבל החלטות). עם ההתפתחות המתמדת בעולם טכנולוגיית המידע, כשכמות המידע הדיגיטלי גדלה כל הזמן, עולה החשיבות של תחום כריית המידע, שבאמצעותו ניתן להפוך מידע לידע.

התוכנה תקבל כקלט וידיאו מהמשתמש ותחזיר תובנות אודות הוידאו לפי סדר המאורעות עפ"י ציר הזמן בסרט, אשר יכללו:

- רגשות
- סנטימנט
- זיהוי הדמויות
- רצף הזמן

התובנות יוצגו למשתמש אשר טען את הסרט.

תרחישי שימוש

- **תיאור: אורח נרשם למערכת**

שחקן: אורח
תנאי קדם : ללא
תרחיש מוצלח : האורח נרשם בהצלחה למערכת לאחר הזנת הפרטים להרשמה.
המערכת תציג הודעת אישור כי ההרשמה בוצעה בהצלחה, ולאחריה אוטומטית יתבצע log in למערכת.
תרחיש לא מוצלח : יש שגיאת הרשמה כתוצאה מהזנת פרטים שגויה כמו -
כתובת אימייל לא חוקית ,
סיסמה המכילה תווים שהם לא מספרים או אותיות באנגלית.
המערכת תציג הודעת שגיאה מפורטת למשתמש.

- **תיאור: אורח מתחבר למערכת**

שחקן: אורח
תנאי קדם : משתמש רשום אך אינו מחובר
תרחיש מוצלח : האורח מתחבר בהצלחה למערכת לאחר הזנת הפרטים להתחברות.
המערכת תציג הודעת אישור כי החיבור התבצע בהצלחה, ולאחריה אוטומטית יתבצע log in למערכת.
תרחיש לא מוצלח : יש שגיאת הרשמה כתוצאה מהזנת פרטים שגויה כמו -
כתובת אימייל לא חוקית , הפרטים אינם תואמים למאגר החשבונות
אשר שמור במערכת , המערכת תציג הודעת שגיאה מפורטת למשתמש.

- **תיאור: משתמש טוען סרט במערכת על מנת לקבל תובנות**

שחקן: משתמש
תנאי קדם : משתמש רשום מחובר למערכת.
תרחיש מוצלח : הסרט נטען בהצלחה למערכת , המערכת תציג למשתמש פידבק שהסרט נטען בהצלחה.
תרחיש לא מוצלח: טעינת הסרט לא תוצלח מהסיבות הבאות -
- הסרט נמחק במהלך טעינתו במערכת - המערכת תשלח הודעת שגיאה
אינפורמטיבית על כישלון בטעינת הסרט.
- ביצוע log out של המשתמש במהלך טעינת הסרט.

- **תיאור: משתמש מתנתק מהמערכת**

שחקן: משתמש
תנאי קדם: המשתמש מחובר למערכת.
תרחיש מוצלח: המשתמש התנתק מהמערכת ולא ניתן לצפות יותר בתובנות ובסרטים שהוא העלאה.
תרחיש לא מוצלח: התנתקות מהמערכת לא הצליחה

- **תיאור: משתמש מייצא תובנות לאחר טעינת הסרט לקובץ**
שחקן: משתמש
תנאי קדם: המשתמש מחובר למערכת לאחר שטען סרט תקין.
תרחיש מוצלח: המשתמש מייצא את התובנות אשר המערכת הניבה עבור הסרט שהמשתמש טען לקובץ JSON/txt - המשתמש יקבל עדכון על הצלחת הפעולה.
תרחיש לא מוצלח: המערכת לא הצליחה לייצא את התובנות לקובץ חיצוני - המשתמש יקבל עדכון על אי ביצוע הפעולה.
- **תיאור: סינון תובנות לאחר ניתוח הוידאו**
שחקן: משתמש
תנאי קדם: המשתמש מחובר למערכת ובנוסף טען סרט תקין.
תרחיש מוצלח: משתמש בוחר בסוג הסינון המבוקש מבין האפשרויות
1. רגשות 2. זיהוי דמויות 3. זיהוי סיטואציה
ומקבל בחזרה רק את המידע המפולטר לפי סוג הסינון שביקש.
תרחיש לא מוצלח: המשתמש מנסה לפלטר את התובנות ולא מצליח המשתמש יקבל הודעה מתאימה מדוע הפעולה נכשלה.
- **תיאור: הפעלה על ידי קובץ קונפיגורציה (ללא GUI)**
שחקן: ---
תנאי קדם: ---
תרחיש מוצלח: המערכת לוקחת את המידע הדרוש (נתיב לסרט) לניתוח הסרט מקובץ הקונפיגורציה ומנתחת את הסרט, המערכת תעדכן שהצליחה.
תרחיש לא מוצלח: המערכת נתקלת בבעיה ולכן לא מצליחה לנתח את הסרט על ידי קובץ הקונפיגורציה, המערכת תציג הודעה מתאימה מדוע הפעולה נכשלה.

דרישות פונקציונליות

- **אורח:**
 1. אורח יכול להיכנס לדף רישום למערכת
 2. רישום למערכת יתבצע על ידי אימייל וסיסמה
 3. אורח יכול להיכנס לדף התחברות למערכת ולהזדהות באמצעות אימייל וסיסמה ששמורים במערכת
 4. אורח יכול לראות פרטים בסיסיים אודות המערכת והמפתחים
- **משתמש רשום:**
 1. משתמש יכול לטעון סרט וידאו על מנת לקבל ניתוח דמויות והסקת תובנות מהסרט.
 2. משתמש יכול לסנן את המידע המתקבל לאחר הניתוח.
סוגי הסינון האפשריים:
-רגשות
-זיהוי דמויות
-זיהוי סיטואציות

- מערכת:

1. המערכת תוכל להתממשק עם ה-API של Azure Media Services Video Indexer על מנת לשלוף את המידע לשימוש שיפור ניתוח הסרט.
2. המערכת תוכל לשייך לכל דמות נתונים כמו מה היא עושה או מה מרגישה
3. המערכת תוכל לזהות דמויות המופיעות בסרט, ולהציג את הזמנים בסרט שבה הדמות מופיעה. הזיהוי של הדמות יתבצע גם על מקרים מורכבים: הופעת הדמות עם הגב למצלמה, זוויות פנים שונות (פרופיל, פנים מסתכלים כלפי מטה / מעלה)
4. המערכת תוכל לזהות רגשות של הדמויות המופיעות בסרט, ולהציג עבור כל דמות את הרגשות שהביעה במהלך הסרט.
5. המערכת תציג על גבי ציר זמן כרונולוגי את התובנות שהופקו, בהתאם לציר הזמן הכרונולוגי של הסרט.

דרישות לא פונקציונליות

- אמינות ויציבות-

1. המערכת תתעד את כל השגיאות ופעולות בקובץ log שישמר.
2. מצב המערכת ישמר על בסיס הנתונים.
3. בעת התאוששות המערכת תשחזר את מצבה ע"י משיכת המידע מבסיס הנתונים.
4. המערכת תביא תוצאות מהימנות עפ"י מבדקי איכות מול מאגר הנתונים המכיל 45 סרטים שתוייגו ידנית.

- ביטחון ובטיחות-

1. המערכת תשמור את הסיסמאות בצורה מוצפנת.
2. במערכת יהיה שני רמות של הרשאות: אורח וחבר.
3. המערכת תציג לכל סוג משתמש רק את הפעולות המתאימות לרמת ההרשאות שלו

- ניידות-

1. המערכת תפעל על מערכות נפוצות ב-web או ב-desktop
2. אפליקציית השרת צריכה להתאים לריצה על גבי מערכת הפעלה windows 10

- שימושיות-

1. מערכת תאפשר שליטה בפעולות ללא כל צורך בהדרכה מוקדמת, בהנחה שיש למשתמש לפחות ניסיון מועט בשימוש באפליקציות רשת.
2. שפת המערכת תהיה אנגלית
3. המערכת תאפשר שימוש בה רק למשתמשים רשומים שיתחברו.
4. המערכת תנפק שגיאה אינפורמטיבית במקרה בו פעולה כלשהי נכשלה ותאפשר תיקון במידת האפשר.
5. המערכת תהיה עם ממשק ידידותי ונוח לשימוש.

- זמינות -

1. המערכת תהיה זמינה לשימוש בכל עת, פרט לנפילות יזומות
2. בעת ניתוק ונפילות המערכת תשלח התראה למשתמשים הרלוונטים, המידע האישי ישמר ויהיה ניתן לפנות שוב אל השרת לאחר תיקון התקלה.

● ביצועים -

1. הגעה לתוצאה כמה שיותר קרובה ל-100%.
2. הגעה לתוצאה טובה יותר מרמת הדיוק של התוכנה הקיימת Microsoft Video Indexer תיחשב כהצלחת הפרויקט.
3. על המערכת לבצע את ניתוח הוידאו לא יאחר יותר מאורך הסרט ועוד שעייתם.
4. על המערכת לנתח סרטים שאורכם אינו עולה על 3 שעות

ניהול סיכונים והערכת הרעיון

סיכונים:

החוקרים:

- מאחר והפרוייקט קשור קשר ישיר למחקר עולים סיכונים שיש לקחת בחשבון:
- החוקרים עשויים לשנות את הדרישות עקב נתונים שעולים משימוש ראשוני במוצר.
- מה יקרה אם באמצע הפרויקט המחקר, יחשף מחקר אחר אשר ישנה את תפיסת החוקרים לגבי תחום המחקר?
- מה אם החוקרים יפסיקו את המחקר מסיבות אחרות?

היתכנות:

- מאחר ואין כיום בשוק כלי שעובד בצורה מספקת ומלאה מכיוון שהבעיה לא פשוטה, צריך לקחת בחשבון שאולי הטכנולוגיה והידע שברשותנו כיום לא מספיק על מנת להצליח לשפר את הדיוק בזיהוי הפנים (תחת הפרעות צילום), זיהוי רגשות והבנת סיטואציות.
- ה- data set המתויג יכול להכיל תיוגים שאינם נכונים.
- הקוד אליו מוזנים הסרטים המתויגים ידנית אינו יהיה זמין יותר בשל עדכונים של ספריות חיצוניות כגון: networkx, pip ועוד.
- ללמוד ספריה של deep learning עלול לקחת זמן רב ואפילו להיכשל עקב רמת קושי גבוהה

תוכנית לאיטרציה ראשונה

- על מנת לחסוך סיכון גבוה להמשך הפרויקט, המשימות הנדרשות באיטרציה הראשונה הם:
- מימוש מערכת בסיסית אשר תציג מס' תובנות ללא שיפור הביצועים מהמערכת הקיימת microsoft video indexer.
 - הממשק משתמש יכיל מסך עם אפשרות להעלאת הסרטון ומסך או קובץ קונפיגורציה עם אפשרות לפלטור התובנות ע"י המשתמש.
 - אפשרות להשוואה של תובנה אחת לפחות בין התובנות מהמערכת microsoft indexer למאגר הסרטים שתויג ידנית, עם נתוני השוואה כמו precision ו- recall.
 - ביצוע אלגוריתם בסיסי לשיפור אחת מהתובנות על מנת לצבור ידע על השיפורים האפשריים.
 - למידה של ספריות deep learning.

נספחים

דו"ח אודות face recognition:

https://docs.google.com/document/d/1azk_iecsxQO2tDlkpnn2HHghs6QpKYHIWSHGWCV-Gag/edit?usp=sharing