

ראייה עמוקה (Deep Vision) הופכת למציאות עם רשתות CNN

רשתות נוירונים מבצעות משימות עיבוד תמונה מורכבות

Martin Cassel, Editor, Silicon Software GmbH <

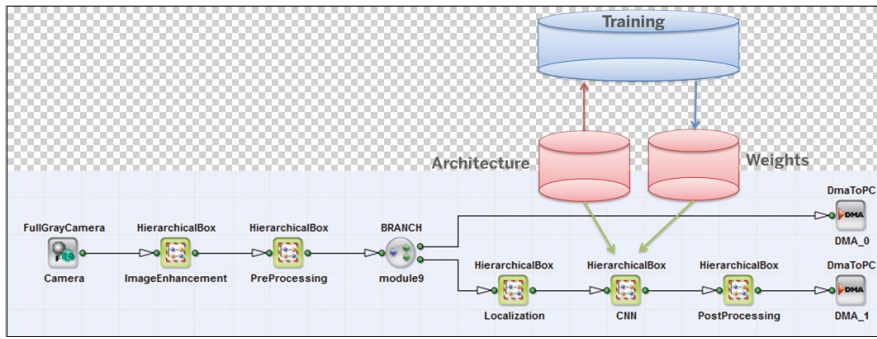
רשתות CNN ממוטבות מאיצות ראייה ממוחשבת

ניסיון מעשי עם רשתות CNN בשנים האחרונות הוביל להנחות מתמטיות ולהפשטות. ניתן למנות כמה, לדוגמה - אגירה ReLu, pooling והימנעות מהתאמת יתר overfitting. הדבר הביא לירידה בהוצאות המחשוב, מה שאיפשר יישום של רשתות עמוקות יותר. צמצום עומק התמונה באותו קצב של זיהוי ואופטימיזציה של האלגוריתם גורמים לכך שרשתות CNN יכולות להיות מואצות משמעותית וכעת הן אידאליות לעיבוד תמונה. רשתות CNN הן בלתי תלויות בהזזת התמונה ובאופן חלקי גם בהגדלתה ובכך מאפשרות את השימוש באותם מבני רשת עבור רזולוציות תמונה שונות. רשתות נוירונים קטנות יותר לעיתים קרובות מספיקות למשימות עיבוד תמונה רבות. בשל הרמה הגבוהה של מקביליות העיבוד, רשתות נוירונים מתאימות במיוחד ל-FPGA (Field Programmable Gate Array), שעליו רשתות CNN יכולות גם

למידה עמוקה בעזרת רשתות נוירונים תהיה בעלת השפעה משמעותית על העתיד של נושא עיבוד תמונות, שכן לגישה זו מספר יתרונות משמעותיים לתוצאות הסיווג והניתוח, וכן לאיכות התמונה הסופית. מכיוון שרשתות נוירונים קטנות מספיקות ליישומי ראייה אופייניים רבים, ניתן ליישם בעילות רשתות נוירונים מלאכותיות (Convolutional Neural Networks) על מעבדים דוגמת FPGA. כתוצאה מכך, קיימות אפשרויות יישום רבות, הרבה מעבר למשימות סיווג, תוך שימוש יעיל בתוך מערכות ראייה משובצות (embedded).

מתמטיים. זה מאפשר לכידה וניתוח של התמונה בתנאים קשים, כגון משטחים מחזירי אור (רפלקטיביים), אובייקטים נעים, זיהוי פנים, רובוטיקה, וכן סיווג קל יותר של המידע בתמונה ישירות מטרם העיבוד ועד לתוצאת הסיווג, דבר שנדרש ביישומים רבים. עם זאת, רשתות CNN אינן יכולות לכסות את כל תחומי עיבוד התמונה הקלאסיים, כגון קביעת מיקום מדויק של אובייקטים - כאן יהיה צורך לפתח רשתות CNN חדשות ומתקדמות.

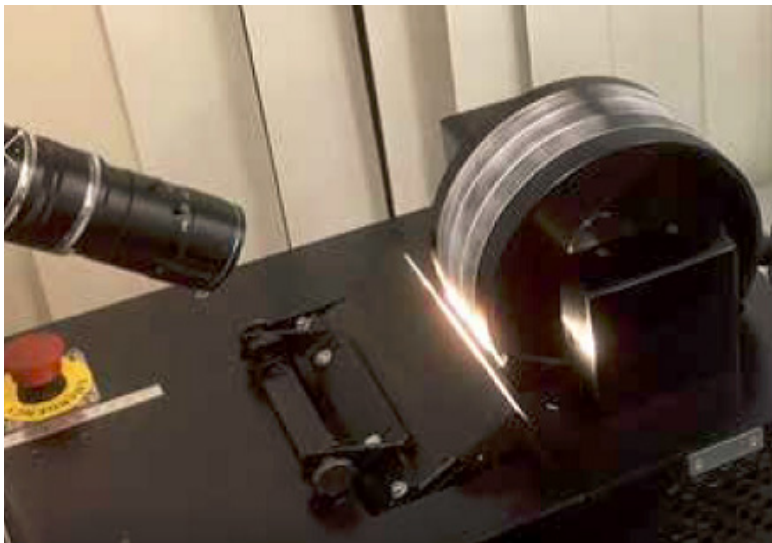
ישומי עיבוד תמונה קלאסיים מוגבלים במקרים הבאים: האובייקטים הנבדקים מעוותים, בעלי צורות לא סטנדרטיות, בעלי וריאציות גדולות, סביבת התאורה אינה מתאימה/מבוקרת או קיימים עיוותים אופטיים. אם, כמו במקרים אלה, לא ניתן לשלוט בתנאים הסביבתיים של רכישת התמונה, גם אלגוריתמים ספציפיים לתיאור מאפיינים כמעט ולא ניתנים ליישום. רשתות CNN, לעומת זאת, מגדירות מאפיינים באמצעות שיטת האימון שלהן, ללא שימוש במודלים



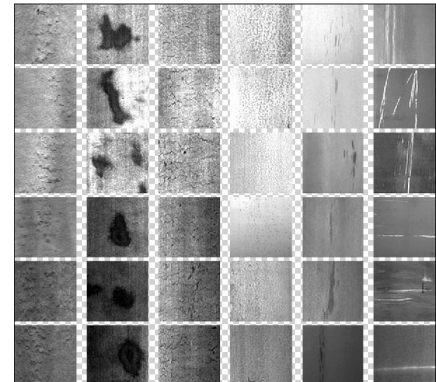
תמונה 2: CNN משולב לתוך VisualApplets. מקור: Silicon Software GmbH

על מערכות משובצות (embedded) או על רובוטים ניידים בעלי פיזור החום הנמוך הנדרש. בעתיד, המגוון והמורכבות של היישומים המבוססים על רשתות נירורנים על גבי מעבדים מיוחדים שפותחו לאחרונה יגדלו עוד יותר. בפיתוח של פתרונות חומרה ותוכנה חדשים ולמען שיתוף תוצאות מחקר, כחלק מפרויקט ה-FPGA "למידת מכונה עבור יישומים תעשייתיים", מתבצע שיתוף פעולה של חברת Silicon-Software עם פרופ' מיכאל הייזמן (Michael Heizmann) מהמכון לטכנולוגיית מידע תעשייתי (IIT) במכון Karlsruhe לטכנולוגיה (KIT). בעתיד צפויים לצמוח חידושים נוספים משיתוף פעולה זה.

בכדי ליישם רשתות נירורנים גדולות במיוחד ליישומי CNN מורכבים, שוחררה גרסת כרטיס לכידת תמונה (frame grabber) שניתן לתכנות, בסדרת Marathon microEnable שמכיל כמעט שלוש פעמים יותר משאבי FPGA לעומת סדרת הכרטיסים הנוכחית והוא צפוי להיות אידיאלי עבור רשתות נירורנים בעלי רוחב פס CNN של יותר מ-GB1 לשנייה. רשתות ה-CNN לא רק מסוגלות לרוץ על גבי FPGAs של כרטיסי הלכידה אלא גם על מצלמות וחיישנים תואמי VisualApplets. מכיוון שה-FPGAs הם עד פי עשר יותר חסכוניים באנרגיה בהשוואה ל-GPUs, יישומים מבוססי CNN יכולים להיות מיושמים בצורה טובה במיוחד



תמונה 3: מדגים לניתוח וסיווג פגמים על משטחים מתכתיים ע"י רשתות CNN. מיושם באמצעות Visual Applets. מקור: Silicon Software GmbH

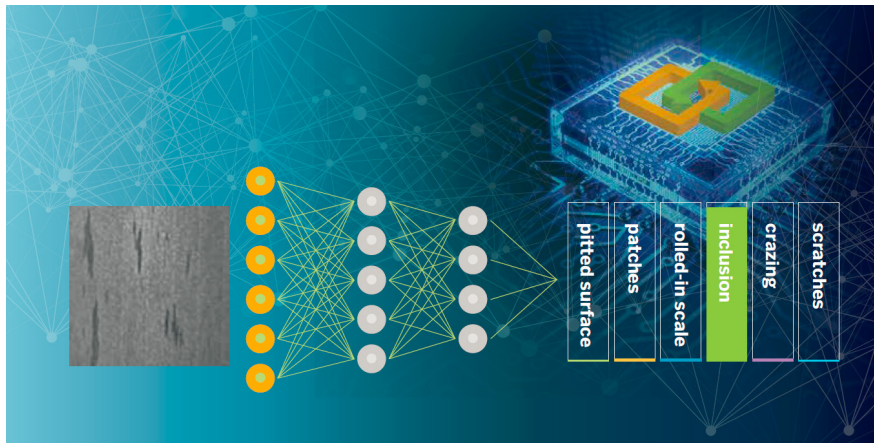


תמונה 1: סוגי פגמים שזוהו (משמאל לימין: אבנית מגולגלת, טלאים, סדקים, משטח מגולוון, הכללה, שריטות). מקור: NEU Surface Defect Database

לנתח ולסווג מידע מהתמונות ברזולוציה גבוהה בזמן אמת. בראיית מכונה (Machine Vision), רכיבי FPGA מתפקדים כמאיצים מסיביים של משימות עיבוד תמונה ומבטיחים תהליכי עיבוד בזמן אמת עם זמני שהייה דטרמיניסטיים. עד כה, המאמץ התיכנותי הגדול והמשאבים הנמוכים יחסית שהיו זמינים ב-FPGA מנעו שימוש יעיל. הפשטה של האלגוריתמים מאפשרת כעת בנייה של רשתות יעילות עם קצבי מידע גבוהים ב-FPGA.

אופרטורים של CNN וכרטיסי לכידה חדשים

ניתן להשתמש בסביבה הגרפית VisualApplets שפותחה ע"י Silicon Software בכדי ליישם רשתות CNN על גבי פלטפורמות חומרה FPGA. אופרטורים של רשתות CNN ב-VisualApplets מאפשרים למשתמשים ליצור ולסנתז תכן יישומי FPGA מגוונים תוך זמן קצר, ללא ניסיון בתכנות חומרה. על ידי העברת הפרמטרים של המשקל והשיפוע (gradient) שנקבעו בתהליך האימון לאופרטור רשתות ה-CNN, תכן ה-FPGA מוגדר למשימה ספציפית ליישום. האופרטורים יכולים להיות משולבים כחלק מתכן תרשים הזרימה של VisualApplets עם מקורות מצלמה דיגיטלית כקלט תמונה ובתוספת אופרטורים של עיבוד תמונה לאופטימיזציה של העיבוד המקדים, לפני העברת התמונה לזיכרון המחשב.



תמונה 3: אילוסטרציה של תמונה המכילה פגם ורשת נירונים המסווגת את הפגם. מקור: Silicon Software GmbH

כדי תנועה על גבי מידע התמונה החדשה שנלכדה, מה שיגדיל עוד יותר את קצב הגילוי ויפשט את תהליך הלמידה. חברת Silicon-Software פיתחה כרטיסי לכידה ועיבוד תמונה ואת תוכנת Visual Applets המיועדת לפיתוח על בסיס חומרה זו. מוצרי החברה משווקים בישראל על ידי Vision1.

זוהו כפגומים, כמו גם סגמנטציה ואחסון, טרם יושמו ביישומים תעשייתיים. תחזית: עד כה, אימון רשת נירונים ושימוש בזמן אמת של רשתות CNN הם שני תהליכים נפרדים. בעתיד, דורות חדשים של FPGA בעלי משאבים גבוהים יותר או המשתמשים ב-ARM / ליבות CPU ו-GPU חזקים יאפשרו אימון תוך

על מנת לקבוע את אחוז הפגמים שזוהו כהלכה בתנאים סביבתיים קשים, אומנה רשת נירונים עם 1,800 משטחים מתכתיים מחוזירי אור, שבה הוגדרו שישה סוגים שונים של פגמים. הבדלים גדולים בשריטות ביחד עם הבדלים קטנים בסדקים, בשילוב עם משטחים בגוונים שונים של אפור בשל שינויים בתאורה ובחומר, הפכו את ניתוח המשטח לכמעט בלתי אפשרי עבור מערכות עיבוד תמונה קונבנציונאליות. תוצאות הבדיקה הראו כי הפגמים השונים סווגו באופן חיובי על ידי רשת הנירונים בממוצע של 97.4%, שהוא ערך גבוה יחסית לשיטות הקלאסיות. קצב עיבוד המידע בתצורה זו היה 400MB לשנייה. לשם השוואה, פתרון תוכנה מבוסס CPU השיג ממוצע של 20MB לשנייה.

יישום הלמידה העמוקה בטכנולוגיית מעבדי FPGA, אשר מתאים באופן אידיאלי לעיבוד תמונה, הוא צעד חשוב. הדרישה התעשייתית לכך שהשיטה צריכה להיות דטרמיניסטית וניתנת לאימות על ידי אלגוריתמים חד-חד ערכיים, ייגרמו לכך שכניסת הטכנולוגיה החדשה לכל תחומי עיבוד התמונה תהיה קשה יותר. כמו כן, אפשרויות לתיעוד אילו אזורים שנבדקו