

草刈り機への人の手巻き込み防止に対する機械学習と その FPGA 向け設計の研究

山本 椋太[†] 北村 偉[†] 稲川 清[†] 阿部 司[†]

苫小牧工業高等専門学校[†]

1. はじめに

近年、農業従事者の高齢化[1]やなり手不足の問題から、農業の効率化が重要となっている。特に北海道においては、農地面積が広大であることが多く、農業機械を導入することで農作業の負担を軽減することが必要となる。しかし、農業機械を用いた農作業中の事故は毎年発生しており[2]、危険の防止が必要である。

安全装置の導入のためには、コスト・消費電力・サイズなどの障壁があり、これらのニーズに合わせたシステムを実現することが必要であると考えられる。

本研究では、安全装置としてシステムを構築する前に、これらのニーズに対応できるかを明らかにしたい。

そこで本研究では、GPU より電力に対する性能の効率がよく、CPU よりも高速に実行可能な FPGA に着目した。FPGA を用いて安全装置を開発することとしたとき、以下のことを検討した。

- 危険な状態を、機械学習を用いて検出するとき、どの程度の作業が必要か。
- 機械学習の結果を FPGA で実現するとき、量子化せずに CNN を実現させると、どのような性能・回路面積になるか。

以上のことについて、農業機械として草刈り機を用いて調査したため、報告する。

2. 関連研究

2.1 これまでの取り組み

我々は以前より、農業機械に対して FPGA を用いることについて検討を進めてきた[3]。物体検出のためにステレオカメラと FPGA を用いており、視差画像から物体との距離の計測は可能だが、物体種別の判別はできない。

そのため、ステレオカメラと本研究の成果を組み合わせることで、将来的には物体までの距離とその物体の判別の機能を持つ安全装置の開発を考えている。

また本研究は、我々による先行研究[4]を前提としている。この研究では、2つの草刈り機を対象に、手を含む画像かどうかの識別を行った。この先行研究において、いくつかの学習データセットで学習の試行を行ったが、機械学習による正解率は68%から92%の幅があり、80%程度の正解率のなっていたため、向上の余地がある。具体的な問題点として、画像の枚数が少ないこと、「ブレ」の画像がないことがある。

また、先行研究の時点では、ハードウェア化が未完了であったため、ハードウェア化についても検討した。

2.2 システムレベル設計環境 SystemBuilder

本研究においては、ソフトウェア（以下、SW と書く）とハードウェア（以下、HW と書く）の協調設計が可能な SystemBuilder を用いて開発を進める。SystemBuilder は、C 言語によって HW・SW を設計する。

C 言語記述について、HW 部分は高位合成ツール、SW 部分はコンパイラへと渡される。また、SystemBuilder ではプロセスと呼ばれる単位で開発を行うが、HW-HW、SW-SW、および HW-SW 間の通信をサポートしている。

この環境を用いることによって、HW 単体の開発では作業コストが大きい設計作業の効率化を図る。具体的には、機械学習の実装におけるデバッグや動作確認を、まずは SW で行ってから HW で実現する。これにより、HW 設計時のデバッグにかかる時間的なコストを低減できる。

3. 実施内容

本研究では、2.1 節で述べた通り、以下について、実装および評価を行う。

- 先行研究の時点よりも画像枚数を増加させた上での機械学習
- 各レイヤの C 言語実装および機械学習結果を C 言語に移行
- FPGA 上で実現

3.1 機械学習の改善

先行研究における機械学習における課題は、2.1 節にて述べた通りである。

Reasoning for prevention of human hand entrapment in mower and its design for FPGA

[†]Ryota Yamamoto, Takeru Kitamura, Kiyoshi Inagawa, Tsukasa Abe, National Institute of Technology, Tomakomai College

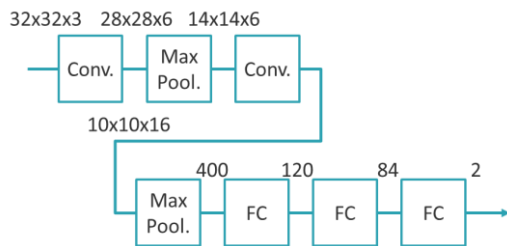


図 1. CNN の構造

具体的な改善方法として、動画撮影を実施した。動画を用いることで、「ブレ」を含む画像の取得も可能となった。本研究においては、「手が手首まで完全に含まれる（手あり）」、「手が全く映り込んでいない（手なし）」、「手を含むが、一部しか映り込んでいないまたは、手が含まれるが、画像にブレがある（曖昧）」の3種類の画像を、各1200枚用意し、学習用画像を1000枚（内200枚が検証用）、テスト用画像を200枚用意した。

CNNの構造を図1に示す。先行研究から、CNNを変更していない。図1の前に、前処理を行っており、画像のリサイズ（トリミング・縮小）のみ行った。実装は、Google Collaboratoryにおいて、Tensorflow+Kerasを用いてPythonde実装した。

学習結果について、表1の通り正解率で評価した。正解率は、テスト用画像に対する正解の枚数の割合である。なお、「正解」とは、手あり・曖昧の画像を「手あり」と識別するか、手なしの画像を「手なし」と識別することを指す。

結果から、先行研究よりも正解率を改善することができた。ただし、分類が「曖昧」の入力画像は、他と比べて正解率がわずかに低い。

3.2 SoC ボード上での実現

本研究では、ZYBO Z7-20をボードとして用いる。このボードのSoCチップには、FPGAとCPU（Arm Cortex-A9）が搭載されている。

まず、このために学習済みパラメータをテキストファイルに変換して、C言語から読み込みやすいようにした上で、各レイヤの実装をPythonからCへ移行した。移行したソースコードをすべてHWに指定したところ、資源不足が生じた。レイヤごとのHW資源の利用状況を表2に示す。表2の通り、全結合層のBRAM使用量が400%を超えること、すべてのレイヤの資源の和を見るとBRAM以外の資源も不足する。

そこで、特定のプロセスのみHW化した場合の実行時間を測定した。画像10枚分の識別処理を行い、その実行時間を10で割った数値を実行時間の平均と呼んでいる。その結果、プーリング層のみ高速化したことがわかった。

表 1. 入力画像種別ごとの草刈り機ごとの正解率

	手あり	曖昧	手なし
刈払機	99.5%	99.5%	91.5%
手押し 草刈り機	91.5%	86.5%	98.0%

表 2. レイヤごとの HW 資源利用状況

ハードウェア化した層	平均処理時間[μ s]
畳み込み層1	75644
プーリング層1	14578
畳み込み層2	40235
プーリング層2	12154
全てソフトウェアの場合	17569

そのため、プーリング層1・2のみハードウェア化した場合についても性能を評価した。

4. おわりに

本研究では、機械学習からHWの設計まで、安全装置に利用する手の検出処理をFPGA向けに設計した。特に、HW設計に関しては、今回高速化として量子化などを施さずに設計したが、高速化を施さない場合には畳み込み層の実行性能が悪くなることがわかり、プーリング層は高速化できることがわかった。また、全結合層はHW化が困難となることがわかった。

今後の課題として、HWの高速化・効率化が必要である。メモリアクセスのレイテンシの隠蔽やプロセスレベルのパイプライン化などを行い、より高性能な設計の検討を進めていく。

謝辞

謝辞 本研究は JSPS 科研費 21K05856 の助成を受けた。

参考文献

[1] 農林水産省, 令和4年度 食料・農業・農村白書, 2023年
 [2] 北海道農作業安全運動推進本部, 令和4年度農作業事故報告書, 2023年
 [3] 横山 港, 稲川 清, 大西 孝臣, 阿部 司, 吉村 斎, 大村 功, 三田村 智行, “OCMを用いた視差画像出力システムの FPGA への実装と物体検出に関する検討”, 2018年度精密工学会春季大会, 精密工学会学術講演会講演論文集, pp. 293-294, 2018年
 [4] 北村 偉, 山本 椋太, 三上 剛, 稲川 清, “深層学習に基づく草刈り機向けの手の巻き込み検出の試作”, 情報処理学会第85回全国大会, pp. 1-117 -- 1-118, 2023