別記様式第４号

2021 年度 独創的研究助成費 実績報告書

２０２２年　３月３１日

<table>
<tr><td>報　告　者</td><td>学科名</td><td>情報通信工学科</td><td>職 名</td><td>助教</td><td>氏 名</td><td>小椋清孝</td></tr>
<tr><td>研 究 課 題</td><td colspan="6">軽量化画像処理DNN のハードウエア化に関する研究</td></tr>
<tr><td rowspan="3">研 究 組 織</td><td colspan="2">氏　　名</td><td>所 属 ・ 職</td><td colspan="2">専 門 分 野</td><td>役　割　分　担</td></tr>
<tr><td>代表</td><td>小椋　清孝</td><td>情報通信工学科・助教</td><td colspan="2">集積回路</td><td>研究全般</td></tr>
<tr><td>分担者</td><td>井口周太郎</td><td>情報系工学研究科
博士前期課程１年</td><td colspan="2">デジタル回路
設計</td><td>回路設計・評価</td></tr>
<tr><td>研究実績
の概要</td><td colspan="6">１．背景と目的
高精細画像の機器間伝送量削減方法として，送信時に一部の画素を伝送せず，受信側で周囲の画素情報から深層学習ネットワークによる推論器で欠損部分の画素を復元する手法を提案し，検討を行っている．対象機器へ組み込む必要性から，この推論器には，できるだけ小型であり，再生遅延を防ぐために低レイテンシーであることが求められる．高精細画像をリアルタイムかつ低レイテンシーで処理するためには，ハードウエア回路による並列処理・パイプライン処理等が必要になる．

推論器の最適なネットワーク構成については現在検討中だが，これらをハードウエア化する際，レイテンシーを押さえるためにカウンタを用いない構成で設計すると小規模FPGAには実装できないほどの回路規模になることがわかっている．
そこで，本研究では，推論器ハードウエアの小型化のために，深層学習ネットワークの軽量化技術の一つである枝刈りを適用してその有効性を検証した．MNISTのような識別問題で枝刈りを適用すると，ほとんど識別率を低下させずにネットワークの重みの数を90%以上削減可能であることは既によく知られている．一方，今回扱う推論器は欠損部分の画素値を予測する回帰モデルであり，このような回帰問題での枝刈りによる出力値への影響については，扱う問題により変わってくると考えられる．よって，今回，本研究で扱うモデルへの枝刈りの有効性の検討を行った．
また，枝刈り後のネットワークは疎な状態になり，ハードウエア化の際に，カウンタによるニューロンやシナプスへの連続アクセスを前提とした回路構成はそのままでは用いることができない．これに対しては，ネットワークをそのまま回路に転写する形のカウンタを用いない構成や，有効なシナプスとニューロンの組み合わせを管理する構成等を現在検討中である．</td></tr>
</table>

※ 次ページに続く

研究実績 の概要	２．方法 全結合ニューラルネットワークモデルの場合，まずモデルの学習を行う．学習済みのネットワークモデルから重みを取り出し，重み係数が閾値以下の部分を0，それ以外を1としたマスクデータを作成する．このマスクデータとネットワークの重みを掛け合わせたものを学習に用いる重みとして，再度追加学習を行うことで，枝刈りしたネットワークによる推論器モデルを得ることができる．このモデルを用いてテスト画像の復元を行い，PSNR等により推論能力の評価を行う． 畳み込みニューラルネットワークモデルの場合も同様に，初期学習で得られた畳み込みカーネルの係数値に対して閾値判定を行い，閾値以下の係数を0にすることで枝刈りを行う． ３．結果 復元対象画素を中心とする周辺5x5の画素情報から画素値推定を行う全結合型ニューラルネットワークモデルを対象に枝刈りの影響を検証した例を示す．ネットワークの構成は，入力54，中間層20ニューロン，出力3の3層構造である．これに対し，入力―中間層間の枝刈りを行った． 初期学習(100epoch)後のloss値(テストデータ)は1.45，50枚のテスト画像(サイズ500x300程度)の平均PSNRは35.48 dBであった．重みデータをとりだし，係数値が−0.1～0.1の範囲の重みを無効化するマスクを作成した．無効化部分は全体の59.4%である．このマスクを用いて枝刈り対象のシナプスを無効化した状態で追加学習(100epoch)を行うと，loss値は1.30，平均PSNRは35.60となり，元の40%のシナプス構成でほぼ同程度の性能を維持するという結果となった．このことから，この研究で扱う推論器のモデルでも枝刈りが有効であるということが明らかとなった．
成果資料目録	