

機械学習を用いた物体認識技術に関する研究

近藤 正一, 全 慶樹, 藤澤 怜央, 堀 武司

Study on Object Detection Method by Machine Learning

Shouichi KONDOU, Keiki ZEN, Reo FUJISAWA, Takeshi HORI

抄 録

近年、物体認識性能は機械学習により著しく向上している。本研究では、機械学習による物体認識手法である畳み込みニューラルネットワーク（CNN）を用いて、一次産業を中心とした様々な分野への適用を試みた。小麦粒の赤かび病発生判定、木材断面領域の検出、じゃがいもの芽領域の検出および動物の姿勢推定について、機械学習データセットを構築した。CNNによる学習および識別を行い、適切なデータセットを整えることで良好な結果が得られることがわかった。

キーワード：機械学習, ディープラーニング, 画像解析, 物体認識, 畳み込みニューラルネットワーク

Abstract

Recently, the performance of object recognition has been improved remarkably by machine learning. In this research, we tried application of convolutional neural networks (CNN) which is object recognition method by machine learning to various fields focus on primary industry. We build machine learning datasets for diagnosing of Fusarium blight of wheat grain, detection of wood cross section, detection of potato-eyes and pose estimation of animals. We trained and tested by CNN, it was found that good results are obtained by preparing appropriate datasets.

KEY-WORDS : Machine Learning, Deep Learning, Image Analysis, Object Recognition, Convolutional Neural Network

1. はじめに

近年、様々な分野で自律的に動作するシステムの開発が行われている。これらのシステムでは自ら環境を認識し、判断し、行動（処理）することが求められる。この中で環境を認識するために不可欠な技術要素として、画像から得られるシーンから物体を抽出したり、それが何であるかを認識する物体認識技術がある。

画像のカテゴリ分類を行う一般物体認識は、画像認識の分野において重要かつ困難な課題として取り組まれてきたが、近年著しく性能が向上し、国際的な大規模画像認識コンペティションILSVRCにおいて2015年には人間を超える精度を達成した¹⁾。現在この分野において最も顕著な成功を収めているのは、畳み込みニューラルネットワーク（CNN）と呼ばれる、脳の視覚野の構造に関する知見をもとに考案された機

械学習手法である。機械学習はデータを元に、識別モデルのパラメータを決定する手法であり、良質な訓練データが必要である。動物や乗り物、機械等の様々な種類の画像1,400万枚、2万クラスを超える大規模な画像データセットであるImageNetの登場、それに伴うCNNモデルの複雑化、計算機性能の向上等の要因により、高い認識性能が達成されたと考えられている。

また、画像のカテゴリ分類より難しい処理である物体検出、物体領域検出、骨格推定等（図1）へもCNNが応用され、これらに対しても従手法より著しく性能が向上し、実用レベルに近づいている。

当場では、これらの物体認識技術を様々な分野において活用する検討を進めている。本研究では、一次産業を中心としたいいくつかの事例について適用結果を報告するとともに活用の可能性について検討した。

事業名：経常研究

課題名：機械学習による一般物体・空間・現象認識

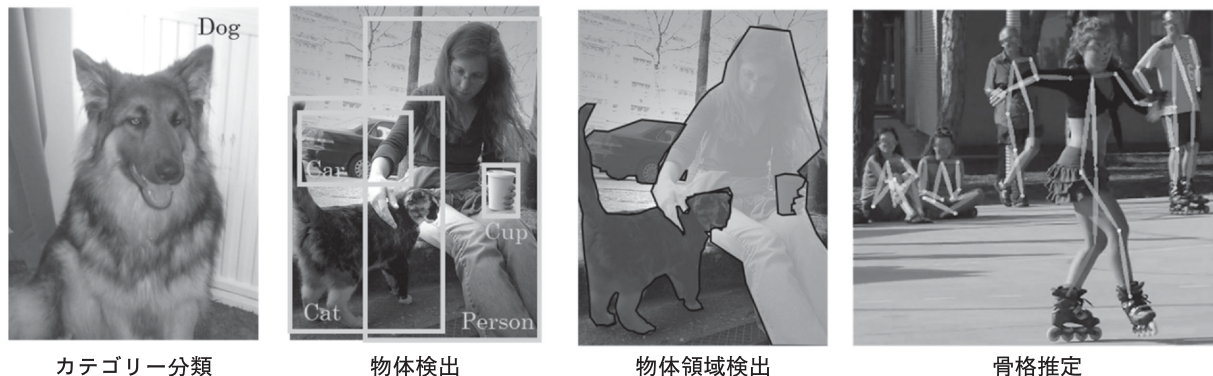


図1 画像認識における処理

2. 機械学習による物体認識技術

2.1 畳み込みニューラルネットワーク (CNN)

畳み込みニューラルネットワーク (CNN) は、画像認識分野において高い性能を達成している機械学習手法である。従来の画像認識手法では、認識・解析する内容に応じて熟練者が適切な空間フィルタの設計をする必要があった。一方、CNNは、フィルタの係数を大量のデータを用いた機械学習により決定することができる。良質なデータによる学習を行うことで高い認識性能が期待できる。

図2はCNNによる画像分類の例である。CNNは大きく特徴抽出層と分類器に分けられる。特徴抽出層では空間フィルタによる特徴量の抽出、分類器では抽出された特徴量による分類を行うことで、入力された画像を対応したカテゴリーに分類する。処理に必要な係数は対象とする物体の画像を学習させることで得られるため、様々な物体の検出・分類に広く用いることができる。特に、対象物の色や形状がさまざま、撮像環境の変化や汚れなどがある一次産業分野では、従来手法では精度向上が困難であったため、今後の自動化を進める上で重要な技術となる。

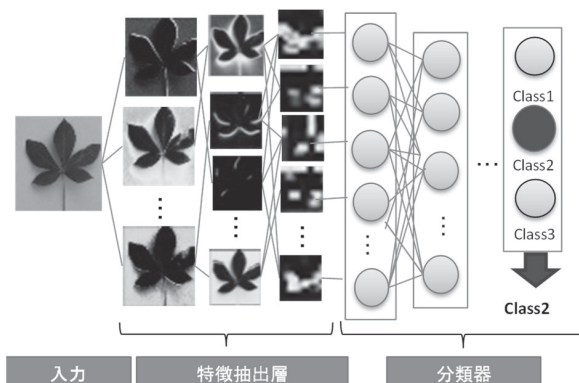


図2 CNNによる画像分類の例

2.2 物体検出手法・物体領域検出手法

物体検出手法は、画像内の物体のカテゴリー分類に加えてその位置と大きさを検出する手法である。また、物体領域検出は、物体の形状を画素レベルでのカテゴリー分類を行って検出する手法である。

CNNを用いた代表的な物体領域検出手法であるMask R-CNN²⁾を図3に示す。入力画像に対して①特徴抽出を行い、得られた特徴量をもとに②領域候補検出を行う。各領域候補内の特徴量に対して畳み込み演算を行うことで、③位置調整、

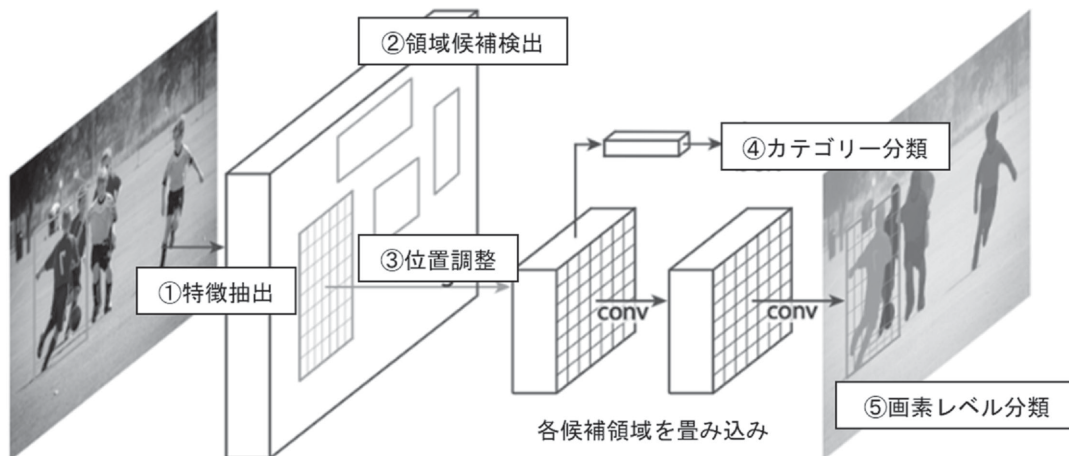


図3 CNNによる物体検出手法 (Mask R-CNN)

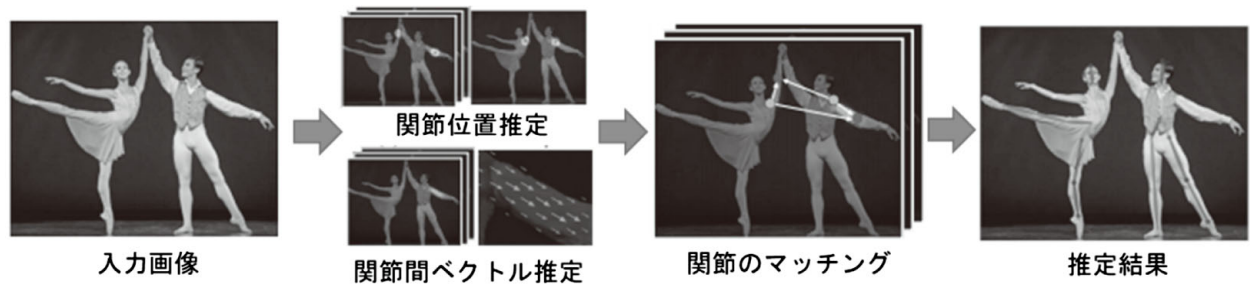


図4 CNNによる骨格推定手法 (OpenPose)

④カテゴリー分類および⑤画素レベルの分類を行い、物体の形状を検出する。

画像分類手法であるCNNにおいて、後段の分類器に相当する部分を置き換えることで、物体の位置や形状等の認識を行うことができる (図3)。

2.3 骨格推定手法

骨格推定手法は、物体の姿勢を推定する手法である。単純化した骨格モデルと関節位置を学習させることで、骨格の状態として姿勢を推定することができる。人間に対しては大規模で良質なデータセットが公開されており、リアルタイムに複数人数の骨格を推定する手法が提案されている。代表的な骨格推定手法であるOpenPose³⁾を図4に示す。入力画像から抽出した特徴量をもとに関節位置推定と各関節をつなぐベクトルの推定を行い、それらの連結性を評価することで骨格の推定を行っている。

姿勢や動きの評価が可能となることから、家畜等の健康管理、人の作業負荷計測などへの適用に応用できる。

3. 機械学習による物体認識技術の適用

これまでに検討を進めた物体認識手法により、実際の課題に対して、適用した結果を報告する。

3.1 学習用データセットの作成

機械学習の適用には適切なデータセット (データと教示値との組合せ) を人手により用意する必要があるため、各課題についてデータセット構築を行った。また、多くのサンプルを収集できなかった場合、拡大、縮小、回転等の処理を施すことでデータセットの拡張を行うことも有効な方法である。本研究では、これらの作業の一部について、ツール作成を行い、効率化をはかった (図5)。

3.2 解析システム

機械学習による解析を行う場合、多くのデータを用いた繰り返し学習が必要となる。このため、本研究では高速な演算機能を有するGPU (Nvidia GeForce GTX 1080 Tiまたは

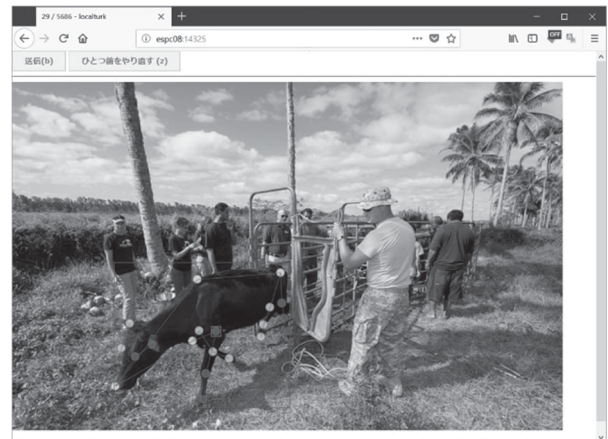


図5 作成したデータセット構成ツール

Nvidia GeForce GTX Titan X) を搭載したPC (OSはWindowsまたはLinux) を用いた。また、ソフトウェアの実装にはオープンソースのCNN用ライブラリであるKerasを利用した。

3.3 小麦粒の赤かび病発生判定

小麦に発生する赤かび病は収量や品質に大きな影響を及ぼす病害であり、赤かび病に強い品種の育成や、適切な時期の適切な薬剤散布による防除等の対策が重要である。

赤かび病の発生有無は外観で判別することが出来る (図6)。外観による判別は熟練者による目視で行われており、画像解析で判定することができれば、より簡便な検出が可能となる。本研究では、機械学習による物体認識技術を用いた小麦粒の赤かび病の発生有無の判別として、小麦粒が並んだ画像から個体分離をし、分離された各小麦粒について赤かび病の発生

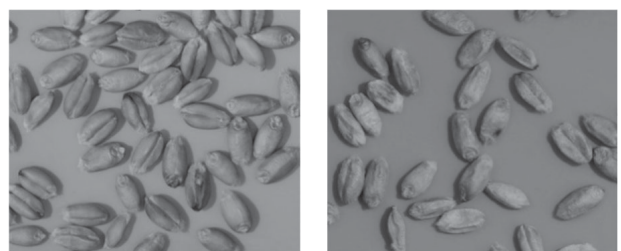


図6 健全粒 (左) と赤かび発生粒 (右)

を判別するという二段階の処理による判別を行った。

個体分離については、物体領域検出手法であるMask R-CNNを用いた。小麦粒が並んだ画像17枚に対して人手で各個体の輪郭を記録したデータと、それらを拡大・縮小・回転等の処理で拡張した1,000枚の画像を学習用データとして用いた。計算には、NVIDIA GeForce GTX 1080 Tiを搭載したPCを用いた。2回の繰り返し学習を行い、学習にかかった計算時間は約50分だった。図7に個体分離の検出結果を示す。

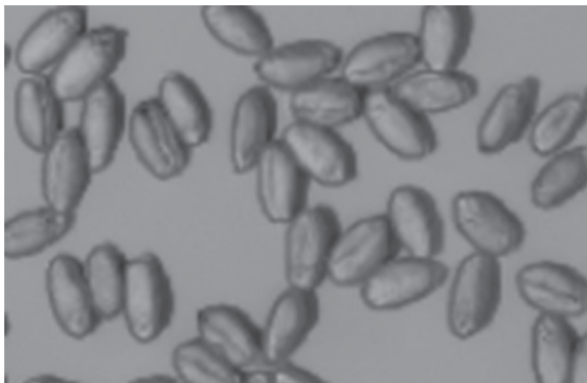


図7 Mask R-CNNによる個体分離結果

赤かび病の発生の判定は、CNNによる画像分類手法のひとつであるResNetを用いた。学習用に1,154枚の小麦粒画像を用い約800回の繰り返し学習を行った。学習にかかった時間は約200分だった。384枚の小麦粒画像を用いて評価した結果、そのうち380枚で検査員の目視による判定結果と一致した。検出率は99.0%である。誤判定した4枚(図8)はすべて、検査員が健全粒と判定したものを赤かび病発生粒と判定したものであった。今後、検査員にこれらの画像を提示し、学習データを厳格化することで判定精度の向上を目指す。また、穀物や豆など、類似した検出を行うケースは多くあると考えられ、コスト等を含めた適用の可能性について、さらに検討を進める。

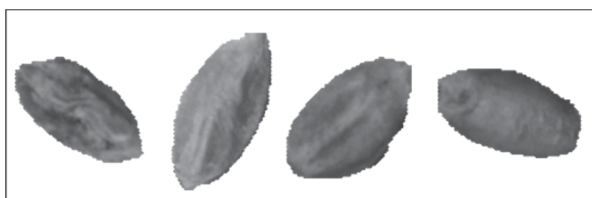


図8 誤判定した個体(健全粒を赤かび粒と判定)

3.4 木材断面領域の検出

林業において、伐採された木は丸太の状態に加工され、その価格は樹種や直径等により決定される。直径の計測は、図9のように土場に積まれた木材をものさし等により手作業で計測する場合が多く、簡便な計測手段による労働負荷の低減が求められている。



図9 土場に積まれた木材

本研究では、土場に積まれた木材の写真に対して機械学習による物体認識技術を適用し、丸太の断面領域検出を行うことで計測方法の簡便化を検討した。

手法はMask R-CNNを用いた。学習用データとして人手で画像内の約90本の丸太の断面の輪郭を記録し、それらを拡大・縮小・回転等の処理で拡張した1,000枚の画像を用いた。計算には、NVIDIA GeForce GTX 1080 Tiを搭載したPCを用いた。2回の繰り返し学習を行い、学習にかかった計算時間は約27分だった。図10に丸太断面の検出例を示す。各丸太の断面形状が検出できており、断面形状から直径を求めることができる。

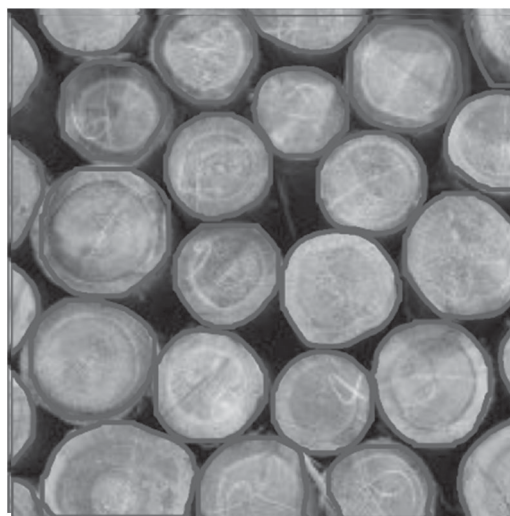


図10 Mask R-CNNによる木材の断面領域検出結果

Mask R-CNNは、上記のような高速な演算機能を有するGPUでは1秒で画像数枚程度の処理が可能であるが、タブレット等の計算性能では1枚の画像を処理するのに数分以上かかる。そこで、現場にて簡便な計測を可能とするために、タブレット等で動作する、計算量を大幅に低減したCNNモデルを用いたアプリケーションの試作を行い、実用化を進めている(図11)。一般的に機械学習の演算は計算コストが高く、特に学習に時間がかかるため、学習時は計算性能の高いGPUを搭載したPCにより計算を行い、作成した機械学習モデルのデータを用いて認識処理のみタブレットで行うことで



図11 タブレット上で動くアプリケーションの試作

実用システムの効率的な開発が可能となる。

今後は、撮影時に大きさのわかる基準となるものを置き、画像から読み取るなどすることで、実用化の取組をはかっていく。

3.5 ジャがいもの芽領域の検出

加工食品用のジャがいもは芽や痛み等の不用部除去が必要であり、現場においてもそれらの自動除去を行うシステムの開発を行っている⁴⁾。除去性能の向上のためには、芽や痛みの位置の正確な計測が必要となる。そこで、物体領域検出技術によるジャがいもの芽領域の検出に取り組んだ。

学習用に皮を剥いたジャがいもの画像180枚を用い、20回の繰り返し学習を行った。計算には、NVIDIA GeForce GTX 1080 Tiを搭載したPCを用いた。学習にかかった計算時間は約47分だった。図12にジャがいもの芽領域の検出結果を示す。ジャがいもと芽の領域の輪郭が検出できることを確認した。今後、より多量のデータを用いた学習による高精度化や芽以外の傷や痛み等の検出への適用を検討する。

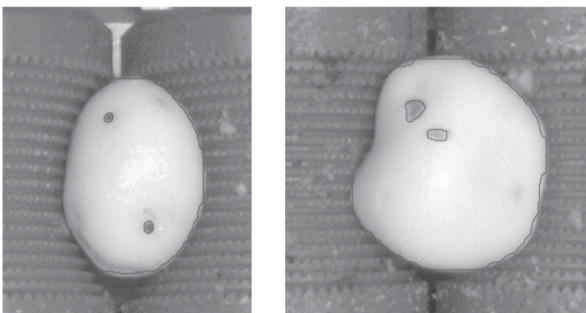


図12 Mark R-CNNによるジャがいもの芽検出結果

3.6 動物の骨格推定

動物の動作解析は、従来マーカやセンサの装着や複数台カメラによる撮影等の環境が必要であった。しかし、CNNによる骨格推定手法の発展により、画像のみからリアルタイムに姿勢推定が行うことが可能となり、特別な環境を準備せず

に動画から動作解析が行える可能性が出てきた。そこで、家畜の健康管理等への活用のため、動物の骨格推定に取り組んだ。

推定する馬の骨格は図13に示すモデルとした。学習用データの構築には、ImageNetおよびMicrosoft COCOデータセットに含まれる馬と牛の画像を用いた。作成したデータセット構成ツールを用いて、画像内の各馬と牛の関節位置の記録をしたものを学習用のデータとした（馬：約1,700枚、牛：約200枚）。

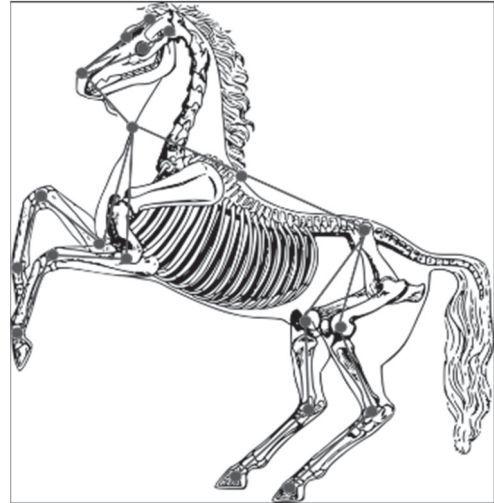


図13 推定する馬の骨格モデル

骨格推定手法はOpenPoseを用いた。計算には、NVIDIA GeForce GTX Titan Xを搭載したPCを用いた。約2,000回の繰り返し学習を行い、学習にかかった計算時間は約7日間だった。

図14に関節位置および関節間ベクトルの推定結果および、それらを用いた最終的な骨格推定結果を示す。画像中の複数の馬について、異なる馬の関節が接続することなく推定できている。また、図15にその他の馬や牛の最終的な骨格推定結果を示す。色や柄の異なる個体に対しても骨格推定が出来ることを確認した。遮蔽により推定できていない関節もあるため、今後データセットの拡充等により改善を検討する。

また、異常動作検知等への応用については、時系列での動作解析を行い、処理手法の検討を進める。

4. おわりに

本研究では、機械学習による物体認識手法である畳み込みニューラルネットワーク（CNN）を用いて、小麦粒の赤かび病発生判定、木材断面領域の検出、ジャがいもの芽領域の検出および動物の姿勢推定を行い、一次産業を中心とした様々な分野への適用を試みた。その結果、実用化には検討すべき課題があるが、計測環境を整えることで良好な結果が得られることがわかった。

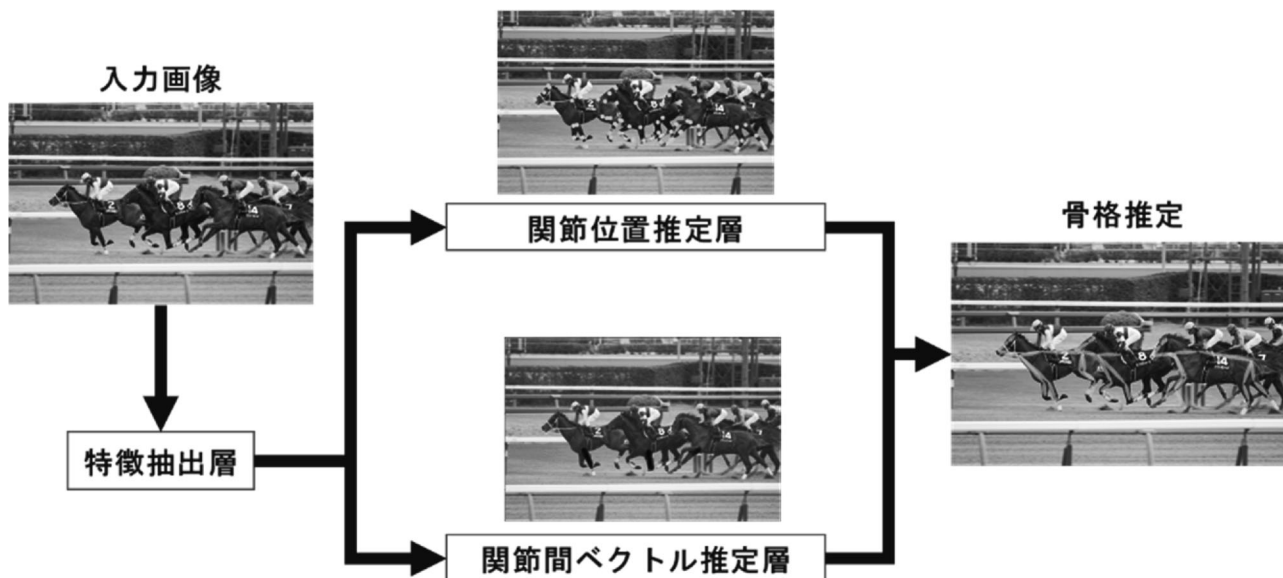


図 14 馬の骨格推定の様子

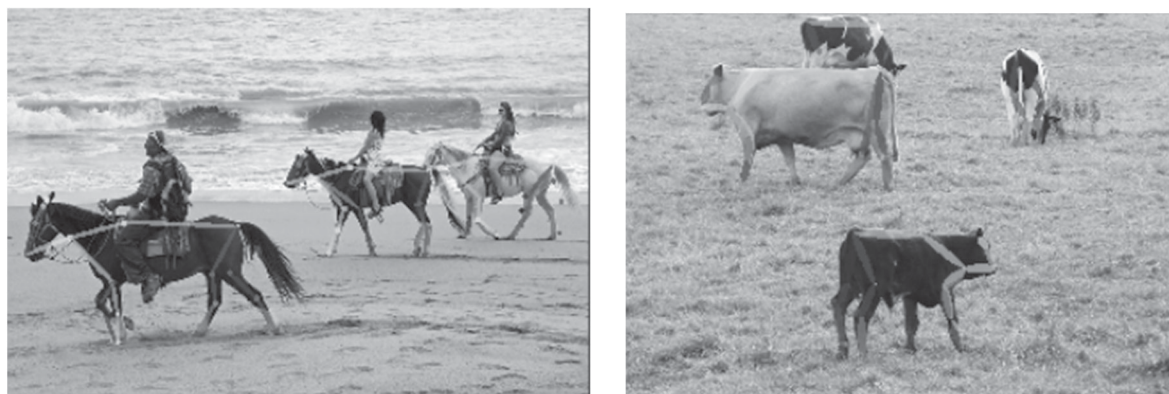


図 15 馬・牛の骨格検出結果

今後は、実用化のための課題を明確にし、各応用事例の実用化の検討を進めるとともに、他の産業応用についても本研究の手法の適用可能性について検討を進めていく。

4) 井川 久・中西洋介・川島圭太他：じゃがいもの不用部除去システムの開発，北海道立総合研究機構工業試験場報告，No.317，pp.23-30（2018）

参考文献

- 1) Kaiming He, Xiangyu Zhang, Shaoqing Ren, et al.: Deep Residual Learning for Image Recognition, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, (2016)
- 2) Kaiming He, Georgia Gkioxari, Piotr Dollar, et al.: Mask R-CNN, The IEEE International Conference on Computer Vision (ICCV), pp. 2961-2969, (2017)
- 3) Zhe Cao, Tomas Simon, Shih-En Wei, et al.: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7291-7299, (2017)