

3.1.2 餌資源の検証調査

平成22年度の調査で、海ワシ類の衝突が、海岸線の崖の風力発電施設で多いことが明らかにされている。その原因として、海岸線の崖は吹きつける風によって上昇風が生じ、それを利用して海ワシ類が飛行するため、飛行頻度が高いことがあげられる。しかし、平成23年度の調査により、海ワシ類が風力発電施設を回避していることが明らかになっており、飛行頻度が高くても必ずしも衝突が生じるわけではない。考えられる衝突の理由の1つとして海岸線に打ち上げられる海ワシ類の餌となる魚類や海獣類などが、海ワシ類を誘引するとともに、海ワシ類が餌に気をとられてしまうことにより、風力発電施設の危険を認識することができず、衝突が生じる可能性が考えられる。一昨年度および昨年度に道東の根室で行った調査から、海岸に餌がある場合とない場合では餌がある場合にワシが下を向いて飛ぶ頻度が上がることが示された。今年度は日本海側の小平で同様の調査を実施し、その反応に地域差があるかどうか検証を行った。

[調査方法]

調査は留萌の北に位置する小平で行った（図3-1-19）。平成28年1月15日から20日および2月8日から11日に実施した。この2回の調査のうち3日間を餌がない状態での海ワシ類の行動を記録し、3日間を餌のある状況での海ワシ類の行動を記録した。餌は小さなものでは海ワシ類が気づく前にハシブトガラス等に食べられてしまったり、ワシ類に簡単に持ち去られてしまうため、カジカを設置した（図3-1-20）。



図3-1-19 調査地点図（小平）



図 3-1- 20 餌として設置したカジカと集まったオジロワシ, カラス, キツネ



図 3-1- 21 撮影したビデオ映像のサンプル

(左: 正面を向いて飛ぶオジロワシ, 右: 下を向いて飛ぶオジロワシ)

餌の有無にともなう海ワシ類の行動の違いは、ビデオ撮影をすることにより記録した。撮影したビデオを後日再生し、海ワシ類の頭が向いている方向（前方あるいは下：図 3-1- 21）別の時間をストップウォッチで記録した。そして、20 秒以上の記録がとれたものを対象に、下を向いて飛んでいる時間の割合を計算した。そして、オジロワシとオオワシの種別に餌の有無による違いがあるかどうかを検討した。

[結果および考察]

餌のない状態とある状態でそれぞれ6日ずつの調査を実施した。ワシの飛翔頻度には、風の影響が大きく影響し、風が崖に吹きつける日には、多くのワシが飛行し、風の弱い日や逆方向から吹いている日にはあまり記録できなかった。この傾向は過年度の調査結果と同じだった。20 秒以上海ワシ類が見ている方向を把握できたのは餌のない状態のオジロワシ 22 例、オオワシ 10 例、餌のある状態でオジロワシ 39 例、オオワシ 23 例だった。

これらの記録について下を見て飛んでいる時間の割合を餌のない場合とある場合で比較すると、根室で行った 2014 年、2015 年と同様に餌のある場合の方が、より長く下を見て飛んでいた（図 3-1- 22、オジロワシ：U=671, P<0.001, オオワシ：U=206.5, P<0.001）海ワシ類は多くの場合は下を見たり、正面を見たりしながら飛んでいたが、下を向いたまま長距離飛ぶこともあった。このような場合は餌の存在により海ワシ類が風車に気づかず、バードストライクを誘発する危険性が考えられる。

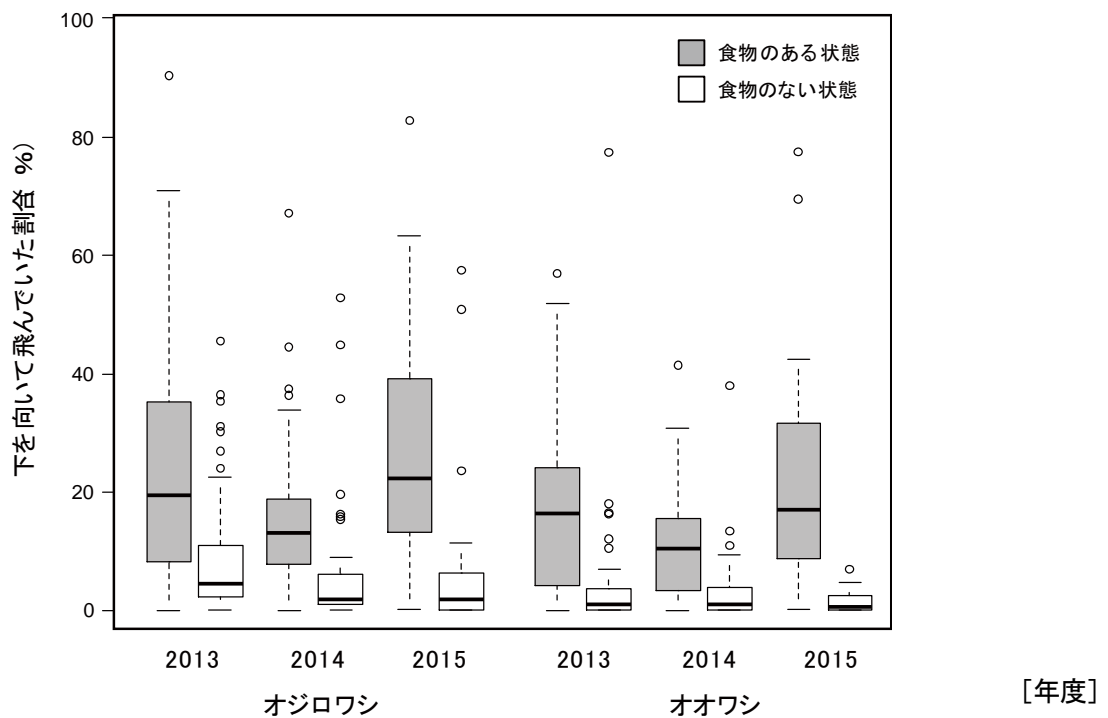


図 3-1- 22 餌がある場合とない場合でのオジロワシとオオワシが下を向いて飛んでいる時間の割合の違い

平成 25 年度の本事業で撮影することのできたバードストライクの瞬間の映像から、海ワシ類が風車に衝突する原因として、飛行中の他個体との干渉が考えられた。そこで、餌のある場合とない場合で、他個体からの干渉の割合に違いがあるかどうかを検討したところ、2014 年のオオワシを除き、餌がない場合とくらべ、餌のある場合に、カラス類あるいは海ワシ類からの飛行中の干渉が有意に多く観察された（図 3-1-23）。したがって、餌の存在は、飛行時の視線方向だけでなく、他個体との干渉をもととして、海ワシ類の風力発電施設への衝突の危険性を高めると考えられる。

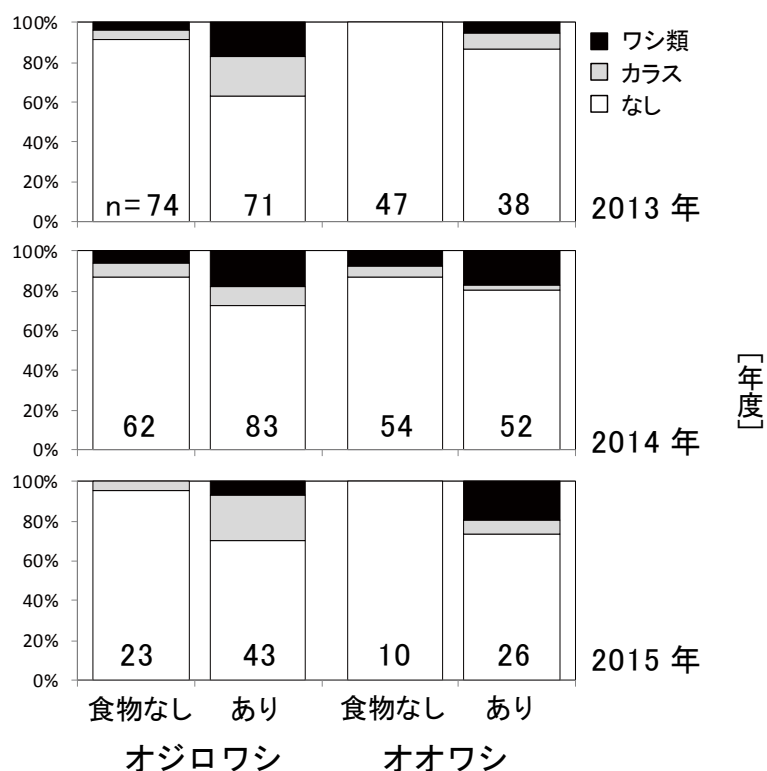


図 3-1- 23 餌がある場合とない場合でのオジロワシとオオワシの他個体との干渉が生じる割合の比較

3.1.3 衝突感知センサの開発・検証調査

衝突を感知するセンサの開発・検証を行う。当該センサは、平成 19～21 年度業務において先行開発されたものがあり、これは風力発電施設基礎部に設置するものであった。

今回は、その成果に基づき、より衝突の感度が上がるブレードに接着するものを開発し、検証を行うこととした。

表 3-1- 1 衝突感知センサの年度別作業項目

| 年度 | 内容 | 結果 |
|-------------|--|---|
| 平成 25 年度 | 感知センサの設計及び試作品の作成 | 試作品の完成 |
| 平成 26～27 年度 | 試作品をブレードに設置し、動作試験を実施する。動作試験の終了時には、試作品を回収し、検証結果を総合的に検討し当該センサを完成させる。 | 【平成 26 年度】 動作試験を行ったところ、感知に問題があることが確認された。 【平成 27 年度】 春、夏、秋の確認試験を経て、冬季の動作試験を行った。 |

今年度について衝突感知センサの部分的改修の方針で進めることとした。

- ①衝突音を検知する方法＝サーフェスマイクロホンの検討
- ②地上に置かれた風車ブレードを利用した打撃実験（信号検出の有無確認）
- ③苫前風車に設置し常時観測、風車運転時の現象と傾向の把握
- ④苫前風車に衝突感知センサを設置し、回転する風車へのソフトボール打撃実験
- ⑤事前実験結果を踏まえ、衝突感知システムとして仕上げた最終形態の機材設置、冬本番観測



図 3-1- 24 サーフェスマイクロフォン

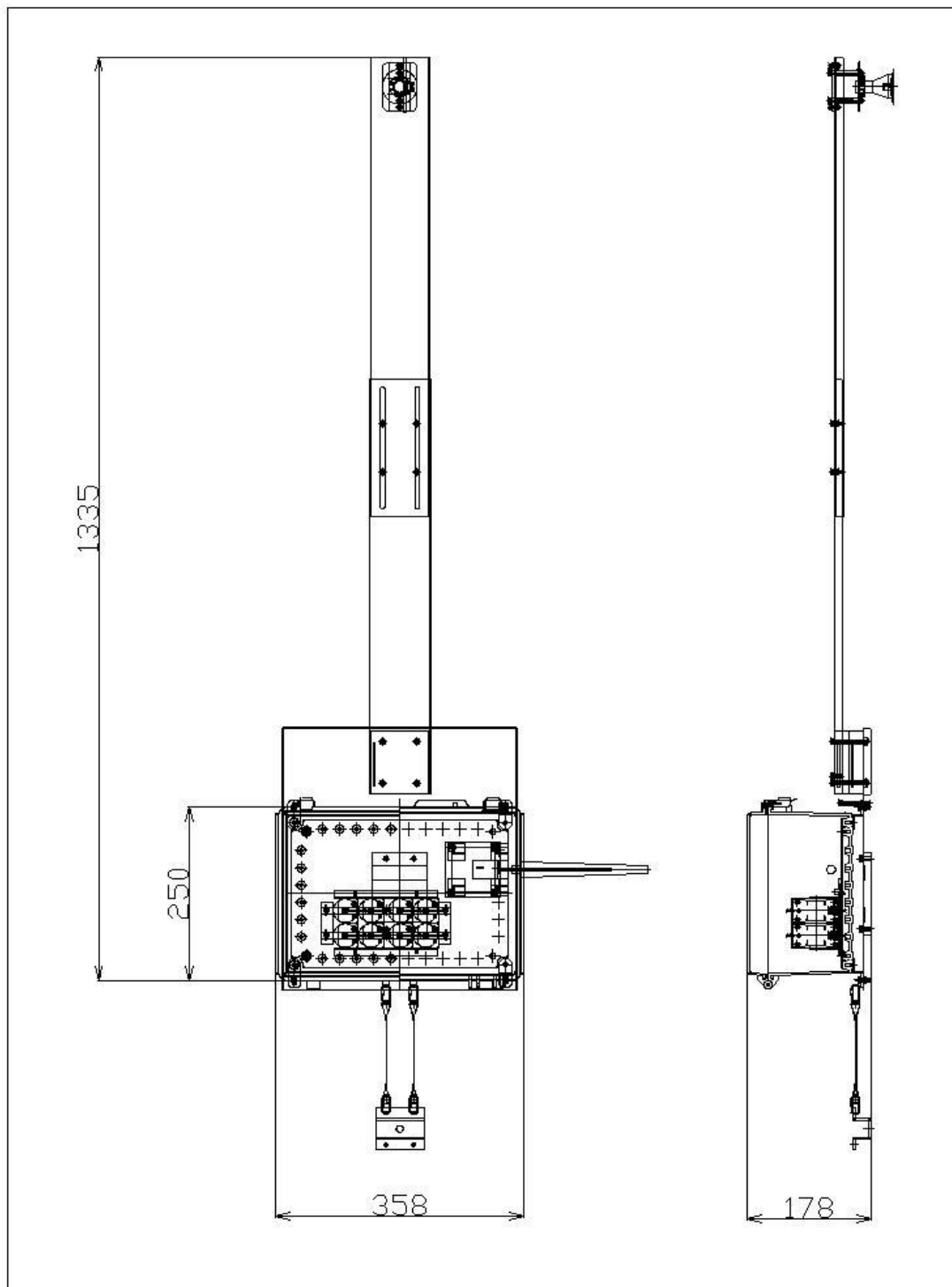


図 3-1- 25 衝突感知センサ 外観図 (改修版)

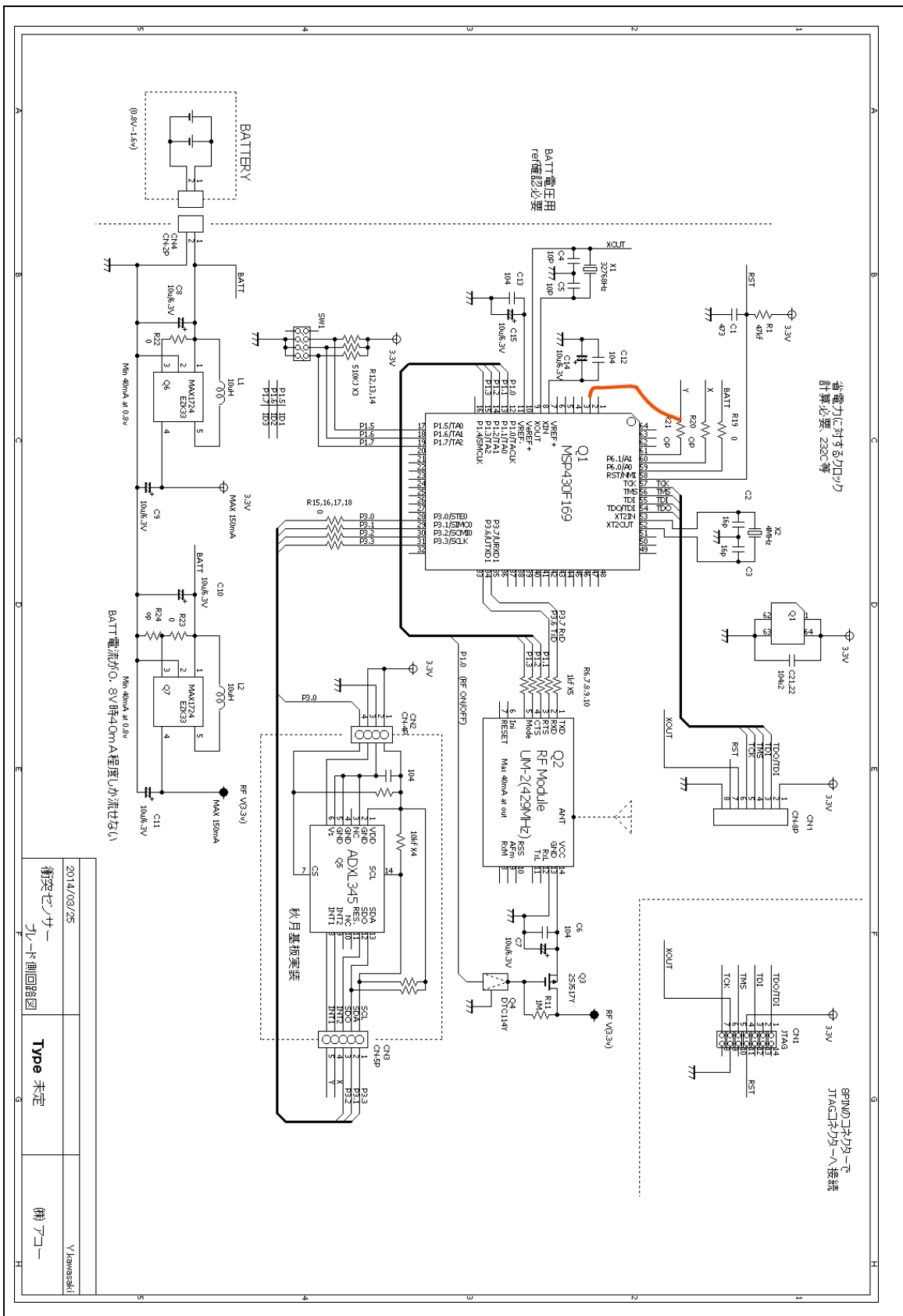


図 3-1- 26 衝突感知センサ回路図 (その 1)

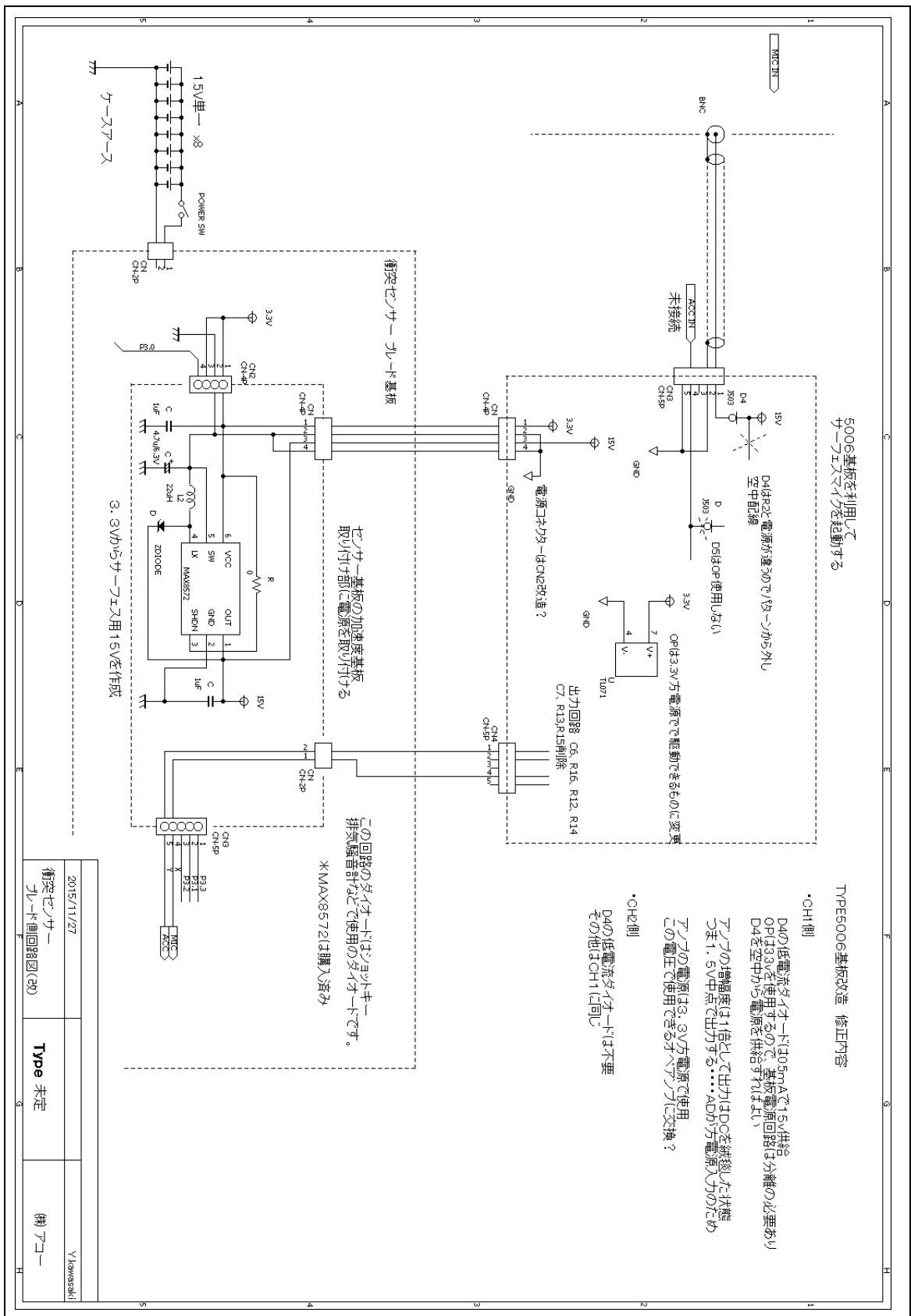


図 3-1- 27 衝突感知センサー 回路図 (その 2)

[地上に置かれたブレードにて打撃実験]

平成 27 年 6 月 2 日に実施した (図 3-1- 28)。

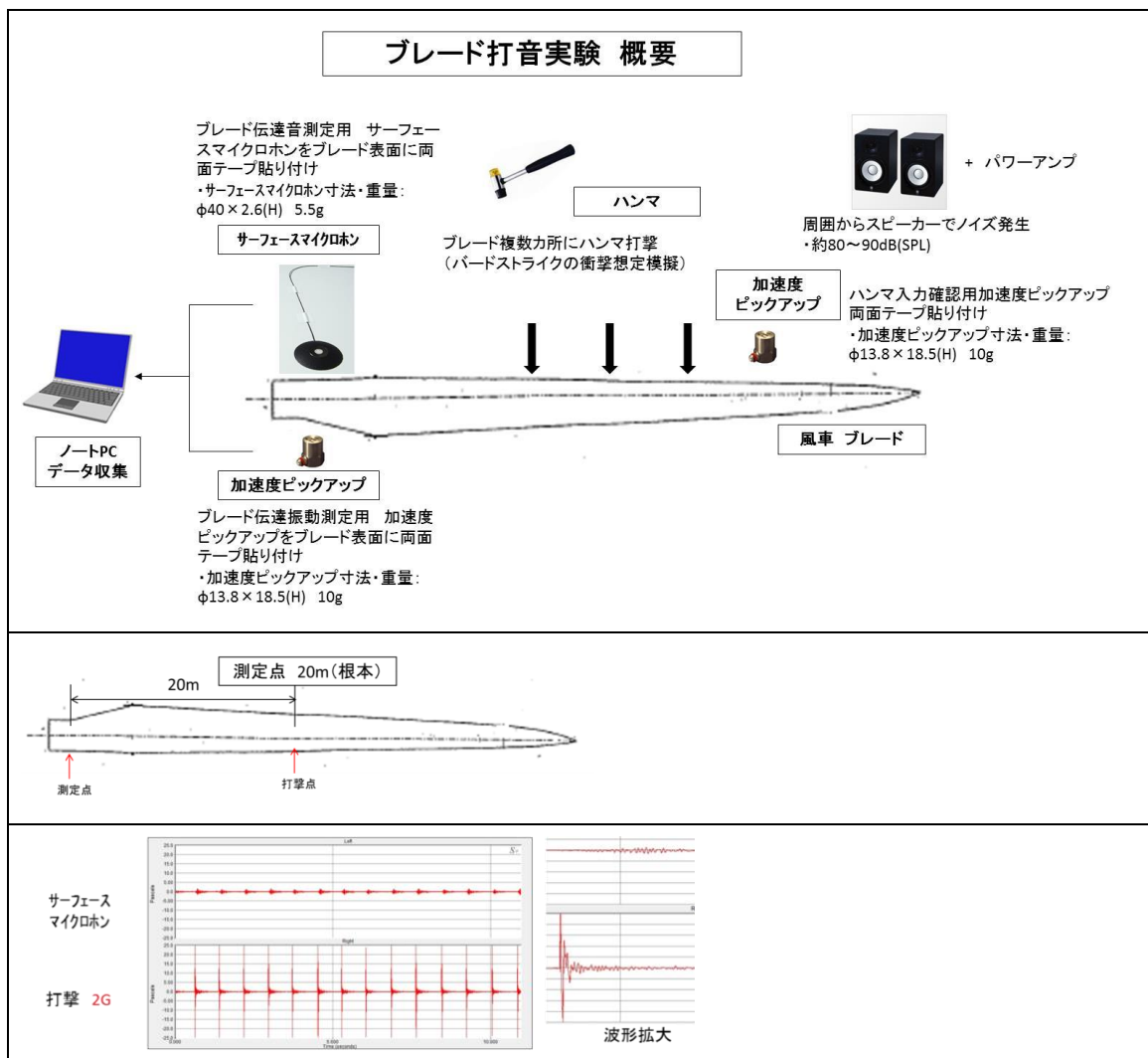


図 3-1- 28 ブレード打音実験の概要

実験結果

- ・ 打撃 1G 程度でも、打撃点から 20m 離れた位置で検出できており、ノイズとの切り分けを考慮しながら利用することで現地での運用に生かせるものと考えられる。
- ・ 暗騒音や暗振動の影響が実際のどの程度となるか、現地で予定している風車ブレード回転の挙動観測データを活用し、加速度センサ、サーフェース・マイクロホン等どちらのセンサが適しているかなど含め検討する必要がある。

[苫前町風車にセンサを装着して、回転する風車装着時のデータ観測]

平成 27 年 7 月 17 日に、苫前町の風力発電施設に、「振動センサ (X, Y, Z 方向)」と「サーフェスマイクロホン」を設置して、回転する風車にセンサを装着したときのデータ記録実験を開始した (図 3-1- 29)。



図 3-1- 29 振動センサ、サーフェスマイクロホンの設置状況

実験の結果、振動センサについては、ブレード回転による目立った信号検出は見られず、X, Y, Z 方向に差異はなかった。なお、設置時にブレードを手でたたいた際には信号検出を確認済みである。

サーフェスマイクロホンについても、風雑音や機械的な音、遠くの車の音が聞こえるなど、音信号の検出可能と考えられた。

【平成 27 年度冬本番に向けて】

振動センサについては、風雑音の影響を受けないことから振動センサを併用する計画。
X,Y,Z 方向に差異が見られないことから、Z 方向の利用のみで検討進めることとした。X,Y
がないことで消費電力軽減にも寄与するものと考えられる。

サーフェスマイクロホンについては、フィルタの活用方法など、本番に向けての検出
アルゴリズムを検討していくこととした。



図 3-1- 30 風車への装着実験で得られた波形データ

[センサ装着、回転する風車に向けピッチングマシンによるソフトボール衝突実験]

平成 27 年 10 月 29 日に苫前町風車のブレードにセンサを装着し、ソフトボールによる疑似衝突実験を実施した（図 3-1- 31）。



図 3-1- 31 ソフトボールによる衝突実験状況

実験の結果、ピッチングマシンからブレードにソフトボールを 3 回衝突させることができた。このうち最初の 2 回分は録音データを再生したところ、耳で聞いて確認することができた。3 回目は確認することができなかった。

時間波形データのうち、フィルター処理前とフィルター処理後を図 3-1- 32 に示す。

音を耳で聞けば確認することができるが、風雑音（暗騒音 110dB～120dB、風速常時約 10m/s 前後）が大きく、単純に波形の大小だけでは打撃部分を判別することはできないかった。ハイパスフィルタ 300Hz をかけてみたところ、風雑音による大きな波形は除去できるものの、衝突波形とそれ以外の波形の切り分けに検討を要することが判明した。

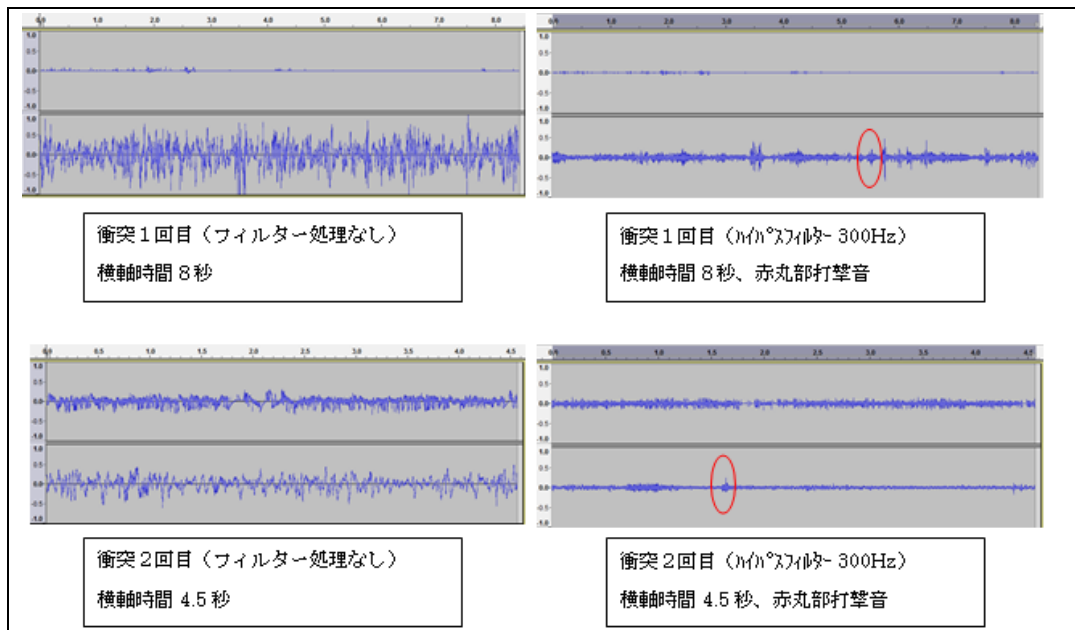



図 3-1- 32 ソフトボール衝突実験で得られた波形
(各波形グラフの上段は振動Z方向、下段はマイクロホン)

[冬本番 衝突感知センサ装着、バードストライクの有無検知モニタリング]


検討の結果、衝突感知センサは「音」を検出するマイクロホンを採用。2015 年夏～秋にかけての現地実験の結果、振動データは衝突実験時の信号を明確に認識することが困難であった。

これに比べ衝突音データは、フィルタによる処理を必要とするものの、衝突実験時の信号が風車ブレード動作時のノイズと区分け可能なレベルにあると判断した。検出アルゴリズム等を図 3-1- 33～図 3-1- 35 に示す。

・衝突感知センサ 装着



装着状態 (ブレード3本とも同一に設置)



無線受信ユニット



レコーダユニット

| 名前 | 更新日時 | 種類 | サイズ |
|-----------------|------------------|------|------|
| A20151214094207 | 2015/12/25 12:05 | ファイル | 5 KB |
| M00151214079308 | 2015/12/25 11:56 | ファイル | 6 KB |

| ファイル名 | 種類 | 書式 | 書式ID | ヘルプ |
|---|----|----|------|-----|
| 1, 20151225095032, 1921, A0102, M0105, JK | | | | |
| 2, 20151225095032, 1929, A0108, M0211, JK | | | | |
| 3, 20151225095032, 1978, A0107, M0095, JK | | | | |
| 1, 20151225195032, 1847, A0100, M0504, JK | | | | |
| 2, 20151225195032, 1851, A0107, M0441, JK | | | | |
| 4, 20151225195032, 1887, A0107, M0054, JK | | | | |
| 1, 20151225195032, 1841, A0099, M0489, JK | | | | |
| 2, 20151225195032, 1845, A0105, M0111, JK | | | | |
| 3, 20151225195032, 1877, A0101, M0279, JK | | | | |

メンテナンス情報ファイルとアラーム記録ファイルはSDカードに記録 (左:メンテナンス情報ファイル、右:アラームファイル、調整時記録)

図 3-1- 33 衝突感知センサによる冬季試験の状況

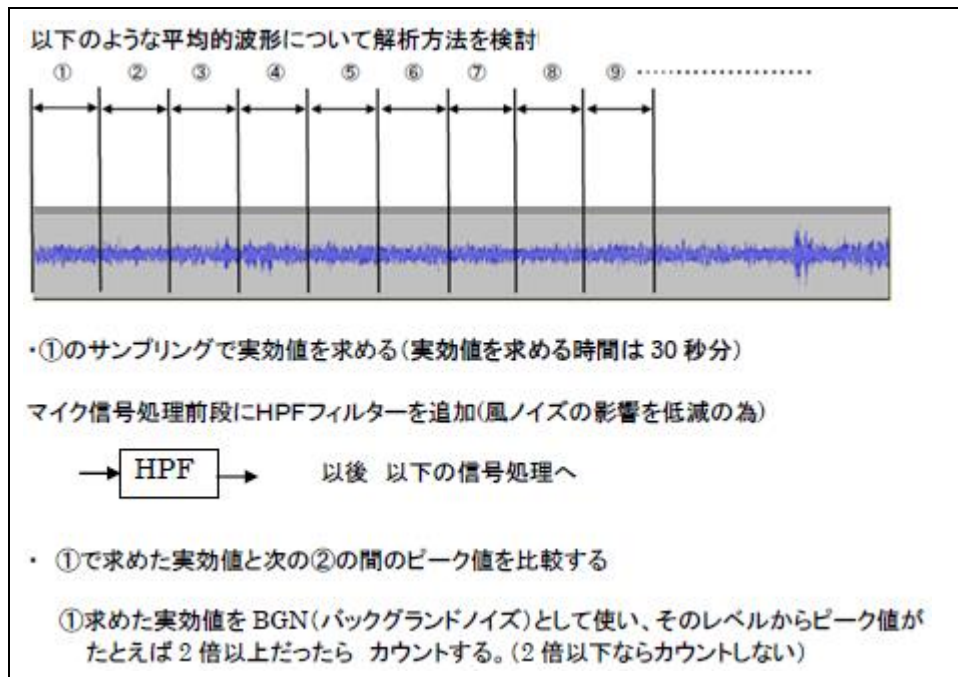


図 3-1- 34 検出アルゴリズム (1)

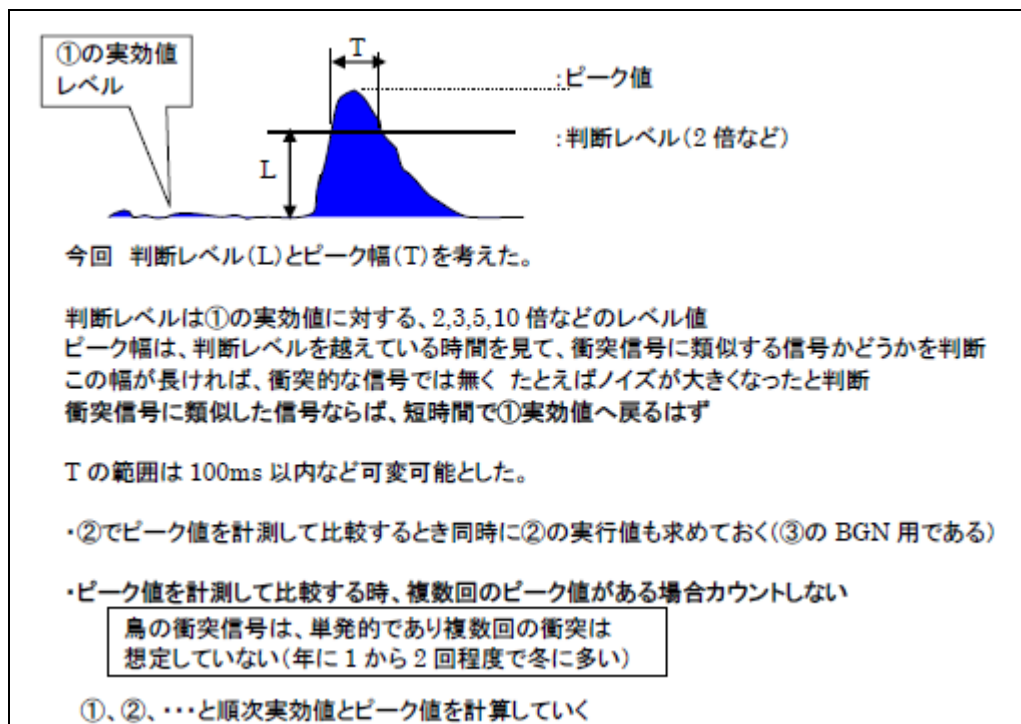


図 3-1- 35 検出アルゴリズム (2)

・装着状況補足

観測期間は、2015年12月25日～2016年2月29日とした。2015年12月25日の設置時、ブレード 3 のみ、センサの設置状態の影響なのか他のセンサとは状況が違っており風車通常回転時で若干アラーム信号が出やすい状態になっているが、現状可能な範囲で判定閾値の設定を実施した。2015年12月25日の設置後、約1～2週間後に一つのブレードのセンサが脱落していることが判明した。レコーダ回収データを下記に示す。



図 3-1- 36 レコーダ回収データの一部

実験の結果、メンテナンスデータより、2015年12月25日～2016年2月29日の期間で最終的に作動していたセンサは3個中No1 だけであった。(1月10日 No2 脱落、1月14日 No3 電池切れ)

アラーム情報データより、頻繁に記録がされていることから閾値のこともあるが、センサ設置状態がよくなかったことが考えられた。これは冬季環境に対するセンサの耐久性の無さが原因と考える。事前実験の夏、秋とは想定以上に条件が大きくかけ離れていたことに原因があると思われる。

[改善案]

1) 判明した点とその改善案

- ①回転する風車へのソフトボール打撃実験の際、ソフトボール打撃信号を同時設置した振動センサでは検知できず、サーフェスマイクロホンでは取得できた事実から、使用するセンサはやはりサーフェスマイクロホンが効果的と考えられる。
- ②風の影響が想定以上に大きかった。

事前実験の約1ヶ月にわたる常時観測の時も風の影響はあったが、冬本番はそれ以上と想定される。

冬本番現地設置時の状況では、アラームレベルをバックグラウンドノイズ（ブレード通常回転時の音）の3倍、ブレード3に至っては20倍に設定しないとアラームが出続けてしまう状況であったので下記の設定としたが、回収した観測データではこれを上回るレベルが頻繁に検出され、アラーム信号が出続けてしまった。

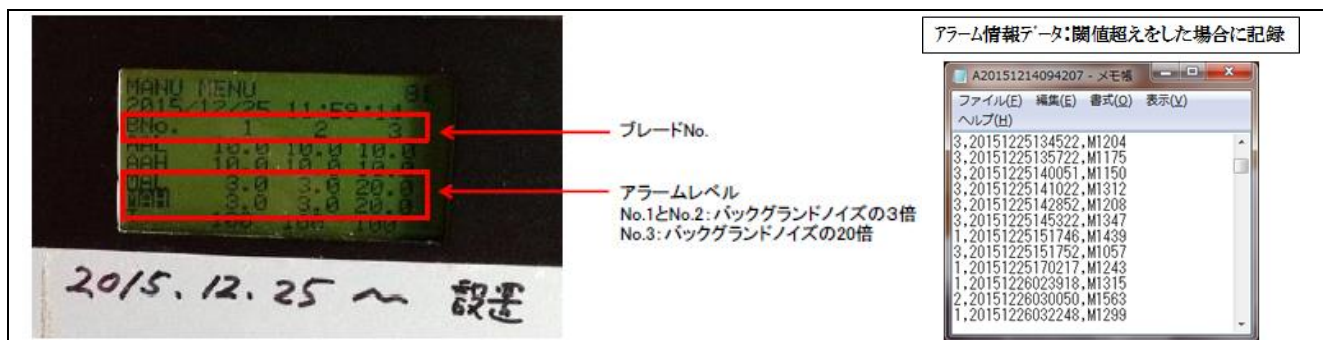


図 3-1- 37 アラームレベル設定状況

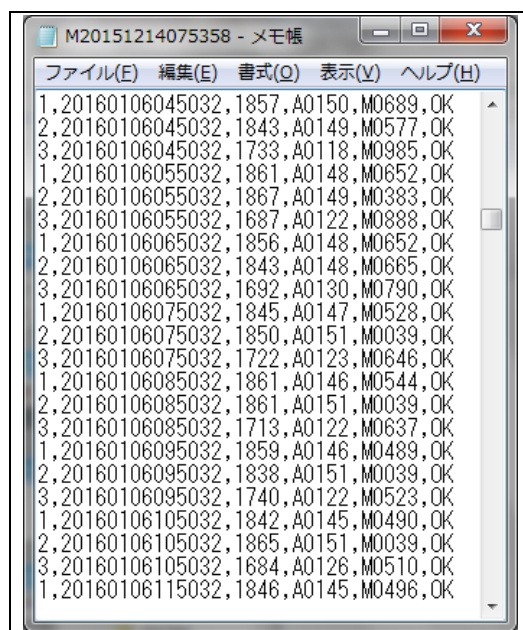


図 3-1- 38 メンテナンス情報データ

メンテナンス情報データについてみると、バックグラウンドノイズ（ブレード通常回転時の音）相当のデータとなる。概略 500 前後のレベル数値で推移しているが、突発的な大きな信号も連続的に多々発生しており、これがアラーム記録につながったものと思われる。

③取り付け方法について、破損という状況が発生したので、想定以上の力が加わったと考えられる。

2) 部分的改善案

①センサに対する想定以上の風影響への対策

各ブレード毎に、もう一台サーフェスマイクロホンを取り付け、一台はセンサ受感部を外側（ブレード面の反対側）に向け、もう一台はセンサ受感部を内側（ブレード面）に向けてセットする。このことにより、風の影響を外側向きのサーフェスマイクロホンで大きく受けることになるので、内側向きのサーフェスマイクロホン信号を差し引くことで風の影響を把握し、衝突信号を計測する手法が考えられる。



図 3-1- 39 センサ受感部

(写真は外側向き、観測ではこの面を下向きにセットして利用した。)

②ケース自体を現在よりもコンパクトにし、ブレード根本の金属部分に取り付けるなど、強風影響を受けにくい構造にする。

3) その他改善案

部分的改善案からは逸脱するが：

- ③電源をバッテリー方式から外部より供給（例えばソーラーなど）出来る方法にする事で、長期にわたり音・振動に限らず色々な信号の検知方法の模索が可能となる。
- ④上記③が実現すれば、小型ドライブレコーダーの様に衝突画像を記録出来ると、後日、衝突信号等と合わせ多角的な状況確認が可能となる。

3.1.4 監視システムの開発・検証調査

[概要]

風力発電が再生可能エネルギーの一つとして注目され始めるとともに、風車と野鳥の衝突、バードストライクが無視できない生態系へのリスクであることが明らかになってきた。バードストライクを回避できる適切なシステム開発が求められている。

従来、野鳥生態、挙動を把握するためには人力に依る部分が大きく、風力発電設備に対する回避システム構築には技術的な解決を模索する必要がある。そこで東京大学では定点デジタルカメラによる自動連続撮影と、画像認識による鳥検出を組み合わせる方法を提案してきた。今後海ワシ類のバードストライク回避に対しては、このシステムを種別分類可能なシステムに改良していく必要があり、またシステムの認識精度を評価する手法を確立する必要がある。これまで高解像度デジタルカメラによる自動撮影と画像認識技術を用いた鳥検出機能を有するシステムを開発し、鳥と非鳥(鳥以外の昆虫や飛行機など)の区別については90%を超える信頼性で実現できることを示してきた。さらに、平成26年度までに、種別の判定・学習を行うための基本的な種別分類データベースを構築し、分類性能の高度化を進めてきた。また、同時に動画による分類と軌跡抽出を実現してきた。

本年度は、更なる高度化に向け、低解像度画像に対する修正アルゴリズムの開発と意味領域分割を組み合わせた物体検出を実現し、検出精度の向上、誤検出の低減を実現した。

用語説明

・バウンディングボックス

「表示領域」「判定領域」とも呼ばれる。画像などのオブジェクト全体を含む境界線のこと。表示領域。文字や複雑な画像をオブジェクトとして処理する場合、そのオブジェクトを含む領域をボックスとして扱うことでそのオブジェクトを表現できる高さや幅が容易に割り出せるようになる(出典：<http://www23.atwiki.jp/sevenlives/pages/1045.html>)

・Haar-like または HOG 特徴量

特徴量とは画像認識において画像を分析する際に抽出する特徴のこと。面積や幅、周囲の長さなどがこれにあたる(出典：https://www.omron.co.jp/r_d/dictionary/dic_word_to.html)。

Haar-like 特徴量：計算対象である矩形中の黒色の領域のピクセル値の和の値から白色の領域のピクセル値の和の値を引いた値(出典：http://news.mynavi.jp/series/computer_vision/011/)

HOG 特徴量：画像の局所領域から輝度勾配・輝度強度を取り出す特徴量(出典：<http://www26.atwiki.jp/hirokatsukataoka/pages/19.html>)

・学習アルゴリズム AdaBoost

YoavFreund と RobertSchapire によって考案された[1]機械学習アルゴリズム(出典：<http://ja.wikipedia.org/wiki/AdaBoost>)。1995年に開発された、教師あり機械学習の代表例である。複数の識別器の重み付き多数決によって分類を行うアルゴリズムであり、特徴としては1.貪欲法によるシンプルかつ高速な学習が可能である。2.弱識別器として任意の学習アルゴリズムを取り込むことができる(メタ

アルゴリズムと呼ばれる) といった点がある。

・畳み込みニューラルネットワーク

畳み込みニューラルネットワーク (CNN) [12]は畳みこみとプーリングと呼ばれる演算の繰り返しによって画像からの特徴抽出と分類を同時に行う手法である。ここでは手書き文字認識[12]のために導入された最初のネットワーク (以下 LeNet と呼ぶ) にいくつかの改変を加えたうえで用いた。

・Fine-grained (な分類)

特徴にほとんど差がない場合の分類手法を指す (出典: <http://hirokatsu16.blog.fc2.com/blog-entry-93.html>)

・セグメンテーション

画像処理によって画像の内容を解析し認識するためには、まず対象物を抽出しなければならない。このための処理は一般に画像のセグメンテーション (imagesegmentation) とよばれる。セグメンテーションは、画像の特徴抽出における最も基本的な処理の一つであり、“一つの対象物に対応した領域では、画像特徴 (明るさ、色、テクスチャなど) またはそれらから推定されるシーン特徴が一様で、異なる対象物の境界部分では特徴が急変する” という仮定の下で、様々なアルゴリズムが考案されている (出典: <http://www.sice.jp/handbook/%E9%A0%98%E5%9F%9F%E5%88%86%E5%89%B2>)

・ステレオマッチング

2次元の画像データから3次元の情報を得る」手法のこと (出典: <http://www.pasco.co.jp/recommend/word/word045/>)

・ロバスト

ロバストネスまたはロバスト性とは、ある系が応力や環境の変化といった外乱の影響によって変化することを阻止する内的な仕組み、または性質のこと。ロバストネスを持つような設計をロバスト設計、ロバストネスを最適化することをロバスト最適化という。

「頑強な」という意味の形容詞“robust”が語源であり、他に頑強性、強靱性、堅牢性、強さ、などと呼称されることもある。(出典: <http://ja.wikipedia.org/wiki/%E3%83%AD%E3%83%90%E3%82%B9%E3%83%88%E3%83%8D%E3%82%B9>)

・クラウドソーシング

狭義では不特定多数の人に業務を委託するという新しい雇用形態。ウェブサービスのトレンドの一つでもある。群衆 (crowd) と業務委託 (sourcing) を組み合わせた造語。

広義では、必ずしも雇用関係を必要とせず、不特定多数の人間により共同で進められるプロジェクト全般を指す場合もあり、その場合オープンストリートマップ (OSM) などが代表例として挙げられる。(出典: <http://ja.wikipedia.org/wiki/%E3%82%AF%E3%83%A9%E3%82%A6%E3%83%89%E3%82%BD%E3%83%BC%E3%82%B7%E3%83%B3%E3%82%B0>)

[低解像度画像による鳥検出・分類]

ここでは深層学習・画像認識手法の評価を行い、低解像度での野鳥の検出・分類について述べる。近年、風力発電所が増設されると同時に、バードストライクの問題が懸念されるようになってきている [1-3]。バードストライク問題に対しては、過去当事業で開発を進めてきたカメラシステム同様、鳥の検出を自動的に行う装置が製品化されつつある。このようなシステムは、カメラやマイクロフォン、レーダーのようなセンサの情報から風車周辺の野鳥の位置を検知し、その後必要に応じてログの記録、警報音の発生などを行う。また野鳥の衝突が予測される場合の風車の減速、停止も可能であ

るとされている[4]。スペインで開発された DTBird[4]、オランダで開発された WTBird[5]などはそのようなシステムの例であり、実際の風力発電所での稼働実績がある。レーダーやサーモカメラによる鳥検出[6-8]は古くから利用された例があり、それぞれ対象までの距離が取得できる、夜間も利用できるといった利点があるが、どちらも利用できる情報が限られ、発見された物体が本当に鳥だったかどうか、また何の種類の鳥だったかなどを推定することは難しくなる。一方で RGB カメラを利用すれば、対象の色などの豊富な視覚的情報を高解像度で利用できるシステムが比較的安価に利用できる。カメラベースの鳥監視システムのボトルネックとなりうるのは、画像を効果的に処理するソフトウェアである。画像情報の実利用はかつて困難な課題であったが、これは近年の技術革新により克服されつつある。カメラベースの鳥監視システムを考える上で重要になる技術は、この 10 年で大きな発展を遂げた画像認識である。近年の画像認識システムは、応用によりさまざまな形態があるが、共通したポイントとしては機械学習を利用する点、画像に適切な前処理を施す点、リアルタイムかそれに近い効率的な処理を可能にする点などがある。まず機械学習における近年の最も大きなイノベーションは、深層学習、またはディープラーニングの発明である。深層学習では、データを効率よく表現するような階層的な構造がニューラルネットワークによって自動的に獲得される。これは画像データの表現と特徴抽出に苦しんでいた画像認識の諸領域を大きく進展させることにつながった。

しかしながら、風力発電所での鳥認識を実用的な環境において評価できるようなベンチマークが存在しなかった、また野鳥の種類を分類するシステムがなかったという点がある。画像認識分野ではすでに鳥画像のデータセット[9-10]が詳細画像認識のためのベンチマークとして利用されているが、Web 上で収集できる画像を利用しており、特定の地点での野鳥の監視という課題からは離れたものである。このため、一般物体認識について人間と同程度の精度を達成したという最新の手法も、野鳥認識においてどの程度に効果的か不明である。そこで本事業では、風力発電所での野鳥監視（海ワシの抽出）という課題に向けた最初の実用的なデータセットを構築した。このデータセットの特徴としては、1 日につき 10、815 枚という多くの画像を含む点、画像全体は高解像度（5616×3744 ピクセル）であるにもかかわらず、検出対象である鳥は低い解像度（平均 20 ピクセル四方ほど）でしか映っていない点、階層的なラベリングにより検出から詳細画像認識までのさまざまな粒度の認識課題をカバーする点、固定カメラによるほぼ一定のシーンのデータが長時間収録されている点などである。

[野鳥検出・分類システム]

野鳥検出・分類システムの概要を示す。このシステムは大きく分けて、自動撮影装置とコンピュータ上での画像処理パイプラインからなる。画像処理パイプラインは候補領域の抽出、野鳥の検出、種の分類の 3 ステージからなり、背景差分法と機械学習にもとづく画像分類器を利用する。まず自動撮影装置の構成を示し、のちに画像処理のパイプラインを示す。

[撮影装置]

このシステムではデジタル静止画カメラ (EOS5D MarkII) を用いて画像の撮影を行う。静止画カメラを用いることで、比較的 low 解像度になる動画カメラと比べ、民生品を用いて容易に高解像度の画像取得が可能となる。解像度は 5616×3744 ピクセル、およそ 5K である。さらに EOS5D MarkII に望遠レンズを組み合わせて、遠距離の鳥を撮影可能とする。レンズは EF70-200mmF4LUSM を使い、焦点距離は 200mm、画角は $27^\circ \times 19^\circ$ とした。一定の範囲をカバーするために、風車から 540m 離れた地点に定点を設置し、これにより風車周辺の 1m の大きさの鳥が 20 ピクセルほどで撮影可能である。カメラはコンピュータに接続され、自動的におよそ等間隔での撮影と画像の転送が行われる。接続にはカメラ発売当時の規格である USB2.0 を用いる。この際にはカメラが提供するアプリケーションプログラミングインターフェース (API) を PC から呼び出して利用する。また高速な転送のために、画像はカメラ内部で JPEG 圧縮を行ったうえで送られる。USB2.0 の最大転送速度 (480Mbps) の影響を受け、フレームレートは 0.5fps (2 秒に 1 枚の画像を撮影・転送可能) に限られている。設置時のカメラの様子と、撮影された画像の例を図 3-1-40 に示した。定点カメラではあるが、画像には雲の出現、照明の変化、風による風車の回転といった変化があり得る。

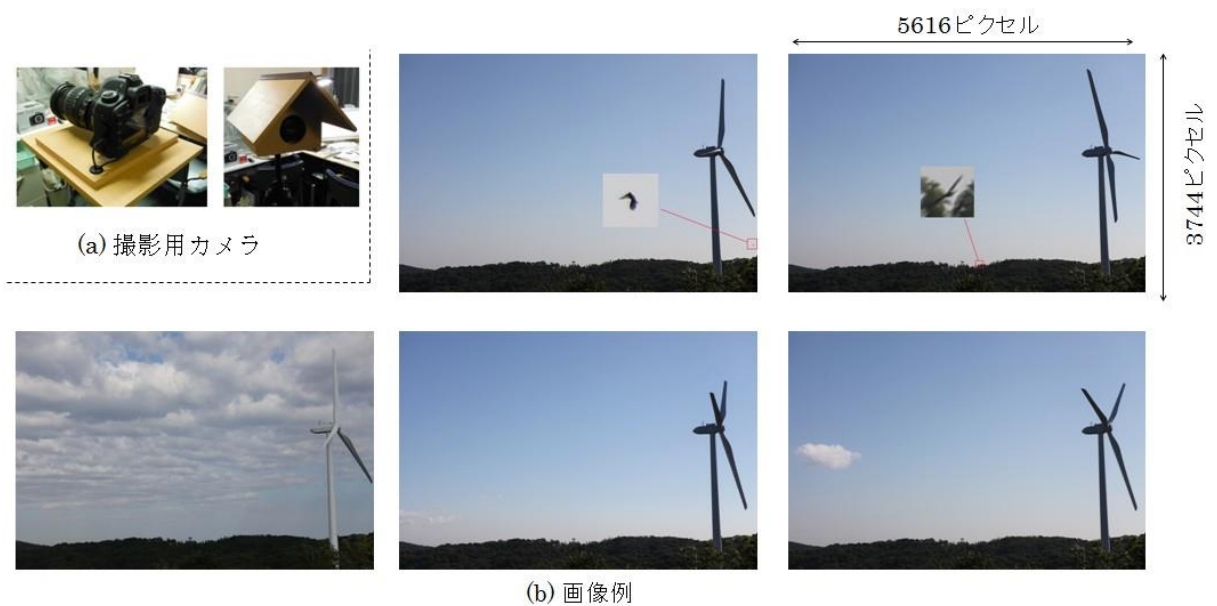


図 3-1-40 撮影装置と画像例

[背景差分法]

画像処理パイプラインでは、まず図 3-1-41 のように背景差分法を用いた鳥の候補領域の抽出を行う。背景差分法とは、背景の色モデルを統計的に構築したのち、現在の色がそれに収まるなら背景、外れるなら前景とする手法である。これにより変化の小さい定点カメラの画像から、効果的な候補領域の生成が可能である。これにより動画像中からの前景抽出、すなわち移動物体の検出が可能である。背景モデルとしては、1 ピクセルごとの輝度値に正規分布性を仮定し、背景を色の平均と標準偏差によって代表する。



図 3-1-41 背景差分法

そのうえで異常度として、この値に閾値をかけて前景マスクを得た。さらにマスク画像にグルーピング、すなわち 4 つの近傍系での単連結な前景領域を一つの物体とみなす処理

を行い、これを鳥の候補領域とした。

[分類器]

背景差分法によって移動物体の抽出が可能ではあるが、それだけでは鳥以外の物体が多く検出されてしまい、また鳥の種類もわからない。そこで背景差分法の後処理として、教師あり機械学習に基づく分類器を鳥の検出・種の分類の両方のために用いる。背景差分法により生成された候補領域を鳥/非鳥に分類することにより鳥の検出、さらに検出された鳥の種の分類が行われる。特に本事業では、AdaBoost[11]と諸々の画像特徴量の組み合わせ、畳み込みニューラルネットワーク(CNN)[12]を用いて分類を行う。

[データセット]

本事業では風力発電所での定点カメラによるタイムラプス画像に基づくデータセットを構築した。画像としては先に述べた撮影装置により取得されたものを利用し、2秒に1回の撮影で1日に8時間・10815枚、3日分の画像にアノテーションを施した。ここではこのデータセットの構築手順の詳細と、内容について述べる。

[アノテーション]

このデータセットでは、画像中に含まれる鳥が低解像度・不鮮明となり、詳細に種類を判断することがより難しくなるという問題がある。そこでアノテーション（ラベル、バウンディングボックスといったメタデータを画像に付与すること）は、野鳥の専門家らに依頼することとした。実際に作業にあたったのは、フィールド調査に参加した経験のある野鳥の会の会員ら12名である。

作業効率のために、アノテーション用インターフェースを開発し(図 3-1-42)、この上で作業を行った。

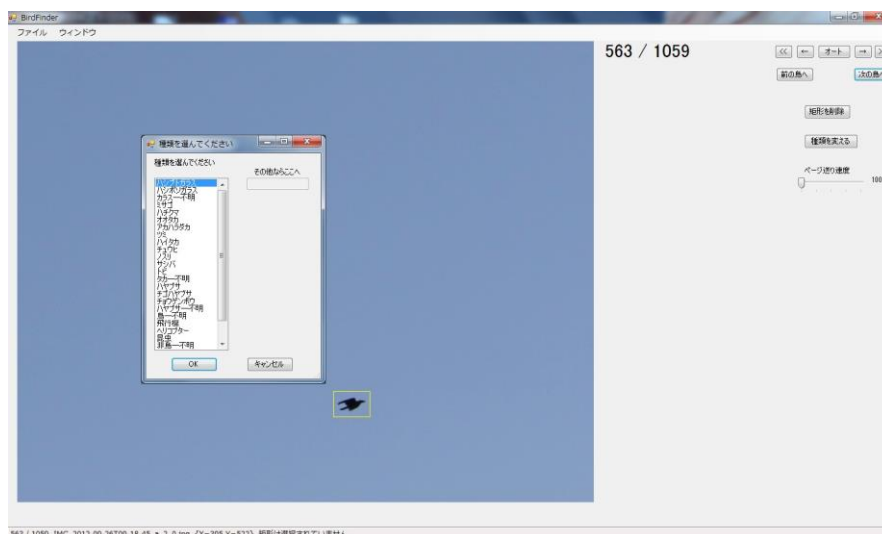


図 3-1-42 画像アノテーションに利用したインターフェース

実際の手順は

- (1) 画像を時系列順にチェックする。
- (2) 鳥や他の飛行物体を発見したら、マウスによるドラッグ・ドロップでバウンディングボックスを指定する。
- (3) ドロップすると種類ラベルを選択するウィンドウが表示されるので、その中からラベルを選ぶ。
- (4) これを時系列の終わりまで繰り返す。

というものとした。加えて画像の巻き戻しも可能とし、判断にあたっては前後の画像を見比べることを可能とした。

出現した物体の大きさ出現した鳥、非鳥領域の大きさの分布を図 3-1-43 に示す。鳥領域の分布のピークは想定通り 20-25 ピクセルに位置した。一方で、マークされた非鳥領域もおおよそ鳥と同様の分布を持っており、見た目の大きさのみでこれらを識別することはできないことも明らかとなった。大きさの分布は、カメラの視野に物体が進入する頻度と、それを人間が発見できる割合によって決まっていると考えられる。つまり、大きく見える物体はカメラの近くに存在し、カメラの視野が狭いため映る頻度が小さい。また小さく見える物体はカメラの視野が広く、より映りやすいものの、画像を目視した人間に見えにくいため、20 ピクセル以下の物体は少なくなる結果となったといえる。

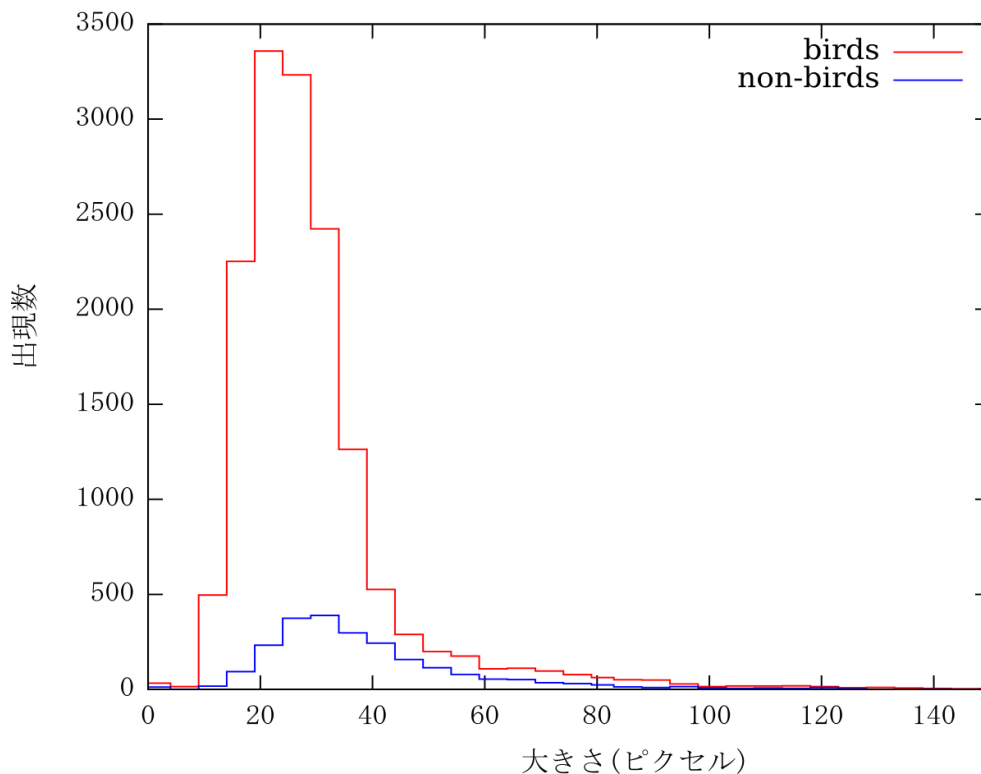


図 3-1-43 鳥・非鳥領域の大きさの分布

カテゴリと画像数

図 3-1-44 に発見された物体のカテゴリと、各カテゴリの出現数を示した。カテゴリは階層的に設計されており、例えば、鳥であると判明したが種は判明しなかったもの (undefined bird)、タカ類であるというところまで判明したもの (undefined hawk)、種名まで判明したもの (black kite など) などがある。鳥の画像は 25,000 枚ほど集まっており、学習・評価に十分な枚数あるといえるものの、より詳しいカテゴリでは枚数が減っていく。種まで判明したものでは、国内では最も普通に見られるタカ類であるトビ (blackkite) のみが 2900 枚と多く集まり、他の種では数十枚ほどに限られている。また 9800 枚ほどの鳥画像は種類は判明せず、undefined bird に分類された。また鳥検出の際に誤検出を起こす可能性のある他の飛行物体、飛行機・昆虫などもマークしておき、学習や結果の分析に利用できるようにした。

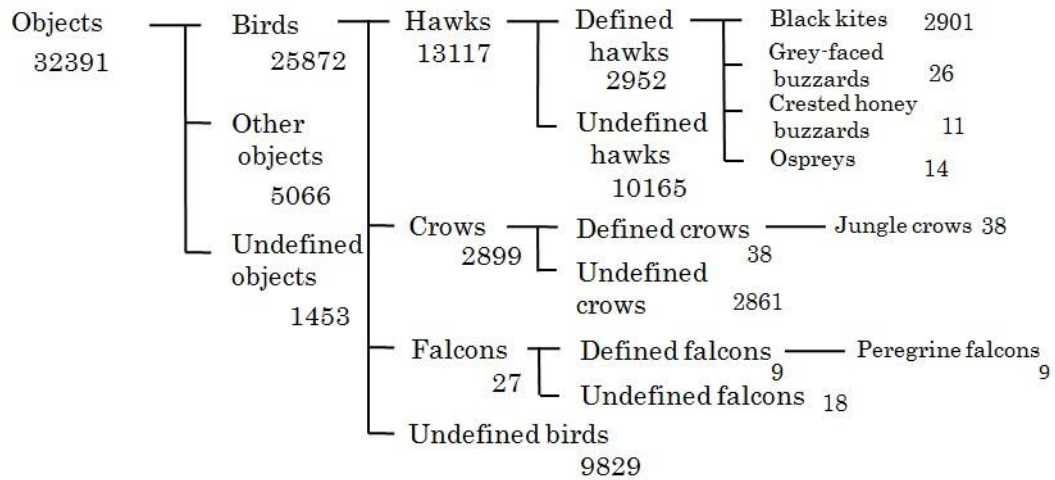


図 3-1-44 カテゴリの階層と画像数

[画像例]

集まった鳥、その他の飛行物体の画像例を図 3-1-45 に示す。タカ類としては(a)トビ、(b)サシバ、(c)チュウヒ、カラス類としては(d)ハシブトガラス、(e)ハシボソガラスの画像例を示した。また(f)メジロといった小鳥類も少数確認された。(g)不明な鳥は、種類を同定する手掛かりとなる部分が不鮮明であるか、映っていないため。分類できなかった。また(h)飛行機、(i)ヘリコプターなどは記録された鳥以外の飛行物体の例であり、視覚的に鳥と類似しているため、システムにより誤検出される懸念がある。また(j)不明な物体は、あまりにも不鮮明なため鳥であるかどうか不明である。

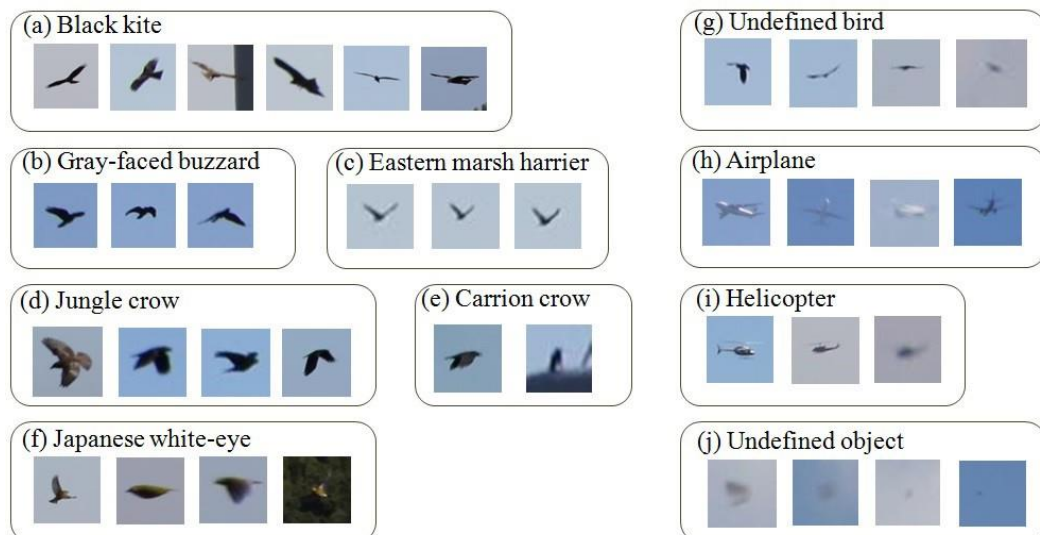


図 3-1-45 データセットに含まれる物体の画像例

[データセットを用いた鳥種別分類]

構築したデータセットを用い、先に述べた画像認識手法を用いて、鳥検出と種類の分類を扱う。

鳥の検出では、2 クラス分類器によって背景差分法で得られた鳥・非鳥領域を分類する。背景差分法のパラメータとしては閾値を 10 とし、1000 フレームごとにその最初の 100 フレームを用いて背景モデルの初期化を行っている。この結果、鳥領域を 8,969 個、非鳥領域を 18,688 個得た。その画像例を図 3-1-46 に示した。これら鳥領域・非鳥領域の 2 クラス分類によって、検出の性能を評価した。上記のデータを用いて 5 分割の交差検証を行った。つまりデータ全体を 5 分割し、そのうち 1 つをテストセットとして残り、他のデータで学習を行い、テストセットでの検出結果を評価した。検出結果は偽陽性率 (false positive rate、以下 FPR) に対して真陽性率 (true positive rate) をプロットした曲線、受信者操作曲線 (receiver operating characteristic curve、以下 ROC 曲線) によって定量化する。FPR と TPR はそれぞれ

$$\text{FPR} = \text{誤検出された負例の数} / \text{すべての負例の数} \quad \dots \text{(式 3-5)}$$

$$\text{TPR} = \text{検出された正例の数} / \text{すべての正例の数} \quad \dots \text{(式 3-6)}$$

として計算した。FPR と TPR はトレードオフの関係にあり、曲線が左上にあるほどよいトレードオフ・高い性能を示す。

種類の分類では、同じく背景差分法で得られた鳥領域の中でのタカ類・カラス類の 2 クラス分類を扱う。学習・評価に十分な数のデータを用意するため、図 3-1-44 のカテゴリのうちタカ類・カラス類を用いた。データは検出に用いた鳥画像のうちから解像度が 15 から 50 ピクセルであるもの、正例としてタカ類の画像 3,036 枚、負例としてカラス類 998 枚とした。タカ類から 2,400 枚、カラス類から 450 枚の画像をランダムに抽出しこれを学習データとし、残りをテストデータとして評価することを 10 回繰り返し、ROC 曲線の平均を計算した。さらに詳細な分析のため、分類の実験では解像度ごとの分類結果も個別に計算した。タカ類、カラス類の画像を 15 から 20、21 から 30、31 から 50 ピクセルの 3 グループに分け、それぞれで同様の実験を行い、各解像度での ROC 曲線を求めた。

評価する手法としては、上記の特徴量 (Haar-like、HOG) と AdaBoost を組み合わせたもの、また特徴量を用いず、画像の RGB 値を直接 AdaBoost に入力したもの (RGB)、CNN の 4 つを用いた。AdaBoost は各特徴量を実数値として受け取ることとし (RealAdaBoost)、またパラメータである弱識別器数の最大値は 800 に固定した。CNN

の学習には確率的勾配法[13]を用い、学習時のパラメータは手書き文字認識[12]と同じとし、加えて全結合層でのドロップアウト率を 0.5 とした。

まず検出の結果の ROC 曲線を図 3-1- 47 に示す。Haar-like により最も高い検出精度が達成され、FPR が 0.02 のときに TPR が 0.995 ほどと、高精度な鳥の検出が可能になっている。Haar-like がこのような条件での鳥検出に適していることは過去本事業でも言及しており、これがより大規模なデータにおいて確かめられた形となっている。一方 CNN は RGB や HOG といった手法より低い精度を示した。

誤検出の画像例を図 3-1- 48 に示した。同図右は誤検出された物体の例であり、風車の一部や木の一部が切り抜き方によって鳥に類似して見える場合、誤検出されるケースがある。また昆虫や飛行機といった飛行物体は鳥に類似しており、誤認識が発生している。

同図左には誤検出が発生した画像の例を示した。誤検出の主な要因はこのように、風車の一部と画面下部の茂みである。このためこの部分を検出対象外とすればより高精度な鳥の検出も可能である。ただし応用上は風車周辺などが監視の重要度が高いため、一概にそうすべきとは言えない。

次に分類の結果の ROC 曲線を図 3-1- 49 に示す。種類の分類は鳥検出より困難な課題であるため、すべての手法が検出のケースより低い性能を示している。また比較的高い精度を示す手法の傾向も変わっており、検出で最も精度が高かった Haar-like は他の手法より顕著に低い精度を示している。逆に CNN は分類においては他の手法と比べ、最も高い性能を示した。出現したカラス類を 0.1 ほど誤認識するときに、0.85 ほどのタカ類を正しく認識することができる。

また解像度ごとの分類結果を図 3-1- 51 に示した。CNN が最も高い精度を示すこと、Haar-like が最も低い性能を示すことは、解像度によって変わっていない。一方で RGB と HOG を比較すると、最も解像度の大きい 31 から 50 ピクセルのグループでのみ HOG が RGB を上回る性能を達成していることが見て取れる。

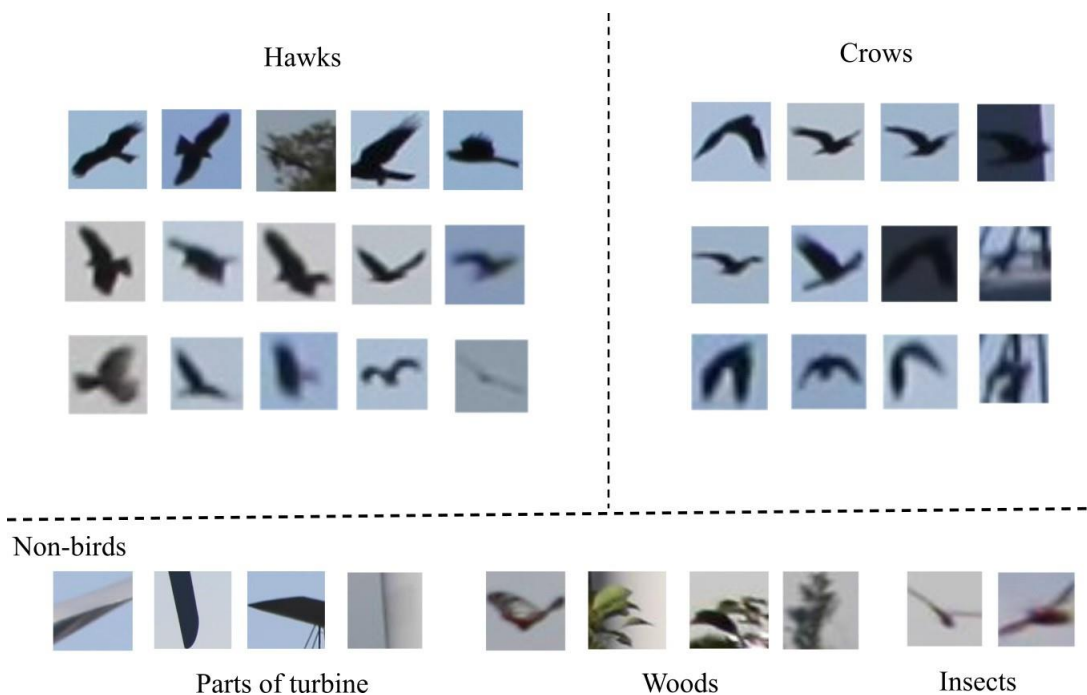


図 3-1- 46 実験に用いる正例・負例の画像例

この結果から、手製の特徴量では手法によって適した解像度があるということが考えられる。HOG の場合では、勾配を劣化させるぼけや低解像度に対して頑健でないといったことありうる。

CNN による分類の結果の画像例を図 3-1- 50 に示した。視覚的には類似した画像がときに正しく認識され、ときに誤認識されていることが見て取れる。CNN は強力な反面、学習の過程が最適化によるブラックボックスであるため、結果の理解が難しい場合がある。いくつかの研究はニューラルネットワークの可視化に努力を払っている[14、15]が、未だにシステムが不可解な出力をするケースがある[16]ことが指摘されている。その点も含め本事業では、意味領域分割を組み合わせた検出精度高度化を試みる。

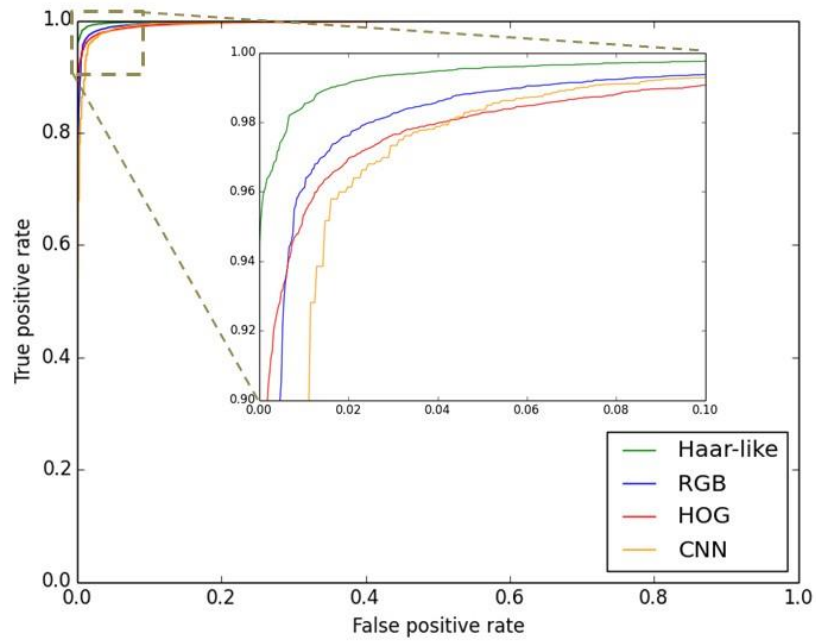


図 3-1- 47 検出結果の ROC 曲線



図 3-1- 48 誤検出の例

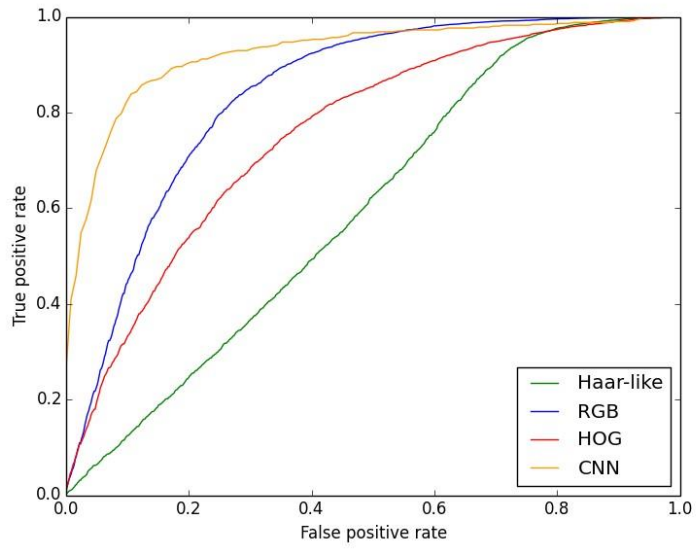


図 3-1- 49 分類結果の ROC 曲線

| True labels Results | | Hawks | | | | | Crows | | | | |
|------------------------|------|-------|--|--|--|--|-------|--|--|--|--|
| | | 20px | | | | | | | | | |
| Correctly classified | 30px | | | | | | | | | | |
| | 50px | | | | | | | | | | |
| Misclassified | 20px | | | | | | | | | | |
| | 30px | | | | | | | | | | |
| | 50px | | | | | | | | | | |

図 3-1- 50 分類結果の画像例

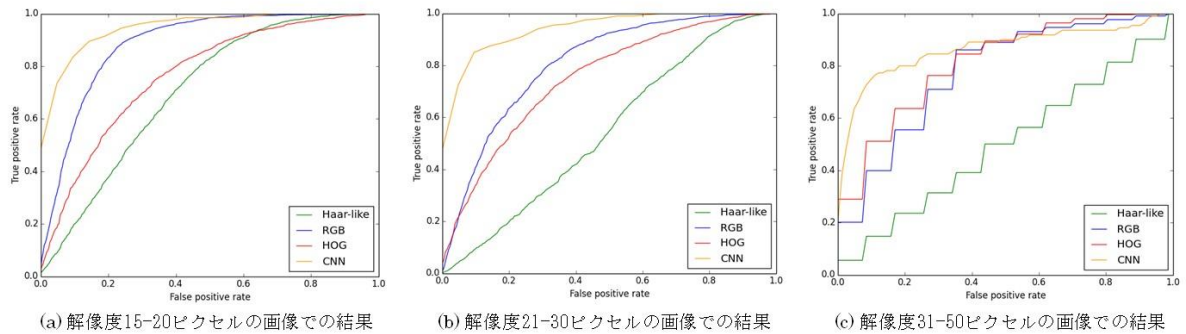


図 3-1- 51 より詳細な分類結果:解像度ごとの分類精度

(より詳細な分類結果:解像度ごとの分類精度)

[意味領域分割の組み合わせによる高度化]

ここでは、先の課題をクリアし、また小物体と背景領域を同時に認識することを可能にするために、CNN を基にした検出器と意味領域分割の手法の 1 つである FCN と Super Parsing を組み合わせる手法を提案する。

提案手法の流れを図 3-1- 52 に示す。まず入力された画像は 3 つの経路でそれぞれ処理が行われる。1 つ目では、背景差分法により物体候補領域を抽出した後、深層特徴を用いた検出器として ResNet を基にした CNN を適用する。2 つ目では、深層特徴を用いた、検出器の意味領域分割のどちらとしても働く手法として FCN を適用する。3 つ目は、従来の画像特徴量を用いた意味領域分割の手法として SuperParsing を適用する。これら 3 つの処理の後に得られたクラス毎の尤度やスコアを SVMs を用いて組み合わせる。最終的に、鳥が検出された領域が出力される。CNN は先に述べた通りである。SVM による各クラスの尤度の結合のため、FCN と Super Parsing は 4 クラス(鳥・空・森・風車)、CNN は 8 クラス (鳥・プレード・タワー・風向風速計・ナセル・ハブ・森・その他) を持つように設計する。Super Parsing と FCN の実装の詳細については別節に記す。各 3 つの経路からは、クラス毎の尤度もしくはスコアが得られる。Super Parsing と FCNs はピクセル毎の各クラスの尤度、一方 CNN はバウンディングボックス毎のスコアが得られる。SVM の訓練の際に画像中の全ピクセルを用いると、実用上計算量が大き過ぎる。そのため、フレーム間差分法により得られた候補領域のバウンディングボックスの中心点のみを訓練に用いる。最初の訓練の後、false positive を減らし全体の性能を向上させるために、誤検出しやすい領域を訓練に加えるハードネガティブ抽出を行う。具体的に述べると、風向風速計・常夜灯・ナセル下部の領域を FCN においてしばしば誤検出していた

ので、SVM の訓練対象に加える。更に、フレーム間差分により得られたピクセルは真の画素分布と統計的に異なるので、CNN と意味領域分割を基にした手法を単純に組み合わせただけでは、全体のフレームワークとしては、CNN による誤検出を多く含んでしまう傾向にある。そのため、候補のバウンディングボックス内の背景領域（空・雲・森・風車）を SVM の訓練対象に加える。

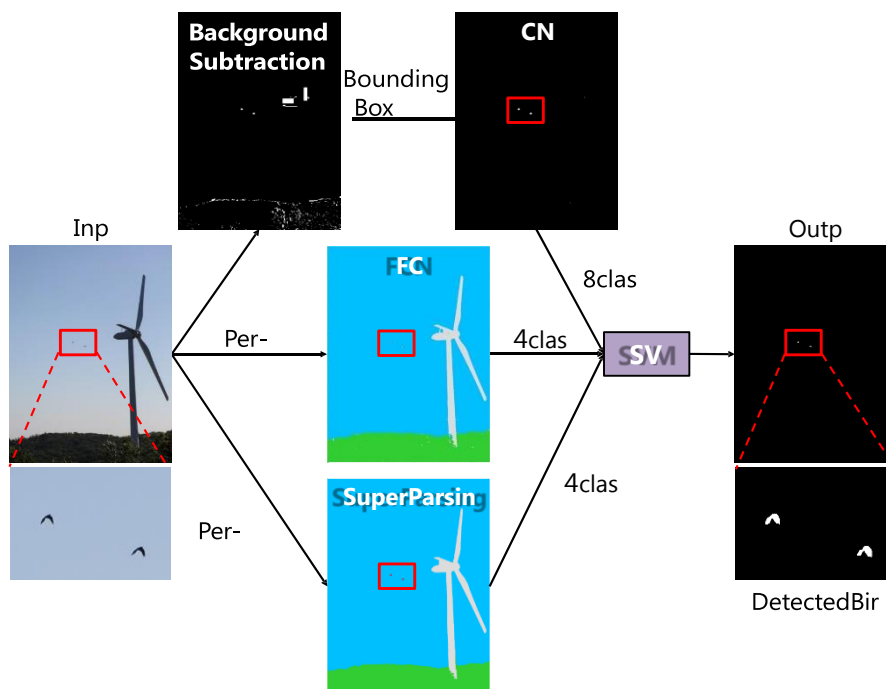


図 3-1- 52 提案手法の流れ

| name | kernelsize | stride | pad | outputsize |
|-----------|------------|--------|-----|------------|
| input | - | - | - | 64×64×3 |
| conv1-0 | 3×3 | 1 | 1 | 64×64×16 |
| conv1-1-1 | 3×3 | 1 | 1 | 64×64×16 |
| conv1-1-2 | 3×3 | 1 | 1 | 64×64×16 |
| conv1-2-1 | 3×3 | 1 | 1 | 64×64×16 |
| conv1-2-2 | 3×3 | 1 | 1 | 64×64×16 |
| conv2-1-1 | 3×3 | 2 | 1 | 32×32×32 |
| conv2-1-2 | 3×3 | 1 | 1 | 32×32×32 |
| conv2-2-1 | 3×3 | 1 | 1 | 32×32×32 |
| conv2-2-2 | 3×3 | 1 | 1 | 32×32×32 |
| conv3-1-1 | 3×3 | 2 | 1 | 16×16×64 |
| conv3-1-2 | 3×3 | 1 | 1 | 16×16×64 |
| conv3-2-1 | 3×3 | 1 | 1 | 16×16×64 |
| conv3-2-2 | 3×3 | 1 | 1 | 16×16×64 |
| conv4-1-1 | 3×3 | 2 | 1 | 8×8×128 |
| conv4-1-2 | 3×3 | 1 | 1 | 8×8×128 |
| conv4-2-1 | 3×3 | 1 | 1 | 8×8×128 |
| conv4-2-2 | 3×3 | 1 | 1 | 8×8×128 |
| gap | 8×8 | 1 | 0 | 1×1×128 |
| fc8 | - | - | - | 1×1×8 |

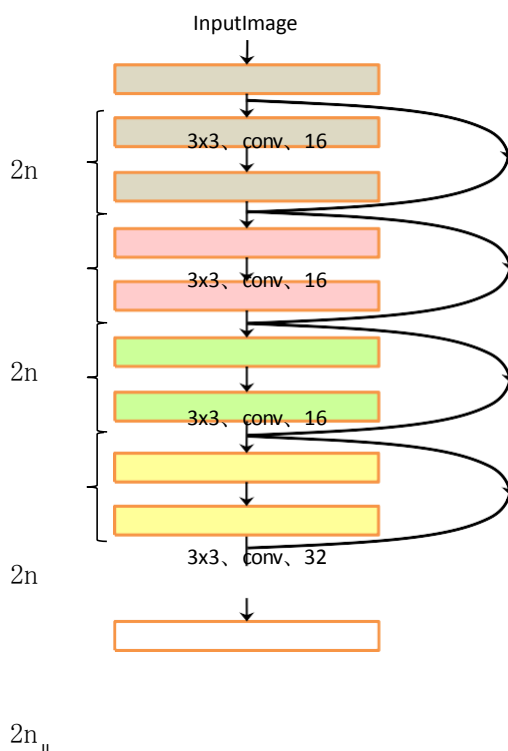


図 3-1- 53 提案手法で用いたネットワークの構造

ここでは、Haar-like 特徴量を用いた AdaBoost をベースラインとして、CNN・FCN・Super Parsing の実装を行った。更に、提案手法といくつかの手法の組み合わせについても訓練を行い、先に構築した広域監視に向けた野鳥のデータセットを用いてそれらの性能を評価した

[評価に使用したデータセット]

まず、データセットから異なる気象条件のシーンを含んだ 82 枚の画像を選択した。データセットは、風車周辺の 2806×3744 ピクセルの画像を含んでいる。選択した画像に対して、鳥・風車・空・森の 4 クラスのラベル付けを全て手動で行った。悪天候のため暗すぎる 5 枚の画像を除外して、残った 77 枚の画像を SuperParsing と FCN を訓練に用いた。特に FCN の訓練においては、そのままのサイズの画像を用いると必要な GPU メモリのサイズが大きすぎて計算処理ができないため、500×500 ピクセルの大きさに切り取って使用した。画像の半分以上の領域が空で覆われているため、画像全体から無作為に切り取ると、空のラベルだけを含む画像が多く得られてしまう恐れがある。このことを考慮して、風車周辺の領域(1306×3744)を集中して切り取ることにした。500×500 ピクセルのウィンドウを 250 ピクセルずつ位置を変えて領域を切り取る事により、1 枚の画像から 70 フレームが得られる。最終的に、FCN のため 77×70=5390 枚のフレームを得た。

ResNet の訓練画像としては、背景差分法により得られた移動物体の候補領域を用いた。訓練画像は鳥と非鳥の領域を含んでおり、鳥のクラスと 7 種類の背景のクラスを用意した。追加した背景のクラスは候補領域に多く含まれており、誤検出しやすいため、ネットワークの訓練に役立つ。Haar-like 特徴量を用いた AdaBoost の訓練には、ResNet の訓練の際と同様に集められた 15,705 枚の鳥画像と 18,688 枚の非鳥画像を用いた。

[評価]

ラベル付けされた 77 枚の画像の内、他より多くの鳥が写っている 44 枚の画像を用いて評価した。全体で 183 羽の鳥が含まれており、手法の性能は精度と再現率の調和平均である F 値を用いて評価した。評価の際に、実際に鳥がいる領域と検出されたバウンディングボックスが重なる領域が少しでもあれば正しく検出したとみなし、全く重なる領域が無ければ未検出とみなした。同様に、領域分割を基にした手法においても、実際に鳥がいる領域と出力された領域との重なりがあれば正しく検出したとみなし、無ければ未検出とみなした。

[評価結果]

広域監視に向けた野鳥のデータセットにおける検出結果の例を図 3-1- 54～図 3-1- 58 に示す。図 3-1- 54 は、結果の代表例を示している。図 3-1- 55 は、CNN は小物体の検出が得意であることを示している。図 3-1- 56、どの手法でも検出できない鳥が存在することと、SP は他の手法と組み合わせた場合に、鳥の形状の認識を助けることを示している。図 3-1- 57、FCN は細い形状の鳥を検出することは得意ではないことを示している。図 3-1- 58、手法によって誤検出しやすい領域には差異があることを示している。

評価対象の truepositive(TP)と falsepositive(FP)をカウントし、それらを用いて精度、再現率、そして F 値を計算した。結果のまとめを表 3-1-2 に示す。簡単のため、Haar-like 特徴量を用いた AdaBoost と SuperParsing はそれぞれ HA、SP と表記する。加えて、SP*と FCN*は手法に SVM を適用したことを表す。通常は、SP と FCN は最も尤度の高いクラスをピクセル毎に出力するが、SVM では学習を通じて全てのクラスを考慮に入れて出力を行う。

表 3-1-2 の上部は個々の手法の F 値を示す。SP は最も高い精度を示した。一方 FCN は最も高い再現率を示した。FCN は SP と CNN の中間の性能を示した。事前の想定通り、CNN は HA の性能を大きく上回った。SP*と FCN*は SVM を用いない場合よりも良い性能を示した。これは、SVM が意味領域分割を用いた手法の性能を最大まで引き出す事ができることを示している。

表 3-1-2 の下部は組み合わせた手法の結果を示す。F 値の点では、組み合わせた手法は全て個々の手法の性能を上回っている。特に、SP を組み合わせた手法はより高い精度を示した。これは、SP は背景をうまく認識することが出来るので、結果として FP を抑えることに繋がったと考えられる。CNN+FCN の結果は、FCN も背景認識ができることを示している。提案手法である CNN+FCN+SP が最も高い F 値を示した。

表 3-1-2 各手法の F 値

| 手法 | 精度 | 再現率 | F 値 |
|-------------------|--------------|--------------|--------------|
| HA | 0.064 | 0.514 | 0.114 |
| SP | 1.000 | 0.366 | 0.536 |
| FCN | 0.684 | 0.519 | 0.590 |
| CNN | 0.598 | 0.902 | 0.719 |
| SP* | 0.989 | 0.508 | 0.672 |
| FCN* | 0.709 | 0.585 | 0.641 |
| FCN+SP | 1.000 | 0.546 | 0.707 |
| CNN+SP | 0.950 | 0.618 | 0.748 |
| CNN+FCN | 0.924 | 0.798 | 0.856 |
| 提案手法 (CNN+FCN+SP) | 0.955 | 0.803 | 0.872 |

[大きさ別の各手法における性能]

提案手法の鳥画像の大きさに対する頑健さを示すために、鳥画像の大きさ別の F 値を表 3-1-3 に示す。先行研究の大きさの基準を参考に、tiny($\leq 15 \times 15$)、small($\leq 45 \times 45$)、normal($> 45 \times 45$)の3種類の大きさに分類した。

個々の手法の結果をみると、tiny の大きさでは CNN が最も高い再現率を示した。small の大きさでは SP*が最も高い精度を示した。一方、CNN は最も高い再現率を示した。FCN*は SP*と CNN よりも高い F 値を示しているが、これは FCN*は二番目に高い精度と高い再現率を示しており、精度と再現率がどちらも比較的良い結果であったからと考えられる。このことは、FCN*は鳥の認識と背景の検出を同時に行うことができることを示している。normal の大きさでは、FCN*が最も高い精度を示した。一方、CNN は最も高い再現率を示した。tiny の大きさの場合を除き、FCN*は最も高い F 値を示した。

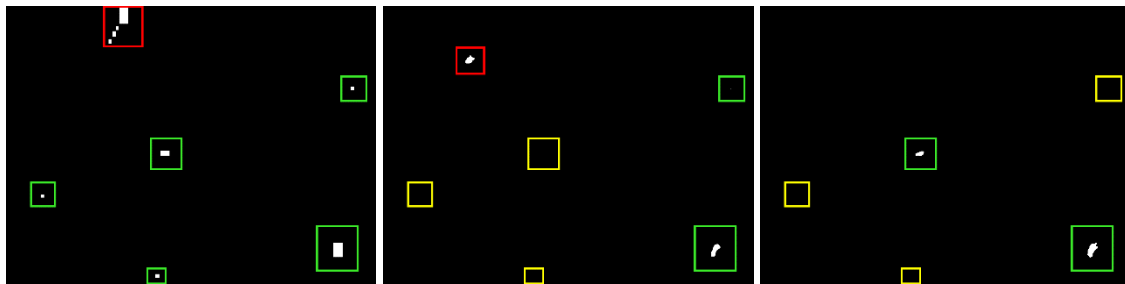
組み合わせた手法の結果をみると、F 値の点では、全ての大きさにおいて提案手法が最も高い値を示した。SP は tiny の物体検出には適していないが、多手法と組み合わせることで効果的な働きをしている。small と normal の大きさの場合を見ると、CNN+FCN がより高い F 値を示し、FCN+SP、CNN+SP が後に続いた。

表 3-1-3 大きさ別の各手法における F 値

| 大きさ | 手法 | 精度 | 再現率 | F 値 |
|--------|-------------------|--------------|--------------|--------------|
| tiny | SP* | - | - | - |
| | FCN* | 0.093 | 0.060 | 0.073 |
| | CNN | 0.541 | 0.896 | 0.674 |
| | FCN+SP | 1.000 | 0.030 | 0.058 |
| | CNN+SP | 0.826 | 0.284 | 0.422 |
| | CNN+FCN | 0.808 | 0.627 | 0.706 |
| | 提案手法 (CNN+FCN+SP) | 0.860 | 0.642 | 0.735 |
| small | SP* | 1.000 | 0.763 | 0.865 |
| | FCN* | 0.932 | 0.863 | 0.896 |
| | CNN | 0.667 | 0.875 | 0.757 |
| | FCN+SP | 1.000 | 0.800 | 0.889 |
| | CNN+SP | 0.969 | 0.775 | 0.861 |
| | CNN+FCN | 0.972 | 0.863 | 0.914 |
| | 提案手法 (CNN+FCN+SP) | 1.000 | 0.863 | 0.926 |
| normal | SP* | 0.970 | 0.889 | 0.928 |
| | FCN* | 1.000 | 0.944 | 0.971 |
| | CNN | 0.583 | 0.972 | 0.729 |
| | FCN+SP | 1.000 | 0.944 | 0.971 |
| | CNN+SP | 1.000 | 0.890 | 0.941 |
| | CNN+FCN | 1.000 | 0.972 | 0.986 |
| | 提案手法 (CNN+FCN+SP) | 1.000 | 0.972 | 0.986 |



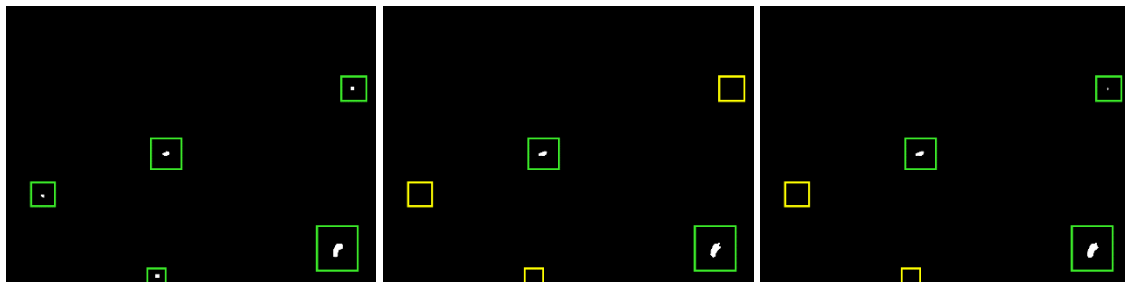
Input



CNN

FCN*

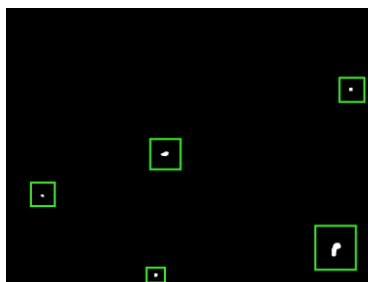
SP*



CNN+FCN

CNN+SP

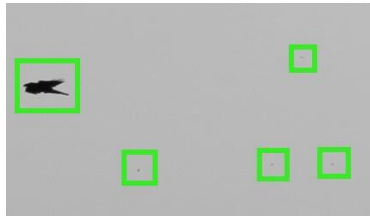
FCN+SP



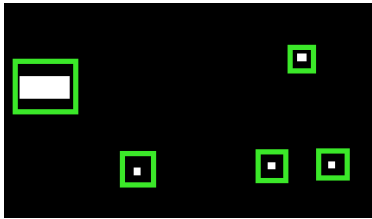
提案手法 (CNN+FCN+SP)

図 3-1- 54 鳥検出結果の代表例

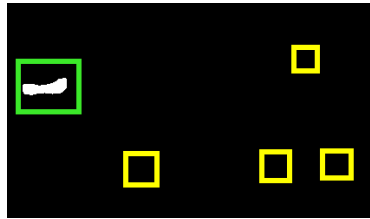
緑色の四角で囲われたものは TP である. 赤色の四角で囲われたものは FP である. 黄色の四角で囲われたものは falsenegative (以下 FN) である. CNN は誤検出も多いが、全ての鳥を検出している. SP*は小さな鳥を見落としているが、誤検出はしていない. FCN は小さい鳥も一部検出でき、CNN と比較して誤検出は抑えられている.



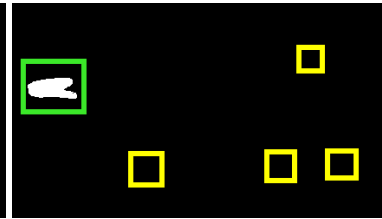
Input



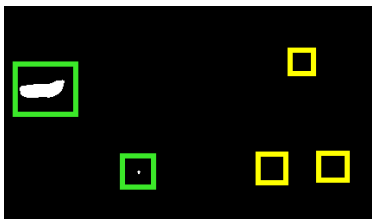
CNN



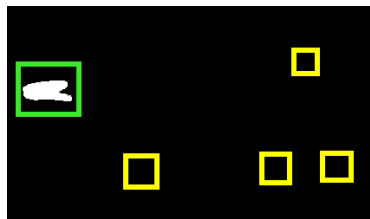
FCN*



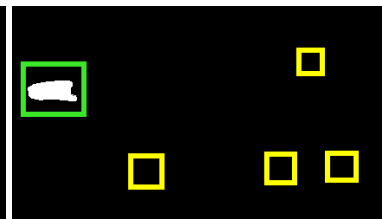
SP*



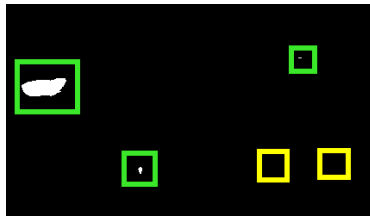
CNN+FCN



CNN+SP



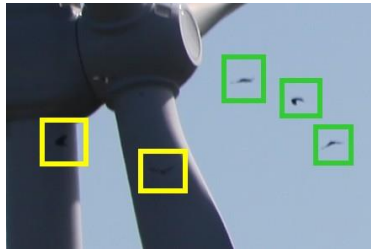
FCN+SP



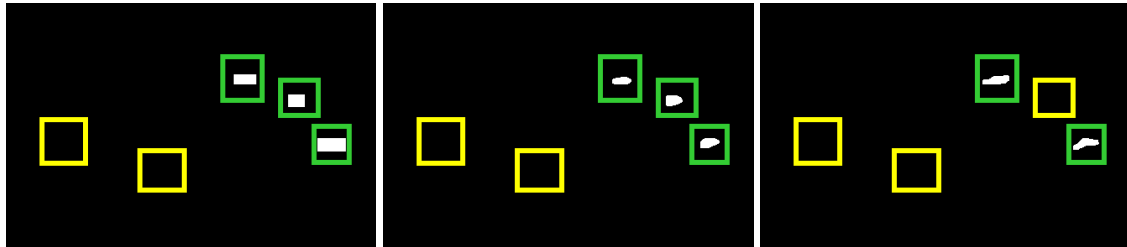
提案手法 (CNN+FCN+SP)

図 3-1- 55 tiny の鳥検出結果の例

緑色の四角で囲われたものは TP である. 黄色の四角で囲われたものは FN である. この場面では, CNN は全ての鳥を検出した. SP は鳥の形状をよく理解している.



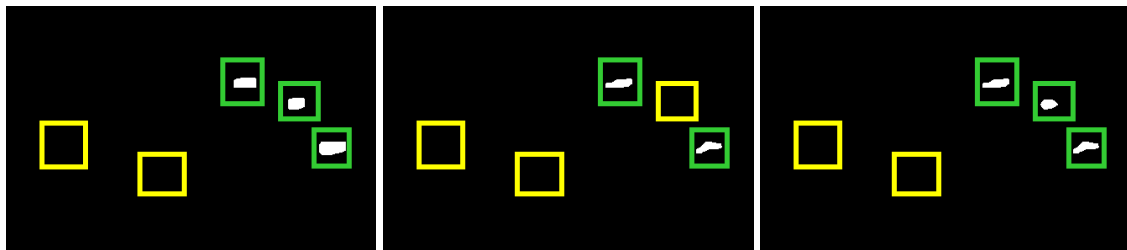
Input



CNN

FCN*

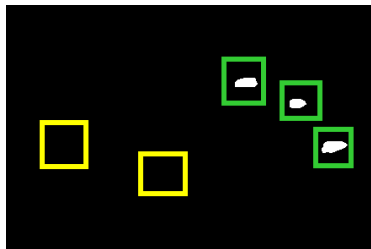
SP*



CNN+FCN

CNN+SP

FCN+SP



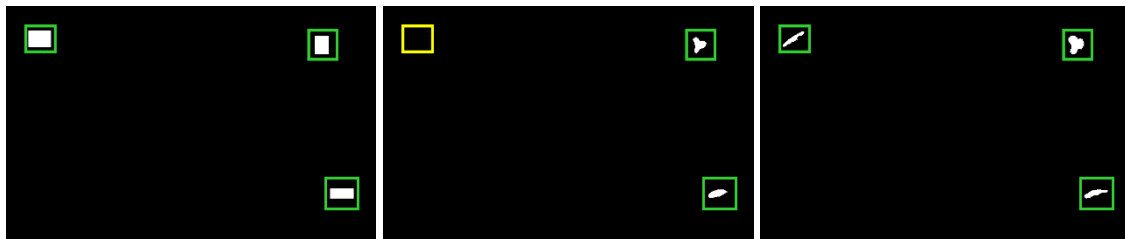
提案手法 (CNN+FCN+SP)

図 3-1- 56 鳥検出結果の例

鳥検出結果の例. 緑色の四角で囲われたものは TP である. 黄色の四角で囲われたものは FN である. 全ての手法において、風車と重なった位置にいる鳥は検出できなかった. SP は鳥の形状の認識を助けている.



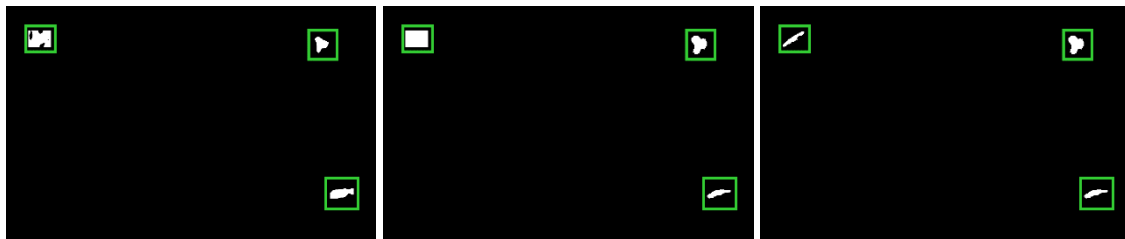
Input



CNN

FCN*

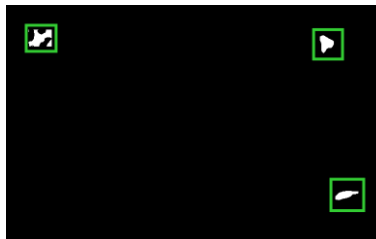
SP*



CNN+FCN

CNN+SP

FCN+SP



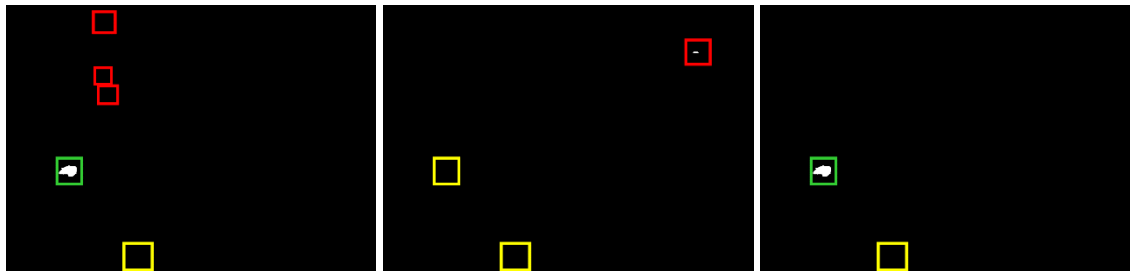
提案手法 (CNN+FCN+SP)

図 3-1- 57 鳥検出結果の例

鳥検出結果の例. 緑色の四角で囲われたものは TP である. 黄色の四角で囲われたものは FN である. FCN*は細長い形状の鳥を見落としているが、他の手法と組み合わせることにより検出できるようになった.



Input



CNN

FCN*

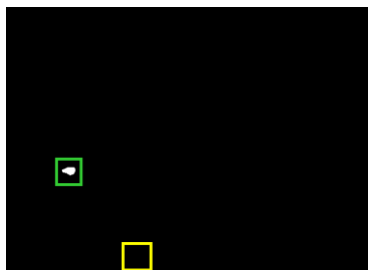
SP*



CNN+FCN

CNN+SP

FCN+SP



提案手法 (CNN+FCN+SP)

図 3-1- 58 誤検出結果の例

緑色の四角で囲われたものは TP である. 赤色の四角で囲われたものは FP である. 黄色色の四角で囲われたものは FN である. CNN はハブ周辺を誤検出している. FCN はプレードの先端を誤検出している.

[まとめ]

当事業では、バードストライク対策技術の一つとして鳥監視カメラシステムの開発を進めてきた。この中で、鳥種別判定機能の充実化を図ることを目的として動画による軌跡検出機能について開発を行ってきた。これまでの調査では、鳥の専門家が種別を判別する方法は、形状で判別する以外に鳥の飛翔パターン（羽ばたき方、動き方、速さ、軌跡）も補足情報として活用していることがわかっている。また、近年の急速な撮像技術高度化により、4Kクラスの解像度を有するカメラが手ごろな価格で民生用として利用できるようになってきた。数年前のWebカメラには適用が困難であったこれまで開発してきた静止画種別判定アルゴリズムが、近い将来高解像度ビデオカメラと連携することで実現できると予想、期待できることから、静止画の種別判定アルゴリズムを活用し、さらにパーティクルフィルターを導入した動画軌跡検出機能の開発を進めた。その結果、誤差2ピクセル未満での軌跡検出が可能となる軌跡検出機能を実現し海ワシの飛翔パターンの検出、学習教師データの取得が可能となった。また、鳥分類の手法として、CNNを基にした検出器とFully Convolutional NetworkとSuper ParsingをSupport Vector Machineを用いて組み合わせることで、大きな画像の中から小さい物体を検出する課題に対して高い性能を示すことが示された。同時に提案手法が高精度に鳥を検出出来ることが示された。

今後、実用的なシステムへの実装を行うことで、バードストライクは低減されることが期待される。

[謝辞]

研究を進めるにあたり、事業を支援してくださった環境省野生生物保護課、日本気象協会に深く感謝申し上げます。また、研究に貢献してくださった苗村健教授、川上玲助教、吉橋亮太君、竹木章人君はじめ苗村研究室の皆さんに感謝いたします。データセットの構築に用いた画像データは、環境省の釧路自然環境事務所、東北地方環境事務所、関東地方環境事務所、北海道地方環境事務所、九州地方環境事務所の各野生生物課から多くの写真提供がありました。NPO法人バードリサーチ 植田睦之代表はじめ鳥データの作成作業に参加してくださった皆様にも深く感謝申し上げます。

【参考文献】

1. W. P. Kuvlesky, L. A. Brennan, M. L. Morrison, B. M. B. K. K. Boydston, and F. C. Bryant, "Wind energy development and wildlife conservation: Challenges and opportunities," *The Journal of Wildlife Management*, vol. 71, no. 8, pp. 2487-2498, 2007.
2. B. Snyder and M. J. Kaiser, "Ecological and economic cost-benefit analysis of offshore wind energy," *Renewable Energy*, vol. 34, no. 6, pp. 1567-1578, 2009.
3. K. S. Smallwood, L. Rugge, and M. L. Morrison, "Influence of behavior on bird mortality in wind energy developments," *The Journal of Wildlife Management*, vol. 73, no. 7, pp. 1082-1098, 2009.
4. A. Rioperez and M. de la Puente, "DTbird: A self-working system to reduce bird mortality in wind farms," *European Wind Energy Association Conference*, 2010.
5. E. Wiggelinkhuizen, S. Barhorst, L. Rademakers, H. den Boon, and S. Dirksen, "WT-Bird: Bird collision monitoring system for multi-megawatt wind turbines," *European Wind Energy Association Conference*, 2007.
6. D. Lack and G. C. Varley, "Detection of birds by radar," *Nature*, vol. 156, pp. 446, 1945.
7. W. L. Flock, "Monitoring bird movements by radar," *IEEE spectrum*, pp. 62-66, 1968.
8. N. Huansheng, C. Weishi, M. Xia, and L. Jing, "Bird-aircraft avoidance radar," *IEEE Aerospace&Electronics systems magazine*, 2010.
9. T. Berg, J. Liu, S. W. Lee, D. W. Alexander, M. L. and Jacobs, and P. N. Belhumeur, "Birdsnap: Large-scale fine-grained visual categorization of birds," *Computer Vision and Pattern Recognition (CVPR)*, 2014.
10. P. Welinder, S. Branson, T. Mita, W. C., S. F., S. Belongie, and P. Perona, "Caltech-ucsd birds 200," 2010.
11. Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Computational Learning Theory*, vol. 904, pp. 23-37, 1995.
12. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, 1998.
13. L. Bottou, "Stochastic gradient descent tricks," *Neural Networks: Tricks of the Trade*, pp. 421-436, 2012.
14. A. Mahendran and A. Vedaldi, "Understanding deep image representations by inverting them," in *Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 5188-5196.
15. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European Conference on Computer Vision (ECCV)*, 2014, pp. 818-833.
16. A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 427-436.
17. J. Tighe and S. Lazebnik, "Superparsing," *International Journal of Computer Vision (IJCV)*, 101(2):329-349, 2013.