

接近検知無線システムを用いた猿の出現予測手法の提案

石倉 功規^{†1} 小林 秀幸^{†2} 中井 一文^{†3} 江崎 修央^{†3} 山端 直人^{†4} 糺谷 斉^{†4} 杉浦 彰彦^{†1}

静岡大学大学院 情報学研究所^{†1} 仙台高等専門学校 情報システム工学科^{†2}

鳥羽商船高等専門学校^{†3} 三重県農業研究所^{†4}

1. はじめに

猿による農作被害は年間十数億円であり、問題視されている。猿は頭が良く、単純な方法では追い払いづらいため、人の手で追い払うことが最も効果的である。しかし、人が一日中農作地を監視するのは大変であり、現実的ではない。

本研究では、あらかじめ猿に対する準備を行えるようにするため、猿の出現を予測するシステムを作成することを目標とする。今回は予測を行うために、猿に電波を発する首輪を取り付け、受信機により動向データを採取する。採取したデータから猿が接近している際のデータを抽出し、そのデータを気象データ等と合わせて、ある時刻に猿が出現するかどうかの予測を行う。出現数の割合による予測とSVM(Support Vector Machine)による予測の2種類について実験を行い、予測精度について比較・検討を行う。

2. 原理

2.1 受信信号強度

猿の接近を検知するためには電波受信時の受信信号強度を測定し、その情報から接近しているかどうかを判断する必要がある。今回の実験で使用した受信機の受信信号強度の値は0から1023までの整数値で表される。

2.2 SVM(Support Vector Machine)

SVMは教師あり学習の手法の一つであり、クラス間のマージンを最大化するという特徴がある[1,2]。本研究では、SVMに猿の接近した際の気象データ等を学習させ、出現予測に利用する。出現予測に使用するパラメータとして、気象データ(気温、湿度、降水量、気圧、天候)と、気象

	0	1	2	3
Temperature[°C]	20-21	22-23	24-25	26-27
humidity[%]	41-50	51-60	61-70	71-80
precipitation[mm]	0-1	2-3	4-5	6-7
atmosphere[hPa]	965-969	970-974	975-979	980-984
weather	shine	cloud	rain	fog
time	0-3	4-7	8-11	12-15
day	Sun	Mon	Tues	Wed

図1 パラメータ例

データ以外の情報として時刻と曜日を使用する。これらのデータをパラメータ毎に6~8つ程度の区画に分け、それぞれに0から7までの数値を割り振る。数値の割り振り例を図1に示す。ある時刻におけるデータSVMに入力する際は、パラメータ毎にそれぞれ対応する数値に割り振り、そのデータをSVMに入力する。

また、猿の出現時のデータと、非出現時のデータを区別し易くするため、我々は、猿の出現時のデータを集計し、出現率が高い区画から順番に番号を割り振る手法を提案する。これにより、猿の出現時のデータは数値が小さくなり易くなり、非出現時のデータと区別し易くなると考えた。

3. 実験

3.1 猿の動向調査

猿の出現予測を立てるために、まず猿の行動データの収集を行う。そのために、猿からの電波の情報を収集するシステムを作成した。このシステムは、猿に取り付けられた首輪からの電波を受信機が受信した場合、サーバへ受信データを送信するというものである。

三重県伊賀市の猿が出没する山を囲うように12箇所にアクセスポイント(以下、AP)を設置し、APからのデータをサーバで集計した。首輪を付けた猿は全部で8匹である。山の長さは約10kmで、AP間の距離は0.5km~1.5kmである。記録したデータは、日時、受信信号強度、APの番号、首輪の個体識別番号である。

“Proposal of Appearance Prediction Method of Monkey Using Contiguity Detection Wireless System”

^{†1} Koki Ishikura, Akihiko Sugiura

Shizuoka University, Graduate School of Informatics

^{†2} Hideyuki Kobayashi

Sendai National College of Technology

^{†3} Kazufumi Nakai, Nobuo Ezaki

Toba National College of Maritime Technology

^{†4} Naoto Yamabata, Hitoshi Kojiya

Mie Prefecture Agricultural Research Institute

3.2 猿の出現予測

3.1で得られたデータを基に猿の出現予測を行う。対象とするデータは68日間のデータ約35万件である。このデータの中から猿が農作地に接近している場合とそうでない場合を区別するため、受信信号強度の値が376以上であったもの約3000件を猿が出現したデータとして抽出する。今回は時間を2時間単位に分割し、2時間毎に猿の出現、非出現を判別する。判別した結果、出現数は56回、非出現数は684回であった。このデータを使用して予測を行う。

3.2.1 出現数の割合による予測

猿が出現していた時のパラメータを集計し、それぞれ6～8つの区画に分類する。そして、区画毎の出現数の割合(以下、係数とする)を算出し、テーブルに纏める。予測する際は、予測対象となる日時における天候等のパラメータを基にテーブルから係数を引き、各パラメータの係数を全て乗算する。乗算した結果、求められた値が判断基準より上ならば猿が出現すると判断する。

3.2.2 SVMによる予測

猿が出現していた時のデータと猿が出現していない時のデータをそれぞれ教師データとし、SVMによりクラス分けする。そして、対象となるデータがどちらのクラスに分類されるかを判断することで、猿の出現を予測する。教師データは、猿の出現時のデータ56件、猿の非出現時のデータ674件である。

また、出現時のデータと非出現時のデータ数を合わせるため、非出現時のデータの中からランダムで56件抽出し、そのデータを教師データとした場合についても実験を行う。

4. 結果・考察

予測結果を表1に示す。「出現判断数」は出現と判断された回数である。「正解数」は出現すると判断された時刻に実際に猿が来ていた場合にカウントされ、最大は56である。「認識率」は正解数/56である。「SVM」の行の「データ数変更前」は非出現時のデータ数が674件の場合、「データ数変更後」は非出現時のデータ数が56件の場合を表している。

結果を見ると、出現数の割合による予測では、認識率を上げるために判断基準を下げると、必要以上に出現すると判断してしまっていることがわかる。これは、実際に猿が出現した時刻とその周囲の時刻で気象パラメータが近くなるため、実際に猿が出現した時刻の周囲の時刻まで出現すると判断されるためだと考えられる。

表1 予測結果

手法		出現判断数	正解数	認識率
出現数の割合	判断基準高	34	12	21.4%
	判断基準低	700	55	98.2%
SVM	データ数変更前	14	14	25.0%
	データ数変更後	269	56	100%

一方、SVMによる手法では出現数の割合による予測と比べ、誤認識数を抑えつつ猿の出現予測ができていたことが分かった。非出現時のデータ数が674件の場合では、誤認識は無いものの正解数は14であり、認識率は25%であった。これは、教師データにおける猿出現時のデータ数と比べ、非出現時のデータ数が多く、判定が非出現に傾き易くなってしまったためと考えられる。

非出現時のデータ数が56件の場合では正解数は56まで上昇し、認識率100%となった。しかし、出現判断数が269まで増え、誤認識数が増加してしまっていることが分かる。これは、教師データにおける非出現時のデータ数が減少したため、出現時のデータに類似した非出現時のデータを判定する際に誤認識が発生し易くなったためだと考えられる。

5. まとめ

本研究では、猿の出現予測を立てることを目標とし、調査として三重県伊賀市に調査用のシステムを設置し、データの収集を行った。収集したデータから出現数の割合による予測とSVMによる予測を行った。その結果、出現数の割合による予測では、判断基準を下げ認識率を上げていくと必要以上に出現すると判断してしまっていることが分かった。SVMによる予測では出現数の割合による予測と比較して誤りが少なく、正解数も多いため予測に適していることがわかった。

今後は、予測精度を更に向上させ、認識率の向上と誤認識の削減を目指す。また、ベイズ推定等の別の予測手法とも比較し、予測に適している手法を調査する。

Reference

- [1] 青木 一真, 黒柳 奨, マウリシオ クグレ, アントサトリヨ スグロホ, 岩田 彰. “Confident Marginを用いたSVMのための特徴選択手法,” 信学論(D), Vol.J88-D-II, No.12, pp.2291-2300, 2005.
- [2] 伊波 靖, 高良 富夫. “危険なシステムコールに着目したWindows向け異常検知手法,” 情処論, Vol.50, No.9, pp.2173-2181, 2009.