特　集「情報セキュリティと AI」

ニューラルネットワークに基づくセキュリティ技術

Information Security Techniques Based on Artificial Neural Network

直江　健介 Kensuke Naoe
慶應義塾大学大学院政策・メディア研究科
Graduate School of Media and Governance, Keio University.
naoe@sfc.keio.ac.jp, http://web.sfc.keio.ac.jp/~naoe, http://www.inas.mag.keio.ac.jp/

田中　秀和 Hidekazu Tanaka
（同　上）
pidekazu@sfc.keio.ac.jp, http://web.sfc.keio.ac.jp/~pidekazu

武藤　佳恭 Yoshiyasu Takefuji
慶應義塾大学環境情報学部
Faculty of Environmental Information, Keio University.
takefuji@sfc.keio.ac.jp, http://www.neuro.sfc.keio.ac.jp/

Keyword: neural network, information security, digital watermark, banknote recognition, virus detection.

1. はじめに

コンピュータをはじめとする情報処理技術の技術革新により，日々の生活の利便性は飛躍的に向上した．その一方で，ウイルス感染によるデータ破壊や P2P ソフトウェアによる違法ファイル共有や個人情報の流出のように，コンピュータやディジタルデータを扱うゆえに起こる被害も増大している．

情報セキュリティ技術は，このような社会の営みのうえに安全性と利便性を与えてくれる技術として期待されてきた．例として，不正プログラムの検出，ネットワークやホストコンピュータの異常検出，ディジタルコンテンツの著作権管理や不正コピー防止，暗号技術や認証システムなどがある．これらは我々人間のコンピュータを用いた社会的行動の基盤を支えている．

情報セキュリティ技術には人工知能の考え方が利用されたものが少なくない．コンピュータは計算処理速度が非常に速いため，人間には不可能な膨大な計算をすることができる．一方，単純な計算の繰返しだけでは機械的でつまらない結果しか得ることができない．それに比べて人間の計算速度は遅いが，熟練した技術者に見られる直感や学習能力には驚くべきものがある．このような人間的な能力をコンピュータで実現することが人工知能の基本的な考え方である．人工知能の研究は人間の知的営みともいえる推論，探索，知識処理，データマイニング，画像認識，音声認識といった分野に発展していった．その中でも，本論文で扱うニューラルコンピューティングとは，生物の脳神経細胞のネットワークを数学的にモデル化して，生物的な知的処理をコンピュータ上で擬似的に再現する研究分野である．

2. ニューラルネットワーク

ニューロンとは神経細胞のことで，脳は多くの神経細胞からなっている．ニューロンの一つ一つは電気信号の入力と出力という単純な作業しかできない．しかし，このニューロンが有機的に接続されたネットワークを形成すると，さまざまな複雑な処理をすることができるようになる [武藤 01].

このニューロンにより形成されたネットワークをモデル化したものをニューラルネットワークと呼ぶ．モデル化の例としては，電気信号の入力およびそれによって上昇する内部エネルギーの変化をモデル化したもの，内部エネルギーがしきい値を超えると電気信号を出力するという構造のモデル化，脳細胞間の接続についてのモデル化などがあり，ニューラルネットワークの細部にわたってのモデル化を行うことができる．人工ニューラルネットワークとはモデル化した生物学的な神経細胞の構造をソフトウェアやハードウェアにて再現・構築したものであり，人工ニューラルネットワークを計算機上で処理することをニューラルコンピューティングと呼ぶ．

2･1　さまざまなニューラルネットワークモデル

ニューラルネットワークは個々のニューロンの性質，結合やネットワークの組み方を変えることでさまざまな種類の問題を解くことができるようになる．

ニューロンとは，ほかのニューロンから信号の受信を行い，単一の出力を行う素子である．ニューロン素子のモデルには，バイナリモデル [Hopfield 85, McCulloch

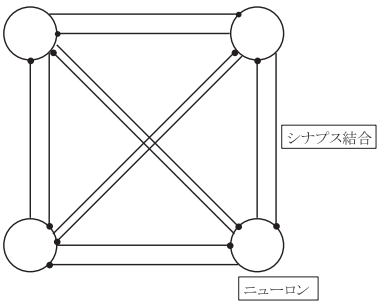


図1 リカレント結合

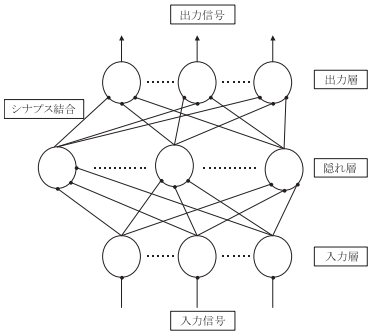


図2 フィードフォワード結合

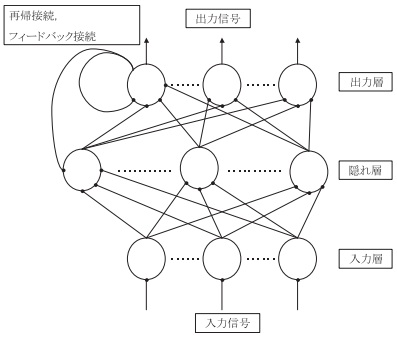


図3 フィードバック結合

43]，シグモイドモデル [Seggern 93]，Radial Basis Function モデル [Moody 89, poggio 90]，競合ニューロンモデル [Kohonen 95, Takefuji 92] などがある．これらのモデルの違いは主にニューロンの入出力関数の違いからくるものである．

またニューロン同士のシナプス結合の違いでモデル化することもでき，このことによりネットワークタイプや出力のダイナミクスに違いが出てくる．結合形態による分類としては主に，ニューロンが相互に結合されるリカレント結合（図1），ニューロンが階層的に接続され信号が単純な順伝播をするフィードフォワード結合（図2），フィードフォワード接続の形態にユニットの逆伝播を行う接続があるフィードバック結合（図3）などがある．

これらの結合のパターンとニューロンモデルの違いによりニューラルネットワークのモデルは状態遷移モデル，競合学習モデル，信号伝播モデルの三つの種類にすべて分類することができる [Kohonen 95].

状態遷移モデルの代表例としては Hopfield モデル [Hopfield 85] がある．このモデルは連想記憶や最適化問題に用いられる．連想記憶に関しては Hebb 則 [Hebb 49] という学習則がある．現在ではこのモデルを改良したアルゴリズムが多く提案されている [Aoba 02, Takefuji 89].

競合学習モデルの代表例としては自己組織化マップ（SOM: Self Organization Map）[Kohonen 82, Kohonen 90] がある．SOM による競合学習は教師なし学習であり，非線形クラスタリング手法としても有名である．自己組織化マップは可視化情報処理機としても有名であり，幅広く応用されている [Tokutaka 99].

信号伝播モデルの代表例としては多層パーセプトロンモデルがある．これはパーセプトロンモデル [Rosenblatt 58] を多層に重ね合わせたものである．三層以上の多層パーセプトロンは非線形関数近似能力をもつとされ，学習法に一般デルタルールと呼ばれる学習則を用いたバックプロパゲーション学習法 [Rummelhart 86] が提案されることで一躍有名になった．そのため非線形分離問題などによく用いられているネットワークモデルである．

その他の著名なニューラルネットワークモデルには RBF ネットワークモデル [Broomhead 88]，パルスニューラルネットワーク [Eckhorn 90, Johnson 94]，Second Order Neural Network モデル [Maxwell 86]，カオスニューラルネットワークモデル [Aihara 90]，Elman-net モデル [Elman 94] などがある．

2･2 セキュリティ分野へのニューラルネットワークの応用

情報セキュリティにはさまざまな分野がある．画像認識問題はニューラルネットワークの得意とするところである．情報セキュリティ技術の中でも画像認識・画像解析という問題として捉えることが出来るとき，その解決手法としてニューラルネットワークがよく使われている．例えば顔認証システム [谷荻 94] や指紋認証 [IPA 04a] に代表されるバイオメトリクス技術などがその例である．

暗号の世界では，カオスニューラルネットワークを使った擬似乱数系での暗号生成という研究 [Kawamura 02] もあり，カオスニューラルの機能を埋め込んだチップの実用化が進んでいる．また，経済の世界ではトレンドの予測や傾向分析にニューラルが用いられているケースも少なくない [Iba 01]．これと同様の仕組みを用いた手法で，計算機を利用するユーザやネットワークの振舞いを観測し，不正なアクセスや異常状態を検出するビヘイビア検出法というヒューリスティックな手法がある [IPA 04b]．ビヘイビア検出の例としてクレジットカードの IC チップにニューラルネットワークを組み込むことで，持ち主の購買傾向や振舞いを学習し，そのパターンと大きくそれるような消費行動を監視するものがある．例えばいつもは買わないような高額な買い物を日に2回

する行動を検出すると，正規のユーザによる利用ではないと判断し，クレジットカード決済を強制的に行えないようにする Falcon と呼ばれるシステムもある [Falcon].

アンチウイルスソフトウェアベンダの一つである symantec ではヒューリスティックスキャンにニューラルネットワークのエンジンを使用しているという報告もある [Symantec 01]．これは定義ファイルにある既知のウィルス以外の未知なウィルスを検出するのに利用されている．もっと身近なところでは Microsoft Outlook2003 で SPAM 判定にニューラルデシジョンエンジンという機能を実装している [Microsoft 03].

このように，ニューラルネットワークを用いたセキュリティ技術は我々の身近なところに存在している．しかし，ニューラルネットワークモデルがすべての問題に対して万能というわけではなく，それぞれの問題に一番適したニューラルネットワークモデルを選択する必要がある．ニューラルネットワークが得意とする問題は探索空間の中に複数個解が存在し，全探索では計算効率が悪い問題などである．最適解を見つけなくてはいけない問題では，局所解に陥った場合にそこから抜け出す仕組みなどを用意する必要があり，前処理や後処理といった本質的でない部分に工夫を凝らす必要が出てくる．そのような問題にはニューラルネットワークは向いていないと考える．

以降の章では，どのようなニューラルネットワークモデルが個々のセキュリティ技術にどのように応用されているかを事例を用いて解説する．また，なぜこのような技術が求められているかを述べる．

3. 事例 1. ユーロ紙幣判別機

3･1 概　　要

紙幣判別機は紙幣の数え上げから両替や券売機などに用いられる非常に有用な機器である．紙幣はその性質上しばしば偽造されることがあるが，コピー機などの性能の向上に伴い精巧な偽札が簡単につくられるようになってきた．現在，偽券・金種の判別を正しく行う紙幣判別システムの需要が非常に高まってきている．我々の社会において紙幣を使い続けるかぎり，紙幣判別は必要不可欠なセキュリティ技術の一つである．

紙幣識別は静止画像認識問題の一種である．したがって，この問題の焦点はどのような識別関数が特徴ベクトルを効果的に分類・判定するのかを考えることである．紙幣判別機に入力される紙幣の特徴のばらつきというのはわかっていないため，ニューラルネットワークの適応性がこのような問題に非常に適しているといえる．

ここで述べるユーロ紙幣判別システム [Aoba 03] では三層パーセプトロンと RBF ネットワークの 2 種類のニューラルネットワークが用いられている．三層パーセプトロンはパターン認識の手法としてよく知られているツール

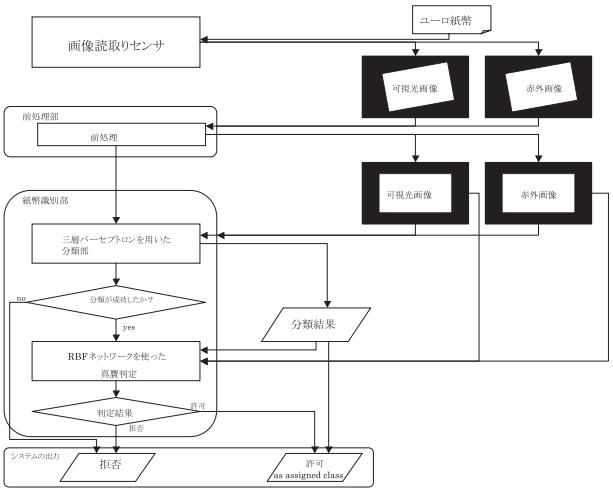


図 4　紙幣判別システムの全体図

であり，紙幣分類には非常に効果的である．三層パーセプトロンは与えられた学習データの分類のための境界線をつくることで紙幣の種別判断を行っている．しかし未知のデータ，つまり偽札に対して必ず拒否することを保証しない．それを保証するために RBF（Radial Basis Function）ネットワークを利用する．RBF ネットワークはサンプルデータの確率分布に沿った出力を得ることができるため，未知の不正なデータを排除することができる．

ユーロ紙幣判別システムは，まず前処理段階でエッジの抽出を行い，紙幣の向きや位置情報を取得する．次にユーロ紙幣を可視光と赤外線を使った読取りセンサで 2 種類の画像データを取り込む．この二つの画像データは，前処理で得られた紙幣の向きと位置情報をもとに補正された後，金種分類部と真贋判別部の二段階で分析される．まず紙幣の金種分類部では学習済み三層パーセプトロンに対して，前処理済みの画像を入力する．この入力データがクラスに正しく分類されたら，処理が引き続き判別部へと渡される．そうでなければ，紙幣ではないとして拒否される．紙幣の真贋判別部では学習済み RBF ネットワークへ前処理済みの画像が入力として渡される．この判別部で受理されたならば正しい紙幣として認識される．そうでなければ，偽券として認識され，拒否される仕組みである．この処理の流れを示した図を以下に示す．

3･2 金 種 分 類 部

ここでは金種分類部について説明する．入力画像はいくつかの大きさに区切った多解像画像とし，それを入力ベクトルとする．このシステムでは，入力ベクトルの次元数を減らすために冗長な入力ベクトルを排除する処理も行っている．これは入力の次元数が増えるとパラメータ推定を難しくしてしまうという，いわゆる次元の呪い問題を回避する工夫である．有用な入力データの選定を行った後に学習済み三層パーセプトロンへ信号を順伝播させ，分類を行う（図 5）．

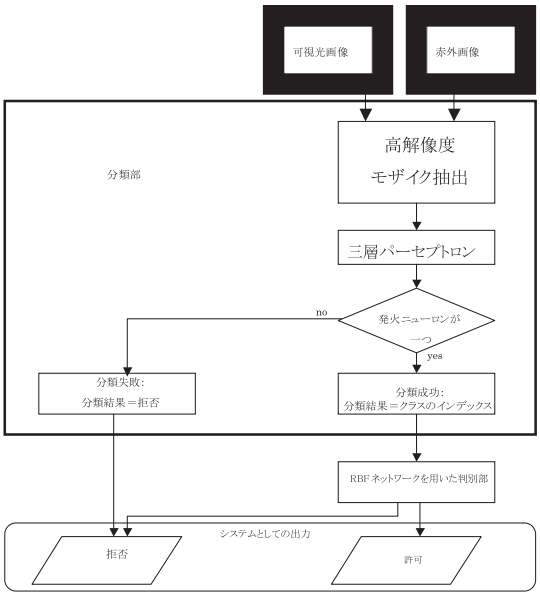


図5 金種分類処理の流れ

3･3 真贋判別部

次に真贋判別部について説明する．可視光画像と赤外画像はそれぞれ小さく区切った領域へと分割される．真贋判別部はいくつかの判別ブロックをもち，それぞれの判別ブロックは与えられたクラスに対応する．また判別ブロックは複数のRBFネットワークをもち，それぞれのRBFネットワークはそれぞれの画像を小さく区切った領域に対応する．つまり，仮にクラスの数がL個で，画像データがK個の小さい領域に分割されたならば，判別部ではL個の判別ブロックによって構成され，それぞれのブロックはK個のRBFネットワークをそれぞれの小さく区切られた画像に対してもつということである．この判別処理の出力のすべてが真であったときのみ，紙幣が真性であると判断される．出力の一つでも拒否を示すと，偽券であると判断される（図6）．

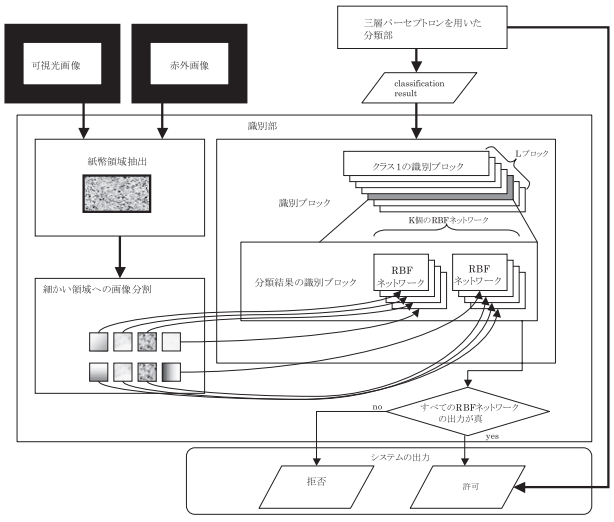


図6 真贋判別処理の流れ

4. 事例2. ウィルス検出技術

4･1 概　　要

コンピュータウィルスは日々増加傾向にあり，それらウィルスの亜種も非常に速いスピードで生まれている．従来のようにウイルスデータベースを構築し，疑わしきプログラムとデータベースとのマッチングによる検出手法では，データベースへのウイルス登録作業が新たに出現するウイルスに対して質的にも量的にも追いつかないため，近い将来コンピュータウイルスに対して我々コンピュータユーザは全く無力になってしまう．そこで昨今，ヒューリスティックな手法や生物学的アプローチによる，より高度なアンチウイルス技術が開発されている．特に未知のウイルスへの対応が求められており，異常検出の手法を扱うものが増えてきている．ここでは正常なブートセクタプログラムと，ウィルスによって感染されているブートセクタプログラムを学習したニューラルネットワークを用いて分類・判別するアンチウィルスシステムを紹介する [Kephart 95, Kephart 96].

このシステムでは，ウイルスであるかどうかを検知することはアルゴリズム的には不可能であるが，実際には，不完全でありながらも，包括的にウイルスを検出することが可能である．判別システムへ入力されたコードが正常であるかどうかを判断するという問題は，パターン分類問題と全く同じ問題として見ることができる．このシステムでは分類方法として階層型ニューラルネットワークモデルが利用されている．

4･2 第一段階分類

ブートセクタ型ウィルスの挙動は機械語の中においてさまざまな形で実行されるかもしれないが，ほとんどの機械語のコードインプリメーテーションは，数少ない限られた単純なパターンと一致する．そこでその単純なパターンを有するブートセクタプログラムを多層階層型ニューラルネットワークを使い分類する．しかしこの分類処理後にウィルスでないと分類されたコードはまだ完全に正規のブートセクタプログラムであると確定できるわけではない．なぜなら，下位メモリに感染するウイルスについての挙動は同じ機能を示すが，大部分のウイルスはコードの中に2バイトの特定なパターンを含んでいる．

4･3 第二段階分類

第一段階の分類処理をすり抜けてしまったウイルスコードを検出するには，さらにウイルスの特徴的なコードのパターンを抽出する必要がある．このシステムは，ブートセクタ感染型ウイルスが自動的にプロシージャを呼び出すことによって現れる特徴に着目している．その特徴は，ウイルス本体がプロシージャを呼び出すことで現れる数バイト長の列である．ウイルス本体のブートセク

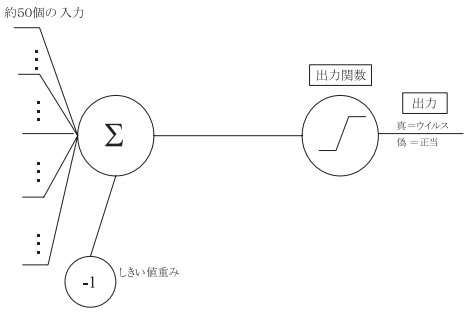


図 7 約 50 個の入力と重み付けを備えた単一層のニューラルネットワーク

タに含まれるバイト列を 3 バイトずつに区切るトライグラムで扱い，トライグラム表を作成する．

作成されたトライグラム表は非常に膨大になる．そのためウィルスの挙動を示す良い特徴を示したトライグラムを選択することがブートセクタ型ウィルスの検出精度に大きく関わる．このトライグラム選択の詳細については [Kephart 95] に述べられているが，このトライグラムの中には，正常なブートセクタの特徴を示すものは一切含まれていない．

図 7 は第 1 段階分類処理をすり抜けたコードを判別するニューラルネットワークであり，学習例上では完全にパターンを分類することができる．しかしこのニューラルネットワークでは，ある一つのウイルス的特徴が入力された場合でさえも発火することがあり得るので，低い確率で誤検出を起こす可能性がある．このようなことはバックプロパゲーション学習を行った後においても同じような形で現れることがあるが，比較的大きい重み付けをすることと入出力関数のシグモイド関数により誤発火を抑えることができる．

それでも検知を回避するウイルスの 10％または 15％のほとんどが，それらのウイルスの特徴を示すトライグラムが含まれていないからではなく，その特徴を示すコードセクションが，ウィルスのさまざまな手法によって隠蔽させられるために検出を逃れている．隠蔽されたコードを独自の手法で検出できれば，その特徴を示すトライグラムを分類器に渡すことができ，これらのウイルスについても検出することが可能となる．

5. 事例 3. 電子透かし

5･1 概　　要

コンテンツビジネスの世界において，電子透かしや電子署名というキーワードを耳にすることは多い．ディジタルコンテンツはその特性上何度繰り返して複製をつくっても品質の劣化がなく，取扱いが簡単である．インターネットの普及もあり，これらのコンテンツの公開・配布が容易になり，インターネット登場以前のような記憶媒体の管理による著作権の保護が難しくなってきた．

ディジタルコンテンツの著作権を主張するためには，著作者が著作物に透かしを施す必要がある．この透かしの抽出方法は著作者のみがわかる仕組みである必要があり，その透かし情報をもって著作者の証明を行う．電子透かしとは著作物である映像，画像，音楽といったディジタルコンテンツに対して著作者情報などといった管理に必要な情報を付与する技術である．この類の付与情報は悪意のある攻撃などにより除去されてはならないため，頑健な電子透かし技術はコンテンツビジネスの世界において重要な情報セキュリティ技術となっている．

電子透かしは絵画などの署名のように目に見えるような形で埋め込むものもあるが，映像や音声といったコンテンツにはその方法は利用できない．また透かしを埋め込むことでコンテンツの質的変化を伴わないものが望ましい．そのため人間の感覚器官では知覚できない信号を透かし情報として埋め込むという方法が一般的にはとられている．つまり埋込み手法には，埋込みによるコンテンツの質的変化を最小限に抑える技術が必要である．

コンテンツに対するメタデータの付与という意味では，付与された情報が除去されてもさほど大きな問題にならない．しかし著作権情報を含む透かしの場合，悪意のある除去攻撃やフォーマット変換に対する耐性が求められる．しかし多くの場合，フォーマット変換や除去攻撃への耐性を優先することが質的変化を伴ってしまう．ここで紹介する電子透かしは，ニューラルネットワークを用いることで，質的変化を伴わずに耐性を有することの両立を目指している．

紹介するシステム [Ando 03, 直江 03] ではニューラルネットワークを利用した透かしの抽出鍵生成処理という特徴をもつ．それは周波数変換後の任意の領域の特徴値を三層パーセプトロンへの入力とし，埋め込みたい透かし情報を教師信号とするバックプロパゲーション学習を行うことである．この特徴情報を入力としたときに出力として透かし情報を得られたなら，正しくクラスの分類が行われたとして学習を終え，学習が収束した際の結合係数を透かし情報の抽出鍵として扱う．

5･2 鍵の生成処理

画像に対する周波数変換後，サブブロック分割されたブロックから一つサブブロックを選択する．この選択されたブロックはその画像の特徴情報を有するサブブロックとして学習を行う．この学習は直流成分を含む斜め方向のすべてのピクセルの値を入力とし，透かし情報を教師信号とするような学習である．例えば 8*8 ピクセルの画像であれば入力は斜め 8 ピクセルの係数である．学習が収束した時点の結合係数を抽出鍵として利用する．これは，結合係数が分類器として機能するため，外部で保持する．

5･3 透かしの抽出処理

透かし情報の抽出に関しては，埋込み処理時に生成さ

れた結合係数を用いたネットワークに対して，学習を行った領域の特徴量を入力とする．この時入力が正しければ，出力として透かし情報を得ることができる．この埋め込み処理は非常に少ない埋込み量，もしくは画像に対して全く情報を埋め込まずに実現することが可能であることを意味する．

バックプロパゲーション学習は基本的に計算コストがかかることが知られている．電子透かしの利用用途を考えると，透かしの埋込み処理の速度は重要ではなく，抽出処理が高速であることが望ましい．このシステムの抽出処理は単純なフィードフォワードによる出力処理であるため，非常に高速である．また，三層パーセプトロンのもつ特性によって，入力ユニットに対するノイズの有無や入力の欠落という問題も，誤り訂正機能とニューロン接続形態により，他ユニットが信号を補完してくれる．このため，出力の閾値の調整をすることで，ある程度の除去攻撃や圧縮処理に対しても耐性があるといえる．

実際に透かしが埋めこめられた画像に対する画像変換・フィルタ後の透かし情報の抽出実験の結果を下図に示す．埋め込んだ透かし情報は10110101という信号である．通常は閾値が0.95付近でも信号を認識できるが，フィルタ処理を施した画像は閾値を0.7付近まで下げる必要がある．しかし，閾値の調整により透かし情報が復元できている．

図8　透かし情報の埋めこめられた画像

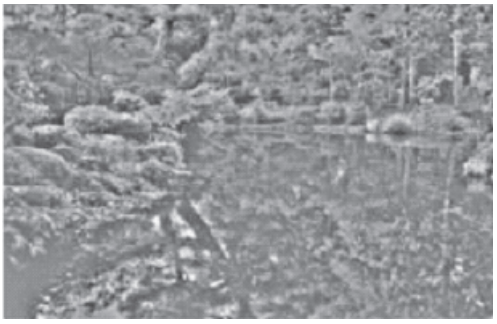


図9　図8に対してハイパスフィルタ処理を施した画像

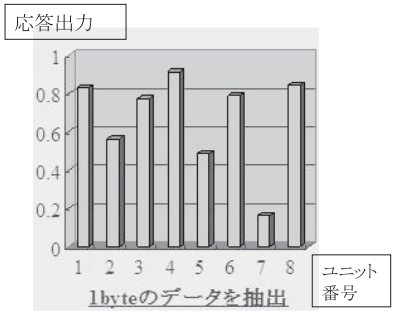


図10　図9から透かし情報を抽出した結果

6. 事例4. 移動物体監視システム

6·1　概　　要

セキュリティ技術には，コンテンツ保護や暗号化などを対象とする情報通信システムのセキュリティと，映像監視や人の入退出管理などを対象とする物理・環境系のセキュリティがある．ここで紹介する事例は物理・環境系のセキュリティ技術の一つに当たる映像監視システムである [Chashikawa 03].

紹介する事例はニューラルネットワークを利用したフレーム間差分法の後処理として用いる移動物体領域の抽出手法である．動画像から移動物体を抽出する方法として，従来からフレーム間差分法が使われている．しかし，フレーム間差分法では移動物体全体が得られないという問題がある．また，撮影環境によっては後処理が有効に機能しない場合も多い．紹介する手法は従来手法に比べてノイズの影響を受けにくく，また抽出領域の欠損が少ないという特徴がある．

6·2　移動物体領域抽出手法

紹介する手法ではニューラルネットワークモデルとしてPulse Coupled Neural Network（PCNN）に基づいた新たなSecond Order Neural Network（SONN）を用いている．一般的なPCNNのニューロンモデルは以下の図で示されるように，Dendritic tree, Linking modulator, Pulse generatorと呼ばれる三つの部分から構成される．

Dendritic treeは，Feeding部とLinking部と呼ばれる二つの部位に分かれており，それぞれLinking modulatorに対して独立した出力を生成する．Linking modulatorはDendritic treeの二つの部位の出力を特

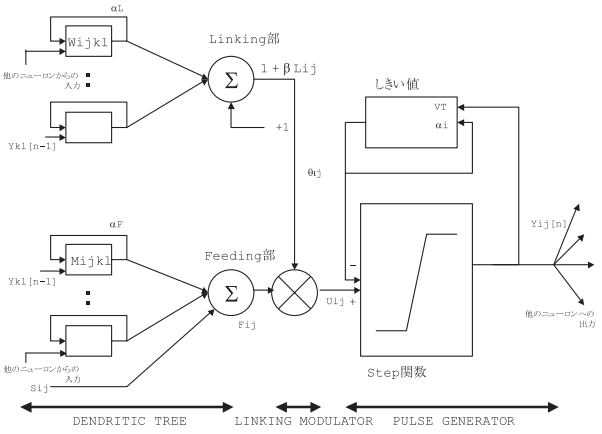


図11　PCNNのニューロンモデル

殊な結合強度に従い結合し内部状態を決定し，Pulse generator は内部状態と動的閾値を比較して出力を決定する．ニューロンがパルスを発生するとしきい値にフィードバックされ，パルスの生成が停止するまでしきい値は急激に増加する．その後，しきい値は再び内部電位の値より小さくなるまで時定数に従って減少する．

このネットワーク一回の動作は，Dendritic tree にてそれぞれの受容野の入力加重和を計算し，Linking 部の出力により Feeding 部の出力の調整を行い，出力パルスを生成させるというものである．この発生されたパルスが，受容野を介してほかのニューロンへ入力される．PCNN モデルは以上のような基本モデルをもっているが紹介するシステムでは，この基本的モデルに対して三つの改良を行っている．

一つは Feeding 部に行った変更である．基本モデルでは Feeding 部は外部刺激と近傍ニューロンの出力を受けているが，ここでは外部刺激と近傍ニューロンの出力の相関を受け取り，外部刺激に過渡応答性をもつような変更が行われている．二つ目の改良は Linking 部に行った変更である．基本モデルでは Linking 部の値は Feeding 部の値を変調する役割を担っているが，ここでは近傍ニューロンが発火していない場合に負の値を出力するような変更がされている．三つ目の改良は閾値に行った変更である．基本モデルの動的しきい値はニューロンに不応期を与えているが，しきい値を定数とすることで固定しきい値をもたせている．

このような構造により，このモデルは SONN モデルの一種と位置づけられるモデルとなっている．SONN はニューロン間の局所的な相互結合によって同期的に発火するセグメンテーション機能を有する．つまり外部刺激が塊として移動しながら与えられる場合，移動方向に生じる同期発火性などにより，その塊の欠損箇所に当たる外部刺激の小さいニューロンも同時発火する仕組みをもつモデルになっている．

移動物体領域の抽出手順としては，カラー動画像から連続フレーム画像を生成し，そのフレーム画像から輝度画像を生成する．正規化した輝度画像をフレーム間差分法で処理し，輝度変化の画像を生成する．その輝度変化の画像を SONN の外部刺激として入力する．下図はその手順を示した図である．

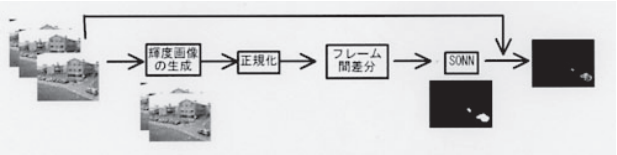


図 12　SONN を使った移動物体領域抽出手順

7. お わ り に

多くのセキュリティ技術は，ニューラルネットワークの特性を利用することでその技術が目的とする問題を解決している．

紙幣識別システムの事例では，精巧な偽札が入力された場合，金種分類部で学習済みのあるクラスとして分類される可能性はあるが，真贋判定部で偽札を拒否することができた．これは分類部と判定部を切り分けたことに大きな意味がある．この分類部では三層パーセプトロンというニューラルネットワーク以外の手法で分類することもできたであろう．しかしながら判定部に関しては，三層パーセプトロンのみでは不適当である．紙幣のサイズの違いというものも考慮しつつ，真性紙幣を受け入れ偽札を拒絶するという未知のデータに対して正しい振舞いをする仕組みは，問題に合わせて適切な構造を有した RBF ネットワークを判別器にしたため解決できたのである．

近年ポリモフィックウイルスやメタモフィックウイルスのように，感染ごとに自分自身のコードの書換えを行うことで亜種を生成するものやコードの暗号化を行うウイルスが出てきている．これらのウイルスの検出は従来のウィルス検出手法のようにウイルスデータベースを用意するやり方では不可能である．このような問題を解決するためにはコードの特徴を表す機械語の学習や特徴語自体をヒューリスティックに探索することで分類や判定を行い，機能としては同じであるがコードの見かけ上は全く異なった亜種ウイルスや未知ウイルスを検出させる必要がある．

電子透かしの例では，透かし情報の抽出方法をある種のパターン分離問題と捉えている．埋込み対象画像の特徴情報学習処理をそのまま抽出鍵の生成処理とさせたことで，コンテンツの品質を損なわずに悪意のある攻撃やフォーマット変換にも耐性を有することができている．

移動物体監視システムにおいては単純なフレーム間差分法を用いた場合，移動物体が実際には移動しているが差分の部分しか検出できないため両端が細切れで検出されてしまうが，紹介した事例ではニューラルネットワークを用いることで，移動物体全体を検出することができ，ノイズに耐性を有することができる．

このように，問題に対する適切なモデルと機能の選択をすることでニューラルネットワークはセキュリティ技術の上においても非常に有効なツールとなり得る．情報セキュリティを例に，さまざまな問題を解決するための手段として今後も幅広くニューラルネットワークが利用されることを期待する．

◇ 参 考 文 献 ◇

[Aihara 90] Aihara, K.: Chaotic neural networks, *Bifurcation Phenomena in Nonlinear System and Theory of Dynamical Systems*, H. Kawakami ed., pp. 143-161, World Scientific（1990）

[Ando 03] Ando, R. and Takefuji, Y.: Location-driven watermark extraction using supervised learning on frequency domain, WSEAS Trans. on computers, Issue 1, Volume 2, ISSN: 1109-2750, pp.163-169（2003.1）

[Aoba 02] Aoba, M. and Takefuji, Y.: Improved Algorithms for Knight's Tour Problems, *Knowledge-Based Intelligent Information Engineering Systems and Allied Technologies*（*KES2002*）, Italy, Crema, Vol. 82, Part 1, pp. 219-223（2002）

[Aoba 03] Aoba, M., Kikuchi, T. and Takefuji, Y.: Euro Banknote Recognition System using a Three-layered Perceptron and RBF Ntworks, 情処学論　数理モデル化と応用（*IPSJ Trans. On Mathematica Modeling and Its Applications*）, Vol. 44, No. SIG 7（TOM 8）, pp. 99-109（2003）

[Broomhead 88] Broomhead, D. and Lowe, D.: Multivariable functional interpolation and adaptive networks, *Complex Systems*, Vol. 2, pp. 321-355（1988）

[Chashikawa 03] Chashikawa, T. and Takefuji, Y.:Extracting Moving Object Areas Based on Second-order Neural Network, *IPSJ,* Vol.44, No.SIG 14（TOM 9）, pp. 31-47（2003）

[Eckhorn 90] Eckhorn, R., Reitboeck, H. J., Arndt, M. and Dicke, P.: Feature linking via synchronization among distributed assemblies: simulations of results from Cat Visual Cortex, *Neural Computing*, Vol. 2, pp. 293-307（1990）

[Elman 94] Elman, J.: Finding structure in time, *Cognitive Science*, Vol. 14, pp. 179-211（1994）

[Falcon] Fair Isaac: Falcon Fraud Manager, http://www.fairisaac.com/Fairisaac/Solutions/Solutions+by+Function/Falcon+Fraud+Manager.htm

[Hebb 49] Hebb, D. O.: *Organization of Behavior*, New York, Wiley（1949）

[Hopfield 85] Hopfield, J. J. and Tank, D. W.: Neural computation of decisions in optimization prolems, *Biological Cybernetics*, Vol. 52, pp. 141-152（1985）

[Iba 01] Iba, T., et al.: Boxed Economy Foundation Model:Model Framework for Agent-Based EconomicSimulations, Terano, T. et al.（Eds.）, *Joint JSAI 2001 Workshop Post-Proceedings*（*LNAI2253*）, Springer, pp. 227-236（2001）

[IPA 04a] 独立行政法人　情報処理推進機構 (IPA): 各国バイオメトリクスセキュリティ動向の調査, http://www.ipa.go.jp/security/fy15/reports/biometrics/index.html, 2004年2月17日.

[IPA 04b] 独立行政法人 情報処理推進機構（IPA）: 未知ウイルス検出技術に関する調査, http://www.ipa.go.jp/security/fy15/reports/uvd/documents/uvd_test.pdf, 2004年4月.

[Johnson 94] Johnson, J. L.: Pulse-coupled neural nets: Translation, rotation, scale, distortion and intensity signal invariances for images, *Applied Optics*, Vol. 33, No. 26, pp. 6239-6253（1994）

[Kawamura 02] Kawamura, S., Yoshida, H., Miura, M. and Abe, M.:Implementation of uniform pseudo random number generator and application to stream cipher based on chaos neural network, *Proc. Int. Conf. Fundamentals of Electronics, Communications and Computer Sciences*, R-18, pp. 4-9（2002）

[Kephart 95] Kephart, J. O., Sorkin, G. B., Arnold, W. C., Chess, D. M. and Tesauro, G. J.: Biologically Inspired Defenses Against Computer Viruses, *Proc. IJCAI' 95, Montreal*, Aug. 19-25（1995）

[Kephart 96] Kephart, J., Chess, D. and White, S.: Neural Networks for Computer Virus Recognition, *IEEE Expert*, Vol. 11, No. 4, pp. 5-6（1996.8）

[Kohonen 82] Kohonen, T.: Self-organized formation of topologically correct feature maps, *Biological Cybernetics*, Vol. 43, pp. 59-63（1982）

[Kohonen 90] Kohonen, T.: The self-organizing map, *Proc. IEEE*, Vol. 78, pp. 1464-1480（1990）

[Kohonen 95] Kohonen, T.: *Self-Organizing Maps*, Springer-Verlag（1995）

[Maxwell 86] Maxwell, T., Giles, C., Lee, T. C. and Chen, H. H.: Nonlinear dynamics of artificial neural systems, *American Institute of Physics* 0094-243x/86/1510299-17, pp. 299-304（1986）

[McCulloch 43] McCulloch, W. S. and Pitts, W. H.: A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133（1943）

[Microsoft 03] Microsoft Research: Office 2003 Security Enhancements, Neural Decision Engine, http://www.microsoft.com/technet/prodtechnol/office/office2003/deploy/secdesn.mspx（2003）

[Moody 89] Moody, J. E. and Darken, C.: Fast learning in networks of locally-tuned processing units, *Neural Computation* Vol.1, pp. 281-294（1989）

[武藤 01] 武藤佳恭，斉藤孝之 監修, 武藤佳恭研究室 編：応用事例ハンドブック ニューラルコンピューティング, 共立出版（2001）

[直江 03] 直江健介，安藤類央，武藤佳恭：周波数領域での非線形適応システムを用いた電子透かしの耐性評価，CSS2003 研究報告（2003）

[Poggio 90] Poggio, T. and Girosi, F.: Regularization Algorithms for Learning that are Equivalent to Multilayer Networks, *Science*, Vol. 247, No. 4945, pp. 978-982（1990）

[Rosenblatt 58] Rosenblatt, F.: The Perceptron: Probabilistic model for information storage and organization in the brain, *Psychology Review* Vol. 65, pp. 386-408（1958）

[Rummelhart 86] Rummelhart, D. E., McClelland, J. L. and the PDP Research Group: *Parallel Distributed Processing*, Vol. 1, The MIT Press（1986）

[Seggern 93] von Seggern, D.: *CRC Standard Curves and Surfaces*, p. 124, CRC Press（1993）

[Symantec 2001]Symantec AVCenter: Symantec White Paper The Digital Immune System, http://www.symantec.com/region/jp/avcenter/reference/dis.tech.brief_j.pdf,（2001）

[Takefuji 89] Takefuji, Y. and Lee, K. C.: A near-optimum parallel planarization algorithm, *Science*, Vol. 245, pp. 1221-1223（1989）

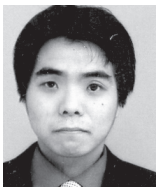
[takefuji 92] Takefuji, Y., Lee, K. C. and Aiso, H.: An artificial maximum neural network: A winner-take-all neuron model forcing the state of the system in a solution domain, *Biological Cybernetics*, Vol. 67, pp. 243-251（1992）

[Tokutaka 99] Tokutaka, H., Kishida, S. and Fujimura, K.: 自己組織化マップの応用―多次元情報の2次元可視化，●●（1999）

[谷荻 94] 谷萩隆嗣, 高野裕昭：カテゴリーを組み合わせたNNによる顔画像認識，信学論（D-II），Vol.J77-D-II, No.11, pp.2151-2159（1994.11）

2006年6月28日　受理

著 者 紹 介



直江　健介

2002 年慶應義塾大学環境情報学部卒業．2004 年慶應義塾大学大学院政策・メディア研究科修士課程修了．修士（政策・メディア）．2004 年～現在，慶應義塾大学大学院政策・メディア後期博士課程．ネットワーク侵入検知，暗号，ウイルス検出に興味をもつ．現在ではニューラルネットワークを用いた音声･動画・画像の電子透かしの埋込み・抽出手法の研究を行っている．

田中　秀和

2005 年慶應義塾大学環境情報学部卒業．2005 年慶應義塾大学大学院政策・メディア研究科修士課程，現在に至る．ウィルス検出・実理証明に興味をもち，現在ではマルウェアに対する有効的な防御手法の研究を行っている．

武藤　佳恭

1978 年慶應義塾大学工学部電気工学科卒業．1980 年慶應義塾大学大学院修士課程電気工学修了．1983 年慶應義塾大学大学院博士課程電気工学修了．博士（工学）．1983 ～ 85 年南フロリダ大学コンピュータサイエンス学科客員助教授．1985 ～ 88 年南カロライナ大学コンピュータサイエンス学科助教授．1988 ～ 97 年ケースウエスタンリザーブ大学電気工学準教授．1992 ～ 97 年慶應義塾大学環境情報学部助教授．1997 年慶應義塾大学環境情報学部教授，現在に至る．NSF-RIA 賞（1989），IEEE Trans. on NN 功労賞（1992），IPSJ 論文（1980），TEPCO 賞（1993），KAST 賞（1993），高柳賞（1995），KDD 賞（1997），NTT tele-education courseware 賞（1999），US-AFOSR 受賞（2003），国際協力機構 JICA 理事長賞（2004），政府顧問 :NCC（フィリピン），VITTI（ベトナム），CTTISC（ジョルダン），タイ，スリランカ，マルチメディア大学（マレーシア），22 冊の本と 200 編以上の科学論文．