

Einige Untersuchungen zu einem Neuronenmodell auf der Basis von dendritischen Spine-Clustern

Holger Arnold

19. Juni 2002

1 Einleitung

Übliche abstrakte Neuronenmodelle betrachten die Neuronen als reine Integrations-Schwellwert-Einheiten, deren Erregung als gewichtete Summe ihrer Eingaben berechnet wird. Dabei bleiben allerdings die komplexe Struktur der Dendriten als Eingabepfade der Neuronen und die Interaktionen der dendritischen Synapsen unberücksichtigt. Wir werden in diesem Artikel ein Neuronenmodell für assoziatives Lernen untersuchen, das auf der Annahme basiert, dass die Assoziation eingehender Signale und damit die Lernfähigkeit der Neuronen auf der Wechselwirkung räumlich benachbarter Synapsen beruht.

2 Das DSC-Neuronenmodell

Die meisten Lernverfahren basieren im weitesten Sinne auf der Hebbischen Lernregel [5], nach der eine synaptische Verbindung verstärkt, also leichter erregbar wird, wenn das presynaptische und das postsynaptische Neuron in zeitlicher Nähe Aktivität zeigen. Die verschiedenen Eingangssynapsen eines Neurons sind nach diesem Modell völlig unabhängig voneinander. Dafür muss eine lokale Rückverbindung vom Ort der Erzeugung der Aktivität (meist der Axonhügel, der sich hinter dem Zellkörper befindet) zur synaptischen Membran angenommen werden.

Über mehrere Stufen haben Alkon, Blackwell und Vogl z.B. in [2], [1] und [3] ein Neuronenmodell für assoziatives Lernen entwickelt, das sich vom Hebb-Modell grundlegend unterscheidet. Experimente mit der Meeresschnecke *Hermissenda* wiesen nach Ansicht der genannten Autoren darauf hin, dass der eigentliche Lernvorgang nicht an einzelnen Synapsen, sondern vielmehr auf kleinen Abschnitten der dendritischen Zellmembran stattfindet. Diese Abschnitte, die sie als *Patches* bezeichnen, bilden sich durch die lokale Wechselwirkung räumlich benachbarter Synapsen auf den Dendriten. Eine Menge von Synapsen die auf diese Weise gekoppelt werden bildet einen *dendritischen Spine-Cluster (DSC)*.

Abbildung 1 stellt ein Neuron nach diesem Modell dar. Auf der oberen Seite des Neurons befinden sich die Dendriten, das Axon verläuft nach unten. Für einen Dendriten sind mehrere Eingangssynapsen dargestellt. Sie befinden sich an Auswölbungen der Dendritenmembran, den sogenannten *dendritischen Spines*. Ziel ist es, einen unbedingten Reiz *UCS* (unconditioned stimulus) mit einer Kombination bedingter Reize CS_i (conditioned stimulus) zu assoziieren. Dazu

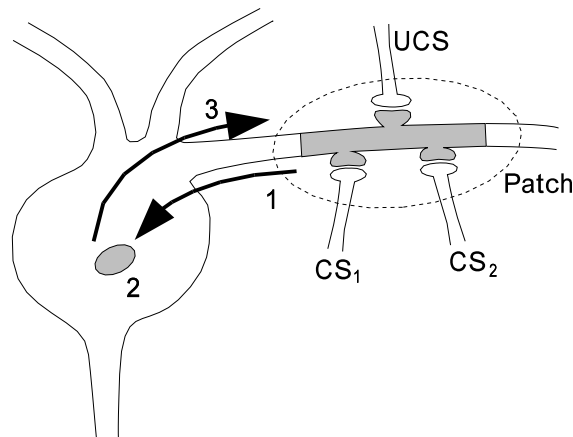


Abbildung 1: Lernen durch lokale Wechselwirkung von Synapsen

sind zunächst synaptische Potentiale an UCS , CS_1 und CS_2 erforderlich, die in Richtung des Zellkörpers weitergeleitet werden (1). Daraufhin kommt es im Zellkern zu Veränderungen, unter anderem kann eine verstärkte RNA-Bildung beobachtet werden (2). Dadurch wird ein Umbau der Zellmembran in der Umgebung der beteiligten Synapsen ausgelöst (3), der zu einer verstärkten Erregbarkeit in diesem Bereich und damit zu einer Speicherung der Assoziation von unbedingtem Reiz und bedingten Reizen führt.

2.1 Der DSC-Lernalgorithmus

Ein einzelnes Neuron nach dem DSC-Modell besitzt eine größere Komplexität als ein Neuron nach dem klassischen Hebb-Modell. Bezüglich der Berechnungsstärke (im Sinne des Kolmogorov-Theorems) ist ein einschichtiges DSC-Netz äquivalent zu einem dreischichtigen Feedforward-Netz nach dem Hebb-Modell.

Abbildung 2 zeigt die ein solches einschichtiges DSC-Netzes. Es besteht aus einer Menge von M Ausgabeneuronen, die untereinander keine Verbindungen besitzen. Jedes Neuron repräsentiert einen unbedingten Reiz (UCS)¹. Die bedingten Reize (CS) die mit einem unbedingten Reiz assoziiert sind, werden durch Patches repräsentiert, die sich während des Lernens herausbilden. Die Menge der bedingten Reize von denen ein Neuron Eingaben erhält, ist eine Teilmenge der Gesamteingaben des Netzes, und wird als *rezeptives Feld* des Neurons bezeichnet. Diese Teilmenge ist für alle Patches eines Neurons gleich, kann sich aber zwischen verschiedenen Neuronen unterscheiden.

Jeder Patch i eines Neurons k besteht aus folgenden Komponenten:

- Dem Patchvektor $P_i = (P_{ij})$, dem arithmetischen Mittel aller CS -Eingabevektoren, die diesem Neuron bisher präsentiert wurden und die zu einer Aktivierung dieses Patches geführt haben.
- Dem arithmetischen Mittel U_{ik} der UCS -Eingabewerte, die diesem Neuron bisher präsentiert wurden und die zu einer Aktivierung dieses Patches

¹Bezogen auf das Problem der Musterklassifikation könnte ein einzelner unbedingter Reiz z.B. für eine bestimmte Klasse von Mustern stehen.

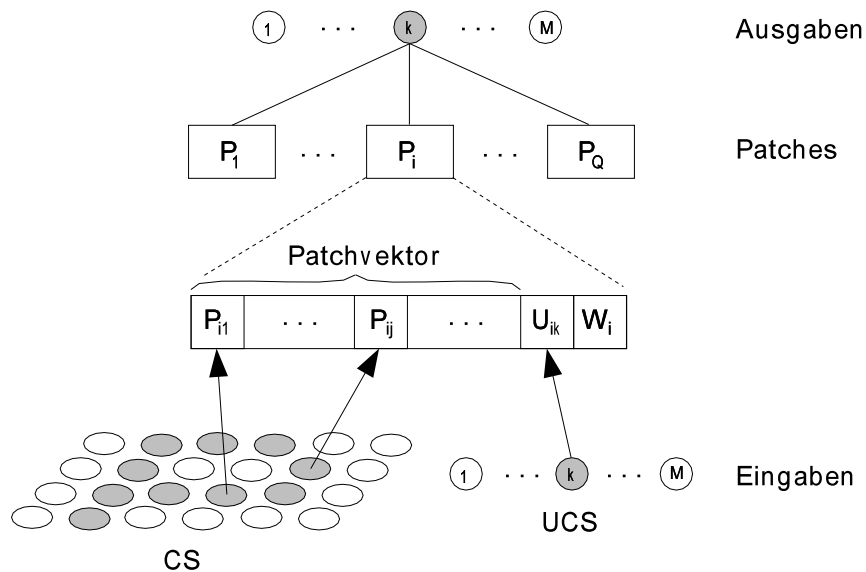


Abbildung 2: Einschichtiges DSC-Netz

geführt haben.

- Einem Gewicht W_i , das aus der Häufigkeit der Aktivierung des Patches gebildet wird.

Zu Beginn des Lernprozesses sind zunächst noch keine Patches vorhanden, diese werden erst während des Lernens nach Bedarf erzeugt. Jedes Neuron lernt unabhängig von den anderen Neuronen. Dabei führt jedes Neuron folgende Schritte aus:

- Die Eingabe CS des Neurons wird mit allen Patchvektoren P_i der Patches des Neurons verglichen. Für jeden Patch i wird eine Größe s_i mit $0 \leq s_i \leq 1$ und $CS = P_i \rightarrow s_i = 1$ bestimmt, die ein Maß für die Ähnlichkeit von CS und P_i ist.
- Es wird ein Patch m gesucht, der folgende Bedingungen erfüllt:
 1. Der Patchvektor P_m ist von allen Patchvektoren dem Vektor CS am ähnlichsten ist: $\forall i. s_m \geq s_i$.
 2. Die Ähnlichkeit zwischen CS und P_m ist größer als ein vorgegebener Schwellwert: $s_m \geq t_{CS}$
 3. Der gespeicherte Wert U_{mk} liegt in der Nähe des Eingabewertes UCS : $|UCS - U_{mk}| \leq t_{UCS}$
- Wird ein solcher Patch gefunden, wird dieser Patch aktiviert und P_m und U_{mk} werden aktualisiert.
- Gibt es keinen Patch, der diese Bedingungen erfüllt, wird ein neuer Patch n mit Patchvektor $P_n = CS$ und $U_{nk} = UCS$ erzeugt.

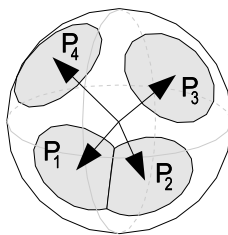


Abbildung 3:

- Die Ausgabe N_k des Neurons kann auf zwei Arten berechnet werden: 1. nach dem Prinzip Winner-takes-all: $N_k = s_m U_{mk}$, 2. als gewichtete Summe der Beiträge der einzelnen Patches: $N_k = \sum_i W_i s_i U_{ik}$.

2.2 Interpretation des Modells

Der oben beschriebene Algorithmus kann als selbstorganisierender Clustering-Algorithmus interpretiert werden, bei dem die Anzahl der Cluster nicht vorher festgelegt ist, sondern dynamisch bestimmt wird. Abbildung 3 veranschaulicht die Situation: sie zeigt vier n -dimensionale Vektoren $P_1 \dots P_4$ und eine n -dimensionale Kugeloberfläche. n ist dabei die Dimension der CS -Vektoren. Jeder der vier Vektoren entspricht dem Patchvektor eines Patches². Die markierten Gebiete auf der Kugeloberfläche stellen $(n-1)$ -dimensionale Schnitte durch die Einflussbereiche der darin liegenden Patchvektoren dar. Die Einflussbereiche der Patches sind disjunkt. Ihre maximale Ausdehnung (im Sinne des Ähnlichkeitsmaßes s) wird bestimmt durch die Größe t_{CS} . Zu beachten ist, dass es vom verwendeten Ähnlichkeitsmaß abhängt, welche Topologie die Einflussbereiche besitzen (ob die Gebiete z.B. wie in der Abbildung zusammenhängend sind oder nicht).

Ein CS -Eingabevektor aktiviert denjenigen Patch eines Neurons, in dessen Einflussgebiet er sich befindet. Dieser wird daraufhin in Richtung des Vektors CS verschoben. Liegt CS nicht im Einflussbereich eines Vektors oder passt der gespeicherte UCS -Wert nicht zur Eingabe, wird ein neuer Patch mit Patchvektor CS erstellt. Dadurch können Einflussbereiche anderer Vektoren verkleinert werden. In der Abbildung ist das an der Grenzlinie zwischen P_1 und P_2 zu erkennen.

Während beim klassischen Hebb-Modell die Klassifikation von Punktmengen durch Separation mit Hyperebenen erfolgt, werden beim DSC-Modell Repräsentanten für Punktmengen gespeichert. Die Einflussgebiete dieser Repräsentanten werden nicht explizit gespeichert, sondern ergeben sich aus der Gesamtmenge der gespeicherten Repräsentanten und der Größe t_{CS} . Das hat zwei bemerkenswerte positive Konsequenzen: Erstens ist der Aufwand, der erforderlich ist, um gegebene Punktmengen zu klassifizieren unabhängig von der Dimension der Vektoren. Zweitens erfolgt eine automatische Adaption der Einflussbereiche der Repräsentanten während des Lernens, d.h. in Gebieten, in denen Punkte verschiedener Klassen eng benachbart liegen, sind die Einflussbereiche kleiner als in Gebieten, in denen Punkte verschiedener Klassen weiter auseinander liegen.

²Die Patchvektoren müssen nicht notwendigerweise die gleiche Länge besitzen, es findet keine Normalisierung der Vektoren statt.

2.3 Ähnlichkeitsmaße

Es ist klar, dass der Wahl des Ähnlichkeitsmaßes s entscheidende Bedeutung für die Leistungsfähigkeit des DSC-Modells zukommt. Das gewählte Maß bestimmt die Topologie der Einflussbereiche der Patchvektoren und trägt nicht unerheblich zum Rechenaufwand des Verfahrens bei. In diesem Abschnitt werden einige Maße vorgestellt, die für die Experimente implementiert wurden. Für die Ähnlichkeit $s(x, y)$ zweier Vektoren x und y muss gelten $0 \leq s(x, y) \leq 1$ und $x = y \rightarrow s(x, y) = 1$. Je ähnlicher sich zwei Vektoren sind, desto näher sollte s an 1 liegen.

Summe der Elemente

Ein einfaches Maß für die Ähnlichkeit zweier Vektoren x und y ist das Verhältnis der Summen ihrer Elemente.

$$s_x = \sum_i x_i, \quad s_y = \sum_i y_i, \quad d_1 = \frac{s_x}{s_y}, s_y \neq 0.$$

d_1 wird zunächst so transformiert, dass der Bereich $[\frac{1}{r}, r]$, $r \in \mathbb{R}$ auf den Bereich $[-1, 1]$ abgebildet wird (wobei $1 \mapsto 0$ gilt) und anschließend auf diesen Bereich beschränkt:

$$d_2 = \log_r d_1, \quad d_3 = \max\{d_2, -1\}, \quad d_4 = \min\{d_3, 1\}.$$

Die Ähnlichkeit der Vektoren wird dann bestimmt als

$$s_{sum} = \left(\frac{1}{2} \cos(\pi d_4) + \frac{1}{2} \right)^2.$$

Winkeldistanz

Der Winkel ϕ zwischen zwei Vektoren x und y ist definiert als

$$\varphi = \arccos \frac{xy}{\|x\| \|y\|}.$$

Die Ähnlichkeit der Vektoren ergibt sich dann durch Normalisierung von φ als

$$s_{ang} = 1 - \frac{\varphi}{\pi}.$$

Korrelationskoeffizient

Zur Bestimmung des Korrelationskoeffizienten zweier Vektoren x und y werden zunächst die Summen s_{xx} , s_{xy} und s_{yy} bestimmt:

$$\begin{aligned} s_{xx} &= \sum_i (x_i - \bar{x})^2 \\ s_{yy} &= \sum_i (y_i - \bar{y})^2 \\ s_{xy} &= \sum_i (x_i - \bar{x})(y_i - \bar{y}). \end{aligned}$$

Dabei sind \bar{x} und \bar{y} die arithmetischen Mittelwerte der Elemente von x und y . Der Korrelationskoeffizient s ist dann definiert als

$$s_{corr} = \frac{s_{xy}^2}{s_{xx}s_{yy}}.$$

Normalisierte Euklidische Distanz

Dieses Ähnlichkeitsmaß wird folgendermaßen definiert:

$$s_{Euklid} = 1 - \frac{\sqrt{\sum_i (x_i - y_i)^2}}{\|x\| + \|y\| + \varepsilon}, \quad 0 < \varepsilon \ll (\|x\| + \|y\|).$$

Die Größe ε dient nur dazu, eine Division durch 0 zu verhindern, falls $x = y = 0$ gilt.

Ein biologisch motiviertes Modell

Wie die Autoren des DSC-Modells in [3] und [4] gezeigt haben, kann die Aktivierung eines bestimmten Patches als Reaktion auf einen Eingabevektor durchaus biologisch erklärt werden. Dazu wurde unter anderem ein Schaltungsmodell eines Dendritenabschnittes analysiert und daraus Aussagen darüber abgeleitet, wie sich das Potential der Dendritenmembran in Abhängigkeit von der synaptischen Erregung verhält. Dabei treten folgende charakteristische Effekte auf:

- Das größte Membranpotential wird erreicht, wenn die synaptischen Eingaben x_i genau den vorher gespeicherten Werten x_i^0 entsprechen.
- Ein *Overmatch* (die synaptische Erregung x_i eines Spines ist stärker als der gespeicherte Wert x_i^0) hat einen stärkeren Effekt auf das Membranpotential als ein *Undermatch* (die synaptische Erregung x_i eines Spines ist schwächer als der gespeicherte Wert x_i^0).
- Ein Overmatch eines Spines kann den Effekt eines Undermatch eines anderen Spines maskieren, so dass dieser nicht zum tragen kommt.
- Eine Abweichung der synaptischen Erregung x_i vom gespeicherten Wert x_i^0 eines Spines hat einen stärkeren Effekt, wenn x_i^0 groß gegenüber den gespeicherten Werten der anderen Spines des Patches ist (d.h. $\forall j. x_i^0 \geq x_j^0$), als wenn x_i^0 klein ist.

Da die Lösung der vollständigen Modellgleichungen für die Verwendung in einem neuronalen Netz zu rechenaufwändig wäre, wird deren Effekt durch folgende vereinfachte Formeln simuliert. Zunächst wird für jedes Spine i (d.h. für jedes Element der Vektoren) dessen Beitrag sc_i zum Membranpotential bestimmt. Dazu wird für jedes Element das Verhältnis des Eingabewertes zum gespeicherten Wert bestimmt. Anschließend erfolgt eine Transformation des Intervalls $[\frac{1}{r}, r]$, $r \in \mathbb{R}$ auf das Intervall $[-1, 1]$ (wobei $1 \mapsto 0$ gilt) und eine Beschränkung auf dieses Intervall:

$$d_1 = \frac{x_i}{x_i^0}, \quad d_2 = \log_r d_1, \quad d_3 = \min\{d_2, 1\}, \quad d_4 = \max\{d_3, -1\}.$$

Der Beitrag sc_i des Spines i zum Membranpotential ist dann gegeben durch

$$sc_i = \left(\frac{1}{2} \cos(\pi d_4) + \frac{1}{2} \right)^2.$$

Wichtig bei diesem Modell ist, dass Abweichungen bei der synaptischen Eingabe eines Spines den Beitrag der anderen Spines des Patches beeinflusst. Dementsprechend werden als nächstes die Überstimmungswerte (*shunting values*) der Spines bestimmt:

$$w_{ovr} = \max_{i:x_i > x_i^0} \left\{ \frac{x_i - x_i^0}{x_i + 10x_i^0} \right\}$$

$$w_{und} = \max_{i:x_i < x_i^0} \left\{ \frac{x_i^0 - x_i}{10x_i + x_i^0} \log \left(\frac{x_i^0}{\bar{x}^0} \right) \right\},$$

wobei \bar{x}^0 der Mittelwert des Vektors x^0 der gespeicherten Werte des Patches ist. Die Ähnlichkeit der n -dimensionalen Vektoren x und x^0 wird dann berechnet als

$$s_{DSC} = (1 - \max\{w_{ovr}, w_{und}\}) \frac{1}{n} \sum_i^n sc_i.$$

Zu beachten ist, dass in diesem Modell nur mit positiven synaptischen Eingaben gearbeitet werden kann (d.h. $\forall i. x_i^0 \neq 0, x_i \neq 0$), was auch der biologischen Realität entspricht. Kann das nicht garantiert werden, muss das Modell entsprechend angepasst werden.

3 Experimente und Ergebnisse

Das Verhalten eines einschichtigen Netzes nach dem DSC-Modell wurde mit mehreren Benchmarks untersucht, deren Ergebnisse im Folgenden vorgestellt werden.

3.1 Spiralenproblem

Das Spiralenproblem ist ein Standardbenchmark für neuronale Netze, bei dem es darum geht, die Punkte der Ebene so in zwei Klassen einzuteilen, dass sich daraus zwei ineinander gewundene Spiralen ergeben. Die Trainingsmenge besteht aus insgesamt 194 Punkten (x_i, y_i) mit zugeordneten Klassen $c(x_i, y_i)$. Die Punkte werden erzeugt durch

$$\varphi_i := \frac{i\pi}{16}, \quad r := \frac{6.5(104 - i)}{104},$$

$$x_i := r_i \sin \varphi_i$$

$$y_i := r_i \cos \varphi_i, \quad \text{jeweils für } i \in [0, 96],$$

und die zugehörigen Klassen sind

$$c(+x_i, +y_i) := +1 \quad \text{und}$$

$$c(-x_i, -y_i) := -1.$$

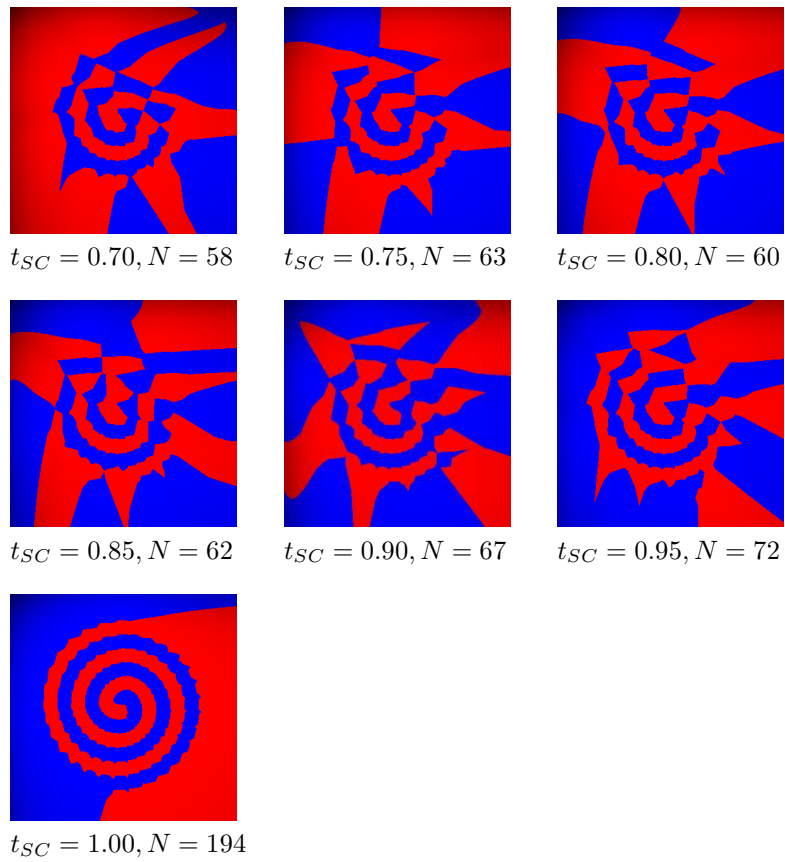


Abbildung 4: Klassifikationsergebnisse des Spiralenproblems bei Verwendung des Ähnlichkeitsmaßes s_{DSC}

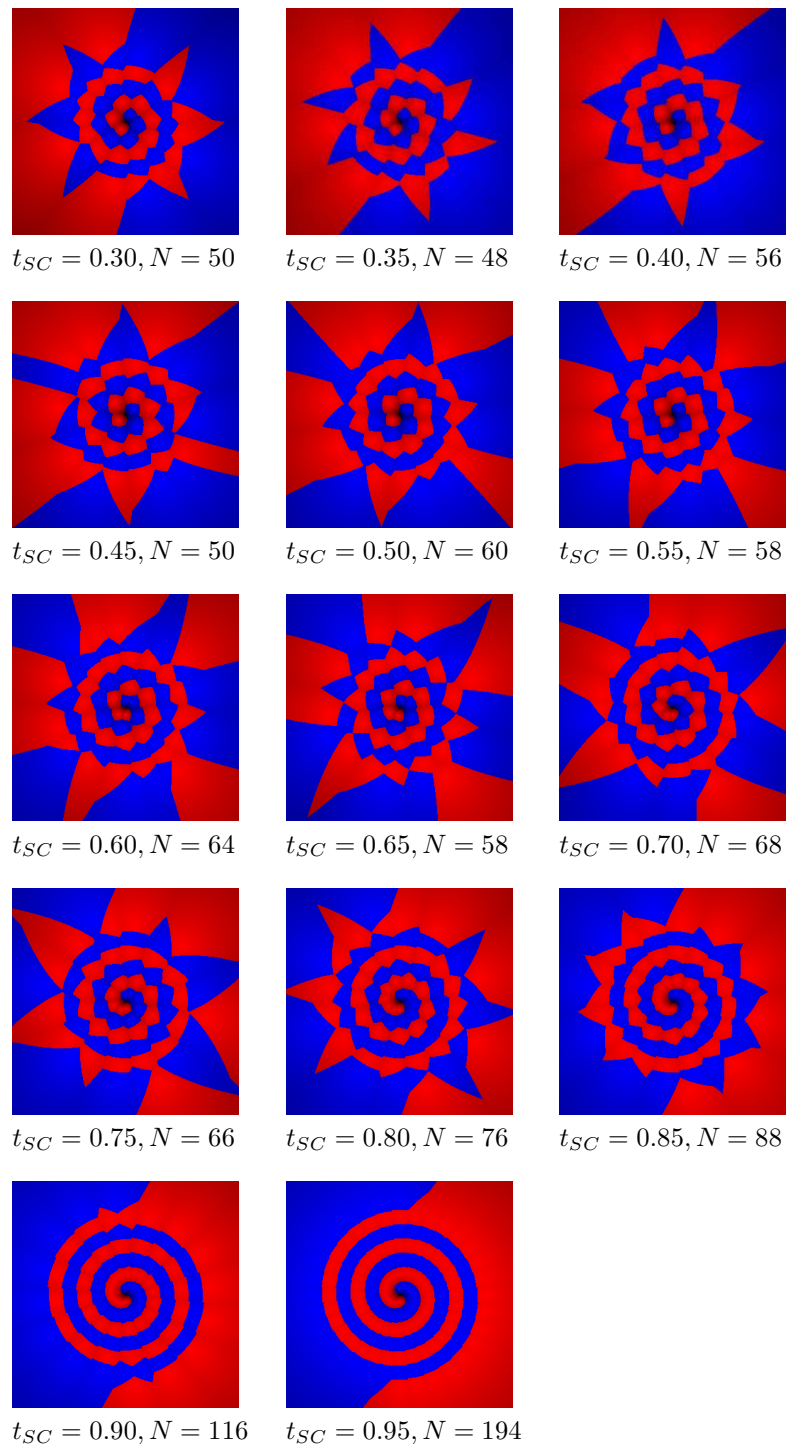


Abbildung 5: Klassifikationsergebnisse des Spiralenproblems bei Verwendung des Ähnlichkeitsmaßes s_{Euklid}



Abbildung 6: Testmuster des Buchstaben „A“ in drei Schriftarten

Die Auswertung der Klassifikationsleistung eines Netzes erfolgt dann durch Darstellung eines Ausschnitts der Ebene, so dass jedem Punkt eine Farbe zugewiesen wird, die der Ausgabe des Netzes mit diesem Punkt als Eingabe entspricht. Eine Klassifikation wird als gut angesehen, wenn in der Ebene zwei deutlich voneinander abgegrenzte Spiralen mit möglichst glatten Grenzlinien erkennbar sind.

Für das Spiralenproblem ist ein DSC-Netz mit nur einem Ausgabeneuron und zwei Eingaben erforderlich. Das Vorzeichen der Ausgabe des Netzes bestimmt die dem Eingabepunkt zugeordnete Klasse. Von den oben beschriebenen Ähnlichkeitsmaßen sind hier nur s_{Euklid} und s_{DSC} sinnvoll. Es ist intuitiv sofort klar, dass s_{Euklid} quasi das natürliche Ähnlichkeitsmaß für dieses Problem darstellt und deshalb die bessere Klassifikation liefern sollte.

Die Klassifikationsergebnisse sind in Abbildung 4 (für s_{DSC}) und Abbildung 5 (für s_{Euklid}) dargestellt. Unter jedem Bild sind der Ähnlichkeits-Schwellwert zur Erzeugung neuer Patches t_{SC} und die Gesamtanzahl der erzeugten Patches N angegeben. Es ist zu erkennen, dass s_{Euklid} die erwartete gute Klassifikation liefert. s_{DSC} arbeitet in allen Fällen schlechter als s_{Euklid} , dennoch sind die Ergebnisse brauchbar.

3.2 Zeichenerkennung

Ein wichtiges Anwendungsgebiet für neuronale Netze ist die Erkennung von Zeichen. Wir wollen hier nur die Eigenschaften einschichtiger DSC-Netze und nicht die Vor- und Nachteile verschiedener Architekturen untersuchen. Daher ignorieren wir zusätzliche Anforderungen wie etwa die Unabhängigkeit des Erkennungsprozesses gegenüber bestimmten Transformationen der Eingabedaten. Stattdessen konzentrieren wir uns auf die Aufgabe, eine Menge hochdimensionaler, stark abhängiger Vektoren in vorgegebene Klassen einzuteilen. Dabei können auch Vektoren einer Klasse zugeordnet werden, die völlig unzusammenhängend sind (z.B. wird der Buchstabe „A“ immer der gleichen Klasse zugeordnet, unabhängig davon, in welchem Schriftstil er gegeben ist).

Die Trainingsdaten für dieses Experiment bestehen aus Darstellungen der Buchstaben „A“ bis „G“ in drei verschiedenen Schriftarten (siehe Abbildung 6). Die Trainingsmuster haben eine Größe von 64x64 Pixeln in 256 Graustufen. Für jeden Buchstaben wird ein Ausgabeneuron verwendet. Zur Beurteilung der Leistungsfähigkeit des DSC-Netzes und der verschiedenen Ähnlichkeitsmaße wurde das trainierte Netz mit verrauschten Bildern der gelernten Buchstaben getestet. Abbildung 7 zeigt die verschiedenen Verrauschungsgrade für einen Buchstaben.

Einige Resultate des Experiments sind in den Abbildungen 8–10 dargestellt. Für die verschiedenen Ähnlichkeitsmaße und verschiedene Werte für t_{SC} wurde mit steigendem Rauschanteil in den Eingabedaten bestimmt, wie viele von insge-

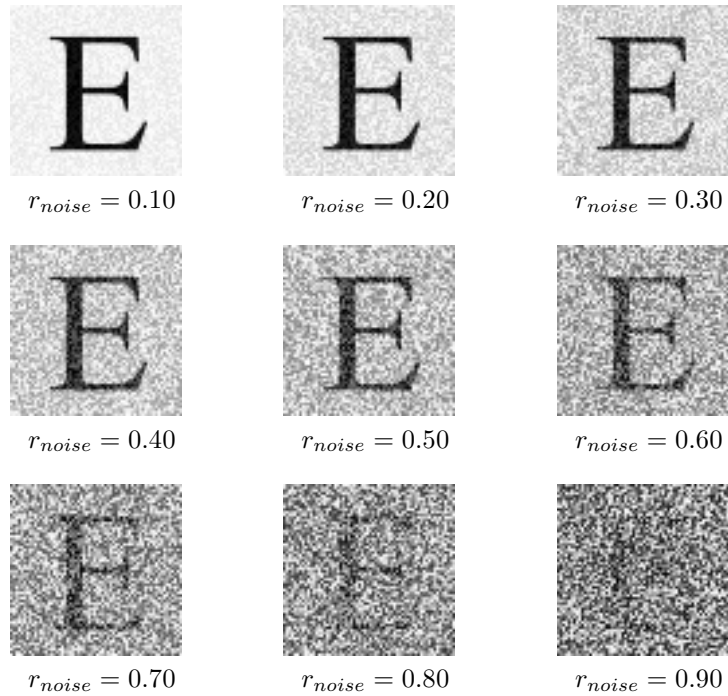


Abbildung 7: Eingabemuster mit wachsendem Rauschanteil

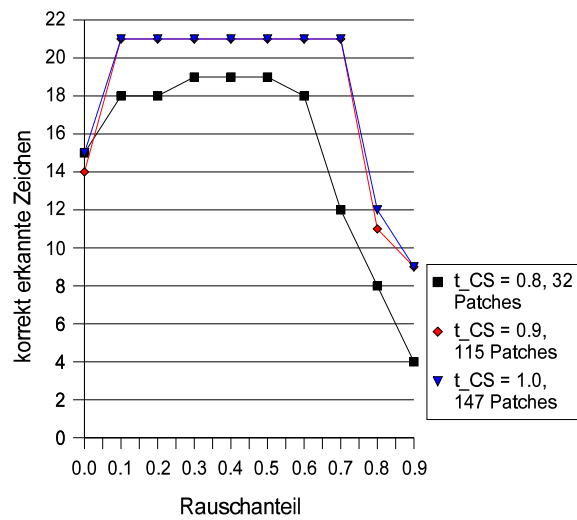


Abbildung 8: Korrekt erkannte Zeichen für s_{ang} (Winkeldistanz)

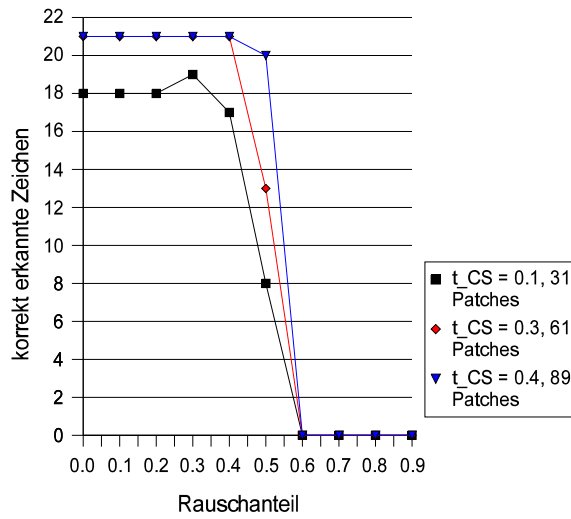


Abbildung 9: Korrekt erkannte Zeichen für s_{corr} (Korrelationskoeffizient)

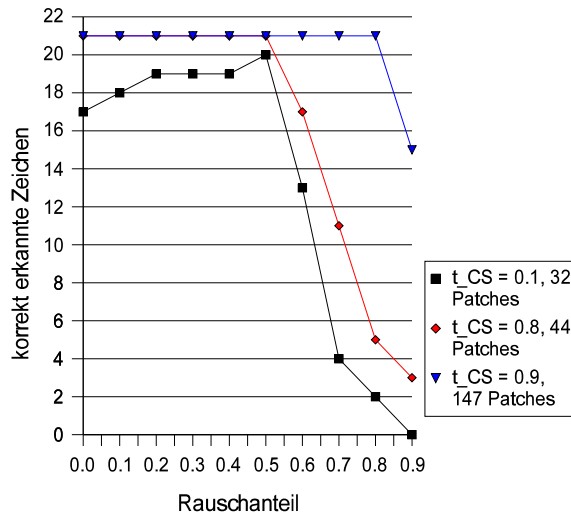


Abbildung 10: Korrekt erkannte Zeichen für s_{Euklid} (Euklidische Distanz)

samt 21 Mustern noch sicher erkannt werden. Dabei galt ein Muster genau dann als sicher erkannt, wenn nur das diesem Muster entsprechende Ausgabeneuron eine Aktivierung von mehr als 0.5 zeigte. Die Ergebnisse dieses Experiments können folgendermaßen zusammengefasst werden:

- Die Erkennung der Zeichen ist ziemlich robust gegenüber zufälligen Störungen der Eingabedaten. Mit entsprechendem Aufwand, d.h. mit einer ausreichenden Anzahl von Patches können z.B. mit s_{Euklid} noch Zeichen sicher erkannt werden, wenn die Eingabedaten zu 80 % aus Rauschen bestehen. Wesentlich weniger Patches sind erforderlich um immerhin noch Störungen von bis zu 50 % kompensieren zu können.
- Ein interessanter und überraschender Effekt ist, dass kleine Störungen offenbar den Erkennungsprozess positiv beeinflussen können. Deutlich zu erkennen ist dieser Effekt z.B. bei s_{ang} in Abbildung 8. Durch die Störung der Eingabemuster gehen offenbar zunächst die Komponenten verloren, die nicht so charakteristisch für ein Muster sind, so dass durch ein leichtes Rauschen von z.B. 10 % tatsächlich die Unterscheidung ansonsten ähnlicher Muster erleichtert wird.
- Enttäuschend war die Erkennungsleistung bei Verwendung von s_{DSC} . Bereits bei sehr kleinem t_{SC} von 0.1 wurde die maximale Anzahl Patches erzeugt. Störungen konnte dieses Verfahren überhaupt nicht kompensieren, kein einziges gestörtes Muster wurde erkannt. Da es sich ansonsten um ein sinnvoll modelliertes Verfahren handelt, ist zu vermuten, dass s_{DSC} nur in einem begrenzten Parameterbereich funktioniert.

4 Zusammenfassung

Das Modell der dendritischen Spine-Cluster ist eine interessante Alternative zu üblichen Modellen neuronaler Netze. Es erklärt in plausibler Weise, wie einzelne Muster durch die Interaktion räumlich benachbarter Synapsen gespeichert werden können. In [6] beschreiben die Autoren ein auf diesem Ansatz basierendes allgemeineres Modell der kortikalen Informationsverarbeitung.

Literatur

- [1] Daniel L. Alkon, Kim T. Blackwell, Garth S. Barbour, Susan A. Werness und Thomas P. Vogl: *Biological Plausibility of Synaptic Associative Memory Models*. Neural Networks, 7[6/7]:1005–1017, 1994.
- [2] D. L. Alkon, K. T. Blackwell, G. S. Barbour, A. K. Rigler und T. P. Vogl: *Pattern-Recognition by an Artificial Network Derived from Biologic Neuronal Systems*. Biological Cybernetics, 62:363–376, 1990.
- [3] K. T. Blackwell, T. P. Vogl und D. L. Alkon: *Pattern matching in a model of dendritic spines*. Network: Computation in Neural Systems, 9:107–121, 1998.
- [4] K. T. Blackwell und K.-H. Yang: *Analog Pattern Matching in a Model of Dendritic Spines*. Technischer Bericht, Institute for Computational Sciences and Informatics, George Mason University, 1998.

- [5] Donald Hebb: *The Organization of Behavior*. John Wiley, New York, 1949.
- [6] Thomas P. Vogl, Kim T. Blackwell und Daniel L. Alkon: *Self-Organization of Cortical Information Processing*. In: *Proceedings of the fifth Appalachian Conference on Behavioral Neurodynamics*, Radford College, Radford, VA, October 1996.