



Universidade Estadual de Campinas

Instituto de Matemática, Estatística e Computação Científica

Departamento de Matemática Aplicada

Memórias Associativas Dinâmicas Baseadas em Números Hipercomplexos

Marcos Eduardo Ribeiro do Valle Mesquita

E-mail: valle@ime.unicamp.br

Resumo

Memórias associativas são modelos matemáticos inspirados na capacidade do cérebro humano de armazenar e recordar por associação. Sobretudo, assim como o cérebro, uma memória associativa também deve ser capaz de recordar um certo padrão mesmo quando uma versão incompleta ou ruidosa de um item armazenado é apresentado como entrada. A rede de Hopfield e as redes recorrentes por correlação de Chiueh e Goodman são exemplos de rede neurais dinâmicas que podem ser usadas para implementar memória associativas para o armazenamento de padrões binários. Nesse projeto de pesquisa investigaremos generalizações da rede de Hopfield, incluindo as redes recorrentes por correlação, para padrões com valores representados por números hipercomplexos, incluindo números complexos e quatérnios. O desempenho dos novos modelos serão avaliados, por exemplo, através de experimentos computacionais com imagens em tons de cinza e coloridas.

Palavras-chave: Inteligência computacional, memórias associativas, redes neurais artificiais, números hipercomplexos, processamento de imagens e sinais.

Campinas, 6 de agosto de 2014.

Metas Atingidas em Relação à Proposta Anterior

No projeto de pesquisa anterior, intitulado “*Estudo sobre Modelos de Memórias Associativas para Imagens Coloridas*”, investigamos modelos de memórias associativas morfológicas (MAMs) para imagens coloridas. Conforme proposto no cronograma de execução, desenvolvemos modelos com melhores tolerância a ruído e de menor custo computacional. Além disso, foram investigadas aplicações das novas MAMs para reconstrução de grandes imagens coloridas. De fato, com o apoio do CNPq, foram publicados ou aceitos para publicação os seguintes artigos em revistas conceituadas de circulação internacional:

1. M.E. Valle: A Robust Subspace Projection Autoassociative Memory Based on the M-Estimation Method, *IEEE Transactions on Neural Networks and Learning Systems*, Volume 25, Issue 07, July 2014. Page(s): 1372 - 1377. (Qualis A1 - Matemática; JCR 4.37).
2. M.E. Valle and P. Sussner: Quantale-based autoassociative memories with an application to the storage of color images, *Pattern Recognition Letters*, Volume 34, Issue 14, October 2013. Page(s): 1589 - 1601 (Qualis A2 - Matemática; JCR 1.06).
3. M.E. Valle and D.M. Grande Vicente: Sparsely Connected Autoassociative Lattice Memories with an Application for the Reconstruction of Color Images, *Journal of Mathematical Imaging and Vision*, Volume 44, Issue 3, November 2012. Page(s): 195 - 222. (Qualis B1 - Matemática; JCR 2.33).
4. M.E. Valle: Complex-Valued Recurrent Correlation Neural Networks, *IEEE Transactions on Neural Networks and Learning Systems*, Accepted for Publication, July 2014. (Qualis A1 - Matemática; JCR 4.37).

Com respeito à formação de recursos humanos, foram concluídas 7 orientações de iniciação científica e a orientação de mestrado em matemática aplicada pela Universidade Estadual de Londrina de Rodrigo V. Costa, com o título “*Memórias Associativas Morfológicas Brouwerianas*”. Por fim, cabe ressaltar que não foram listados acima artigos publicados em revistas com menos prestígio ou de circulação nacional. Também não foram listados acima trabalhos completos publicados em anais de congresso. Contudo, tais informações podem ser encontradas no *Currículo Lattes* do proponente.

1 Introdução Geral

Memórias associativas (AMs, *associative memories*) são modelos matemáticos inspirados na habilidade do cérebro humano de armazenar e recordar informações [4, 9, 11, 26]. Tais como as redes neurais biológicas, uma AM deve ser capaz de recuperar uma informação memorizada mesmo diante da apresentação de um item incompleto ou distorcido. Aplicações de AMs incluem classificação [47, 59], previsão [48, 49, 50, 46], reconhecimento de padrões [57, 58], controle e guiagem autônoma [8, 27, 28, 45].

Modelos de AMs têm sido investigados por inúmeros pesquisadores desde a metade do século passado [13]. Com efeito, pesquisas em AMs iniciaram com os trabalhos de Steinbuch sobre as memórias associativas matriciais [44]. Em 1972, Anderson, Kohonen e Nakano introduziram independentemente um modelo de memória associativa linear em que o armazenamento por correlação, também referido como aprendizado de Hebb, é usado para sintetizar a matriz dos pesos sinápticos [2, 25, 38]. A memória associativa linear ótima (OLAM, *optimal linear associative memory*), que representa a melhor memória associativa linear no sentido dos quadrados mínimos, é obtida substituindo o aprendizado de Hebb pelo armazenamento por projeção [26]. Apesar da ótima capacidade absoluta de armazenamento no caso auto-associativo, a OLAM apresenta uma frágil tolerância a ruído pois corresponde à um modelo linear.

Em 1982, Hopfield introduziu um modelo de AM não-linear para padrões binários [18]. A rede de Hopfield é implementada por uma rede neural recursiva de camada única constituída por neurônios de McCulloch e Pitts [13]. Ela possui muitas características atraentes incluindo: fácil implementação em *hardware*, caracterização em termos de uma função energia e uma variedade de aplicações, incluindo visão computacional [41] e otimização [17, 43, 42, 33].

Apesar de suas características atraentes, a rede de Hopfield é projetada para armazenar e recordar padrões binários. Todavia, muitas aplicações de AMs, incluindo a recuperação de imagens em tons de cinza ou coloridas, requerem o armazenamento de vetores com valores hipercomplexos [15]. Com efeito, diferente dos modelos reais, uma rede neural com valores hipercomplexos processa dados multi-dimensionais como uma única entidade [7, 39, 15]. Em particular, aplicações de redes neurais baseadas em números hipercomplexos inclui controle [7, 3], previsão [6, 34] e processamento de imagens [31, 35, 53, 56]. Em vista dessa observação, nesse projeto de pesquisa propomos investigar modelos de AM descritas por redes neurais artificiais recorrentes baseadas com números hipercomplexos, em particular, números complexos e quatérnios.

2 Identificação e Caracterização do Problema

2.1 Redes Neurais Recorrentes e Memórias Associativas

Primeiramente, lembramos que uma rede neural recorrente, também chamada rede neural dinâmica [10, 12, 14], pode ser usada para definir uma sequência $\{\mathbf{x}(t)\}$ através da equação de diferenças

$$x_j(t+1) = \psi_j(\mathbf{x}(t)), \quad \forall t = 0, 1, \dots, \quad (1)$$

em que $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ representa o estado da rede no tempo t e ψ_j é uma função com valores reais ou hipercomplexos para cada índice $j = 1, \dots, n$. Um vetor coluna $\mathbf{u} = [u_1, u_2, \dots, u_n]^T$ é chamado estado estacionário ou ponto fixo da rede neural se $u_j = \psi_j(\mathbf{u})$ para todo $j = 1, \dots, n$.

Uma função com valores reais E no espaço dos estados é chamada *função de Lyapunov* ou *função energia* da rede neural recorrente descrita por (1) se E é limitada inferiormente e

$E(\mathbf{x}(t+1)) < E(\mathbf{x}(t))$ sempre que $\mathbf{x}(t+1) \neq \mathbf{x}(t)$. A existência de uma função energia garante que a sequência $\{\mathbf{x}(t)\}_{t \geq 0}$ dada por (1) converge para um estado estacionário independente do estado inicial $\mathbf{x}(0)$ [16]. Sobretudo, para qualquer entrada \mathbf{x} , a rede neural fornece uma aplicação \mathcal{M} definida por

$$\mathcal{M}(\mathbf{x}) = \lim_{t \rightarrow \infty} \mathbf{x}(t), \quad (2)$$

em que $\{\mathbf{x}(t)\}_{t \geq 0}$ é a sequência dada por (1) com o estado inicial $\mathbf{x}(0) = \mathbf{x}$. Interessantemente, a aplicação \mathcal{M} pode ser usada para realizar uma memória associativa (AM, *associative memory*).

Uma AM é um sistema capaz de armazenar e, posteriormente, recordar um conjunto de vetores $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^p\}$, chamado *conjunto das memórias fundamentais* [26, 9, 4]. Formalmente, uma AM corresponde a uma aplicação \mathcal{M} , chamada *aplicação associativa*, tal que $\mathcal{M}(\mathbf{u}^\xi) = \mathbf{u}^\xi$ para todo $\xi = 1, \dots, p$ [11]. Além disso, deseja-se que uma AM exiba alguma tolerância a ruído no sentido que $\mathcal{M}(\mathbf{x}) = \mathbf{u}^\xi$ sempre que \mathbf{x} corresponder a uma versão corrompida ou distorcida de \mathbf{u}^ξ . Por fim, intuitivamente, referimos como *capacidade de armazenamento* da memória associativa o maior natural p tal que $\mathcal{M}(\mathbf{u}^\xi) = \mathbf{u}^\xi$ para todo $\xi = 1, \dots, p$. Uma definição formal da capacidade de armazenamento de AMs binárias pode ser encontrada em [36].

Exemplo 1 (Rede de Hopfield Binária). Dado um conjunto de memórias fundamentais $\mathcal{U} = \{\mathbf{u}^\xi : \xi = 1, \dots, p\} \subseteq \{-1, +1\}^n$ e um vetor de entrada $\mathbf{x}(0) \in \{-1, +1\}^n$, a rede de Hopfield binária é descrita pela equação de diferenças

$$x_j(t+1) = \text{sgn}(v_j(t)), \quad v_j(t) = \sum_{k=1}^n m_{jk} x_k(t), \quad \forall j = 0, 1, \dots, n, \quad (3)$$

em que a função sinal é avaliada da seguinte forma

$$\text{sgn}(v_j(t)) = \begin{cases} +1, & v_j(t) > 0, \\ x_j(t), & v_j(t) = 0, \\ -1, & v_j(t) < 0, \end{cases} \quad (4)$$

e os valores w_{jk} são determinados através da seguinte equação para quaisquer índices j, k :

$$m_{jk} = \frac{1}{n} \begin{cases} \sum_{\xi=1}^p u_j^\xi u_k^\xi, & j \neq k, \\ 0, & j = k. \end{cases} \quad (5)$$

Do ponto de vista biológico, $x_1(t), \dots, x_n(t)$ correspondem aos valores de entrada enquanto $x_j(t+1)$ representa a saída do j -ésimo neurônio da rede neural artificial no instante t . O termo $v_j(t)$ em (3) é chamado *potencial de ativação* do neurônio. Observe que $x_j(t+1) = +1$ se $v_j(t) > 0$. Em outras palavras, o j -ésimo neurônio é ativado sempre que o potencial de ativação for positivo. O coeficiente w_{jk} , chamado *peso sináptico*, representa o peso com que a entrada $x_k(t)$ contribui para a saída $x_j(t+1)$. A equação (5) é referida como *armazenamento por correlação* ou *aprendizado de Hebb*. Por fim, pode-se mostrar que a rede de Hopfield

sempre converge para um estado estacionário se apenas uma única componente do vetor $\mathbf{x}(t)$ é atualizado a cada instante t . Precisamente,

$$E(\mathbf{x}) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n m_{ij} x_i x_j, \quad (6)$$

é uma função de Lyapunov da rede neural recorrente descrita por (3).

2.2 Redes Neurais Recorrentes Baseadas em Números Complexos

A primeira generalização significativa da rede de Hopfield usando números complexos no círculo unitário foi proposta em 1996 por Jankowski et al. [24]. Resumidamente, eles consideraram uma *rede neural dinâmica complexa* (CV-DNN, *complex-valued dynamic neural network*) com uma única camada de neurônios que calculam a função sinal-complexa no potencial de ativação e cujos pesos sinápticos são determinados usando uma versão complexa do armazenamento por correlação.

Exemplo 2 (CV-DNN de Jankowski et al.). Seja K um inteiro positivo. Defina $\theta_K = 2\pi/K$ e $\mathbb{S}_K = \{1, e^{i\theta_K}, e^{2i\theta_K}, \dots, e^{i(K-1)\theta_K}\}$, em que i (em negrito) denota a unidade imaginária. Note que \mathbb{S}_K é um subconjunto finito do círculo unitário complexo com K elementos. Em particular, para $K = 2$, tem-se $\mathbb{S}_2 = \{-1, +1\}$. Agora, dado um conjunto de memórias fundamentais $\mathcal{U} = \{\mathbf{u}^\xi : \xi = 1, \dots, p\} \subseteq \mathbb{S}_K^n$ e um vetor coluna complexo $\mathbf{x}(0) = [x_1(0), \dots, x_n(0)]^T \in \mathbb{S}_K^n$, tal como a rede de Hopfield, a CV-DNN de Jankowski et al. define recursivamente

$$x_j(t+1) = \text{csgn}_K(v_j(t)), \quad v_j(t) = \sum_{k=1}^n m_{jk} x_k(t), \quad \forall j = 0, 1, \dots, n. \quad (7)$$

A função sinal-complexa, introduzida por Aizenberg et al. em 1971 [1], é definida através da seguinte equação para qualquer $z \neq 0$:

$$\text{csgn}_K(z) = \begin{cases} 1, & 0 \leq \text{Arg}(z) < \theta_K, \\ e^{i\theta_K}, & \theta_K \leq \text{Arg}(z) < 2\theta_K, \\ \vdots & \vdots \\ e^{i(K-1)\theta_K}, & (K-1)\theta_K \leq \text{Arg}(z) < 2\pi. \end{cases} \quad (8)$$

Define-se $z_j(t+1) = z_j(t)$ quando $v_j(t) = 0$ em (7). Além disso, os pesos sinápticos são determinados através da seguinte equação que corresponde a versão complexa do armazenamento por correlação:

$$m_{jk} = \frac{1}{n} \begin{cases} \sum_{\xi=1}^p u_j^\xi \bar{u}_k^\xi, & j \neq k, \\ 0, & j = k. \end{cases} \quad (9)$$

De um modo similar ao modelo de Hopfield,

$$E(\mathbf{x}) = -\frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n m_{jk} \bar{x}_j x_k, \quad (10)$$

é uma função de Lyapunov da rede neural recorrente descrita por (7) [24]. Em vista desse fato, a CV-DNN of Jankowski et al. pode ser usada para implementar uma memória associativa projetada para armazenar vetores com valores em \mathbb{S}_K^n . Contudo, o modelo de Jankowski et al. apresenta uma baixa capacidade de armazenamento, ou seja, poucos vetores podem ser armazenados e posteriormente recuperados com sucesso por esse modelo.

Em vista da baixa capacidade de armazenamento da CV-DNN de Jankowski et al., muitos pesquisadores apresentaram versões melhoradas ou generalizações do modelo descrito por (7)-(9). Por exemplo, tanto Müezzinoğlu et al. como Lee desenvolveram estratégias que aumentam a capacidade de armazenamento das CV-DANNs baseadas na função sinal-complexa [37, 32]. Kuroe et al. bem como Tanaka e Aihara desenvolveram CV-DNNs na qual a função sinal-complexa é substituída por outras funções de ativação [30, 29, 52]. Suzuki et al. empregaram limiares fortes com objetivo de reduzir o número de memórias espúrias da CV-DNN de Jankowski et al. [51]. Além disso, o proponente investigou recentemente uma classe de redes neurais recorrentes com valores complexos chamadas redes neurais recorrentes por correlação complexas (CV-RCNNs, *complex-valued recurrent correlation neural networks*). Além de generalizar o modelo de Jankowski et al., algumas CV-RCNNs apresentam uma excelente capacidade de armazenamento e tolerância a ruído [56, 54].

2.3 Redes Neurais Recorrentes por Correlação Complexas

Em termos gerais, as redes neurais recorrentes por correlação complexas (CV-RCNNs, *complex-valued recurrent correlation neural networks*) generalizam a CV-DNN de Jankowski et al. adicionando uma camada de nós que computam uma medida de correlação entre o padrão de entrada e os itens armazenados. A função de ativação dos neurônios nessa camada caracterizam a CV-DNN. Sobretudo, a capacidade de armazenamento de certas CV-RCNNs aumenta exponencialmente com o número de padrões para algumas funções de ativação [54]. Além da grande capacidade de armazenamento, as CV-RCNNs apresentam excelentes tolerâncias a ruído.

Formalmente, seja $\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\}$ o conjunto dos números complexos no círculo unitário. Suponha que é dado um conjunto de memórias fundamentais $\{\mathbf{u}^1, \dots, \mathbf{u}^p\} \subseteq \mathbb{S}^n$ e uma função contínua monótona não-decrescente $f : [-n, n] \rightarrow \mathbb{R}$. Dado um padrão de entrada $\mathbf{x}(0) = [x_1(0), \dots, x_n(0)]^T \in \mathbb{S}^n$, uma CV-RCNN define recursivamente a seguinte sequência para $t = 0, 1, \dots$:

$$x_j(t+1) = \sigma \left(\sum_{\xi=1}^p w_{\xi}(t) u_j^{\xi} \right), \quad \forall j = 0, 1, \dots, n, \quad (11)$$



Figura 1: Imagens originais em tons de cinza.

em que os pesos $w_1(t), \dots, w_p(t)$ são dados por

$$w_\xi(t) = f(\Re\{\langle \mathbf{u}^\xi, \mathbf{x}(t) \rangle\}), \quad \forall \xi = 1, \dots, p. \quad (12)$$

Aqui, $\Re\{z\}$ denota a parte real do número complexo z e a função $\sigma : \mathbb{C} \rightarrow \mathbb{S}$ generaliza a função sinal como segue com $v_j(t) = \sum_{\xi=1}^p w_\xi(t) u_j^\xi$ representando o potencial de ativação do j -ésimo neurônio na iteração t :

$$\sigma(v_j(t)) = \begin{cases} v_j(t)/|v_j(t)|, & v_j(t) \neq 0, \\ z_j(t), & v_j(t) = 0. \end{cases} \quad (13)$$

Nesse projeto de pesquisa investigaremos generalizações das CV-RCNNs usando números hipercomplexos. Estudaremos também aspectos teóricos e práticos dos novos modelos. Do ponto de vista teórico, por exemplo, analisaremos a convergência da sequência definida em (11). Do ponto de vista prático, realizaremos experimentos computacionais envolvendo aplicações das CV-RCNNs. O seguinte exemplo apresenta um experimento computacional preliminar com aplicação em reconstrução de imagens em tons de cinza.

Exemplo 3 (Experimento Computacional com Imagens em Tons de Cinza). Considere as oito imagens em tons de cinza mostradas na Figura 1. Estas imagens possuem dimensão 128×128 e $K = 256$ tons de cinza. Para cada uma dessas imagens, geramos um vetor coluna fazendo uma leitura por linhas da imagem e aplicamos a função $r_K : [0, K) \rightarrow \mathbb{S}$ dada por

$$r_K(x) = e^{2\pi ix/K}, \quad (14)$$

em cada componente para obter vetores $\mathbf{u}^\xi \in \mathbb{S}^{16384}$, $\xi = 1, \dots, 8$, com valores no círculo complexo unitário.

Primeiramente, sintetizamos duas CV-RCNNs para armazenar os oito padrões $\mathbf{u}^1, \dots, \mathbf{u}^8$ usando em (12) as funções exponencial $f_e(x) = e^{\alpha x/n}$ e potencial $f_p(x) = 1/(1 - x/n)^L$ com

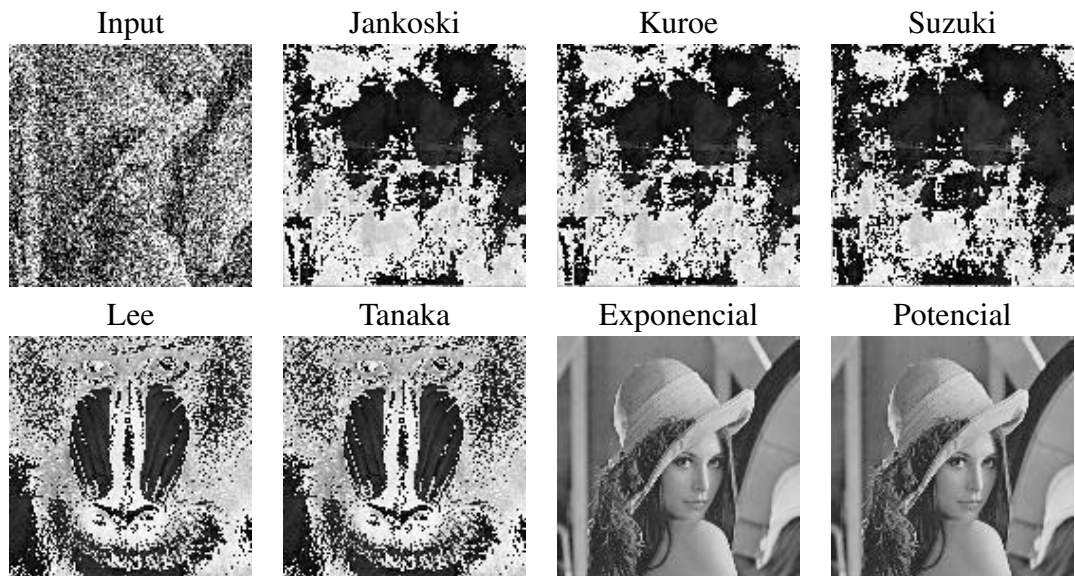


Figura 2: Imagem corrompida com ruído Gaussiano com média 0 e desvio padrão 100 seguido de imagens recuperadas por diferentes memórias associativas complexas.

os parâmetros $\alpha = 10$ e $L = 5$, respectivamente, no lugar de f . Apresentando os padrões originais como entrada obtemos os próprias imagens originais com uma única iteração de (11). Em outras palavras, ambas CV-RCNNs implementaram uma memória associativa.

Posteriormente, avaliamos a tolerância das duas CV-RCNNs com respeito aos ruídos Gaussiano e sal-e-pimenta. As Figuras 2 e 3 apresentam imagens corrompidas seguidas de uma interpretação visual das imagens recordadas pelas duas CV-RCNNs. Para efeito de comparação, essas duas figuras também apresentam as imagens recordadas pelos seguintes modelos: CV-DNN de Jankowski et al. [24], CV-DNN de Lee [32], CV-DNN de Kuroe and Taniguchi com a função de ativação com valores complexas referida por Modelo B [29], CV-DNN de Tanaka and Aihara [52], e a CV-DNN com bias forte de Suzuki et al. [51]. Por simplicidade, usamos o sobrenome do primeiro autor para referir a uma certa CV-DNN. Além disso, as palavras “Exponencial” and “Potencial” referem respectivamente as CV-RCNNs baseadas na função exponencial f_e e função potencial f_p .

Observe que as CV-DNN não foram capazes de restaurar a imagem “lena” corrompida com ruído Gaussiano. De um modo similar, as CV-DNNs de Jankowski et al., Kuroe e Suzuki et al. não foram capazes de recuperar a imagem “airplane” corrompida com ruído sal-e-pimenta. Em contraste, as CV-RCNNs produziram imagens visualmente similares as imagens originais mesmo na presença de ruídos densos. Esses resultados revelam parte do potencial dos novos modelos de AMs que serão investigados nesse projeto de pesquisa.

2.4 Redes de Hopfield Quaterniônicas

Generalizações da rede de Hopfield usando quatérnios foram extensivamente investigadas nos últimos anos [20, 21, 19, 22, 40]. Uma revisão compreensiva de diversas generalizações das redes de Hopfield quaterniônicas (QHNN, *quaternionic Hopfield neural network*) pode ser en-

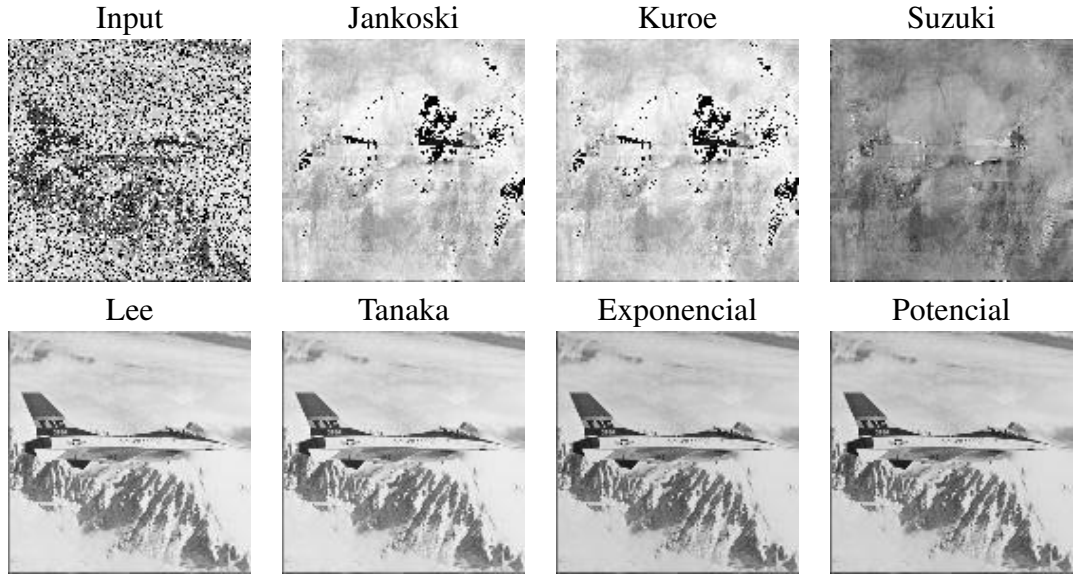


Figura 3: Imagem corrompida com ruído sal-e-pimenta com probabilidade 0.4 seguido de imagens recuperadas por diferentes memórias associativas complexas.

contrada em [23]. Resumidamente, a principal diferença entre as muitas QHNN encontra-se na função de ativação. Por exemplo, a QHNN binária é definida aplicando a função real sgn em cada componente do quatérnio que representa o potencial de ativação de um neurônio [20, 21, 19]. Outra QHNN introduzida por Isokawa e colaboradores, chamada QHNN multivalores (MV-QHNN, *multi-valued QHNN*), generaliza a CV-DNN de Jankowski et al. [23, 24, 32]. Em termos gerais, a função de ativação da MV-QHNN retorna um quatérnio representativo numa seção da hipersfera de todos os quatérnios unitários. Na prática, a implementação de tal função de ativação é trabalhosa pois o potencial de ativação de um neurônio deve ser escrito usando a representação ângulo-fase de um quatérnio [5, 23]. Além disso, a convergência para um estado estacionário é garantida somente se as componentes da representação ângulo-fase não são atualizadas simultaneamente.

Recentemente, o proponente investigou uma QHNN cuja saída dos neurônios é determinada normalizando o potencial de ativação [55]. A nova QHNN, referida como QHNN com valores contínuos (CV-QHNN, *continuous-valued QHNN*), corresponde ao limite da MV-QHNN quando o número de seções da hipersfera quaterniônica tende ao infinito. Especificamente, considere um conjunto de memórias fundamentais $\mathcal{U} = \{\mathbf{u}^1, \dots, \mathbf{u}^p\}$, em que $\mathbf{u}^\xi = [u_1^\xi, \dots, u_n^\xi]^T$ é um vetor coluna quaterniônico com componentes $u_i^\xi = u_{i_0}^\xi + u_{i_1}^\xi \mathbf{i} + u_{i_2}^\xi \mathbf{j} + u_{i_3}^\xi \mathbf{k}$ e $|u_i^\xi| = 1$. Dado um vetor de entrada $\mathbf{x}(0) = [x_1, \dots, x_n]$, a CV-QHNNs define recursivamente a sequência de vetores $\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2), \dots$ através da seguinte equação para todo $j = 1, \dots, n$:

$$x_j(t+1) = \begin{cases} x_j(t), & v_j(t) = 0, \\ v_j(t)/|v_j(t)|, & \text{caso contrário,} \end{cases} \quad (15)$$

em que

$$v_j(t) = \sum_{k=1}^n w_{jk} x_k(t), \quad (16)$$

é o potencial de ativação do j -ésimo neurônio na iteração t . Similar aos modelos binários e complexos, os pesos sinápticos podem ser determinados usando o armazenamento por correlação ou projeção [23].

A versão quaterniônica do armazenamento por correlação é muito semelhante à equação complexa (9). No armazenamento por projeção, que produz memórias associativas superiores aos modelos obtidos usando armazenamento por correlação, os pesos sinápticos são determinados através da equação

$$w_{jk}^p = \frac{1}{n} \sum_{\mu=1}^p \sum_{\nu=1}^p u_j^\xi (W)_{\mu\nu}^{-1} \bar{u}_k^\xi, \quad (17)$$

em que $(W)_{\mu\nu}^{-1}$ denota a μ, ν entrada da inversa da matriz $W = (w_{jk})$ fornecida pelo armazenamento por correlação.

Por fim, tal como o modelo original de Hopfield e muitas de suas generalizações, a CV-QHNN sempre converge para um estado estacionário sob condições fracas sobre os pesos sinápticos. Precisamente, vale o seguinte resultado cuja demonstração pode ser encontrada em [55]:

Teorema 1. *A sequência produzida por (15), com uma única componente atualizada a cada iteração t , é convergente independente do estado inicial $\mathbf{x}(0)$ se a matriz dos pesos sinápticos satisfaz $w_{jk} = \bar{w}_{kj}$ e $w_{jj} > 0$ para qualquer $j, k \in \{1, \dots, n\}$.*

O Teorema 1 revela que a CV-QHNN pode ser usada para implementar uma AM para vetores com valores no conjunto dos quatérnios. Contudo, pouco se sabe sobre a capacidade de armazenamento e tolerância à ruído da CV-QHNN. Logo, nesse projeto de pesquisa propomos investigar esses modelos, tanto do ponto de vista teórico como computacional. Propomos também confrontar a CV-QHNN com outros modelos como, por exemplo, a MV-QHNN [23].

3 Objetivos, Plano de Trabalho e Cronograma

Nesse projeto de pesquisa investigaremos generalizações da rede de Hopfield usando números hipercomplexos. Em particular, estudaremos melhor propriedades teóricas e computacionais tanto das redes recorrentes por correlação com valores complexos (CV-RCNN) como da rede de Hopfield quaterniônica com valores contínuos (CV-QHNN). Também proporemos outras generalizações das redes neurais dinâmicas e CV-RCNNs, por exemplo, usando quatérnios. Precisamente, serão executadas as seguintes atividades:

1. Atividade 1 – Revisão Bibliográfica:

Existe na literatura muitos trabalhos sobre redes neurais artificiais usando números complexos e hipercomplexos. No primeiro semestre faremos uma extensa revisão bibliográfica atualizada de generalizações de rede neurais artificiais usando números hipercom-

plexos. Para efeito de comparação, também faremos uma revisão bibliográfica de modelos de memórias associativas para imagens em tons de cinza e multi-valores, tais como imagens coloridas. Esses modelos serão implementados em softwares matemáticos como GNU Octave e MATLAB.

2. Atividade 2 – Rede de Hopfield Quaterniônica com Valores Contínuos:

No segundo semestre investigaremos propriedades teóricas e computacionais da rede de Hopfield quaterniônica com valores contínuos (CV-QHNN) apresentadas brevemente na Subseção 2.4. Em particular, realizaremos experimentos computacionais para avaliar melhor a capacidade de armazenamento e a tolerância à ruído desses modelos. Também investigaremos a aplicação das CV-QHNN para a recuperação de imagens coloridas.

3. Atividade 3 – Memórias Associativas Recorrentes por Correlação Quaterniônicas:

No terceiro e quarto semestres estenderemos as redes neurais recorrentes por correlação complexas (CV-RCNNs) apresentadas brevemente na Subseção 2.3 usando quatérnions. Em particular, apresentaremos resultados sobre a convergência desses modelos e realizaremos experimentos computacionais para avaliar melhor a capacidade de armazenamento e a tolerância à ruído. Os modelos desenvolvidos nessa etapa serão aplicados, por exemplo, na reconstrução de imagens coloridas.

4. Atividade 4 – Desenvolvimento de outras generalizações:

No quinto e sexto semestre estudaremos outras generalizações das CV-RCNNs e CV-RCNNs. Por exemplo, investigaremos modelos obtidos usando números hiperbólicos, tessarinos, octonios bem como modelos baseados em álgebras de Clifford. Tal como nas atividades anteriores, resultados teóricos e computacionais com respeito a convergência, capacidade de armazenamento e tolerância a ruído serão investigados.

As quatro atividades listadas acima seguirão o seguinte cronograma:

| Mar/2015 a Ago/2015 | Set/2015 a Fev/2016 | Mar/2016 a Fev/2017 | Mar/2017 a Fev/2018 |
|---------------------|---------------------|---------------------|---------------------|
| Atividade 1 | Atividade 2 | Atividade 3 | Atividade 4 |

Durante todas as atividades, pretendemos divulgar os resultados obtidos em eventos científicos da área bem como envolver alunos de graduação e pós-graduação nessa área de pesquisa. Também serão redigidos artigos científicos que serão submetidos para revistas conceituadas, preferencialmente de circulação internacional.

4 Contribuições e Resultados Esperados

Nesse projeto de pesquisa investigaremos modelos memórias associativas implementadas usando redes neurais recorrentes baseadas em números hipercomplexos. Conforme mencionado no primeiro parágrafo da introdução, modelos de memória associativa foram usados com sucesso

em problemas de classificação, previsão e reconhecimento de padrões. Portanto, esse projeto de pesquisa deve contribuir nessas aplicações. Contudo, nos concentraremos principalmente nas aplicações relacionadas à reconstrução de imagens em tons de cinza e coloridas.

É importante ressaltar que os todos resultados desenvolvidos terão forte fundamento matemático. Sobretudo, os resultados desenvolvidos nesse projeto de pesquisa serão submetidos na forma de artigo para um periódico de circulação internacional como *IEEE Transactions on Neural Networks*, *Neural Networks*, *Neurocomputing*, dentre outros similares.

Referências

- [1] AIZENBERG, N. N., IVASKIV, Y. L., AND POSPELOV, D. A. A certain generalization of threshold functions. *Doklady Akademii Nauk SSSR* 196 (1971), 1287–1290.
- [2] ANDERSON, J. A Simple Neural Network Generating Interactive Memory. *Mathematical Biosciences* 14 (1972), 197–220.
- [3] ARENA, P., FORTUNA, L., MUSCATO, G., AND XIBILIA, M. Multilayer Perceptrons to Approximate Quaternion Valued Functions. *Neural Networks* 10, 2 (1997), 335–342.
- [4] AUSTIN, J. Associative Memory. In *Handbook of Neural Computation*, E. Fiesler and R. Beale, Eds. Oxford University Press, 1997, pp. F1.4:1–F1.4:7.
- [5] BÜLOW, T. Hypercomplex Spectral Signal Representations for Image Processing and Analysis. Tech. Rep. NR 9903, Kiel University, Aug. 1999.
- [6] CHE UJANG, B., TOOK, C., AND MANDIC, D. Quaternion-Valued Nonlinear Adaptive Filtering. *IEEE Transactions on Neural Networks* 22, 8 (Aug 2011), 1193–1206.
- [7] FORTUNA, L., MUSCATO, G., AND XIBILIA, M. An hypercomplex neural network platform for robot positioning. In *Circuits and Systems, 1996. ISCAS '96., Connecting the World., 1996 IEEE International Symposium on* (May 1996), vol. 3, pp. 609–612 vol.3.
- [8] FREITAS, G. A., AND VALLE, M. E. Uma Breve Comparação de Modelos de Memórias Associativas Fuzzy em Problemas de Guiagem Autônoma. *SEMINA: Ciências Exatas e Tecnológicas* 32, 2 (2011), 151–166.
- [9] HASSOUN, M. H., Ed. *Associative Neural Memories: Theory and Implementation*. Oxford University Press, Oxford, U.K., 1993.
- [10] HASSOUN, M. H. *Fundamentals of Artificial Neural Networks*. MIT Press, Cambridge, MA, 1995.
- [11] HASSOUN, M. H., AND WATTA, P. B. Associative Memory Networks. In *Handbook of Neural Computation*, E. Fiesler and R. Beale, Eds. Oxford University Press, 1997, pp. C1.3:1–C1.3:14.

- [12] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Upper Saddle River, NJ, 1999.
- [13] HAYKIN, S. *Neural Networks and Learning Machines*, 3rd edition ed. Prentice-Hall, Upper Saddle River, NJ, 2009.
- [14] HECHT-NIELSEN, R. *Neurocomputing*. Addison-Wesley, Reading, MA, 1989.
- [15] HIROSE, A. *Complex-Valued Neural Networks*, 2nd edition ed. Studies in Computational Intelligence. Springer, Heidelberg, Germany.
- [16] HIRSH, M. W. Dynamical Systems. In *Mathematical Perspectives on Neural Networks*, P. Smolensky, M. Mozer, and D. Rumelhart, Eds. Lawrence Erlbaum Associates, Publishers, Mahwah, NJ., 1996.
- [17] HOPFIELD, J., AND TANK, D. Neural computation of decisions in optimization problems. *Biological Cybernetics* 52 (1985), 141–152.
- [18] HOPFIELD, J. J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences* 79 (Apr. 1982), 2554–2558.
- [19] ISOKAWA, T., HISHIMURA, H., KAMIURA, N., AND MATSUI, N. Associative Memory in Quaternionic Hopfield Neural Network. *International Journal of Neural Systems* 18, 02 (2008), 135–145.
- [20] ISOKAWA, T., NISHIMURA, H., KAMIURA, N., AND MATSUI, N. Fundamental Properties of Quaternionic Hopfield Neural Network. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)* (Vancouver, Canada, 2006), pp. 218–223.
- [21] ISOKAWA, T., NISHIMURA, H., KAMIURA, N., AND MATSUI, N. Dynamics of Discrete-Time Quaternionic Hopfield Neural Networks. In *Artificial Neural Networks – ICANN 2007*, J. M. Sá, L. A. Alexandre, W. Duch, and D. Mandic, Eds., vol. 4668 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007, pp. 848–857.
- [22] ISOKAWA, T., NISHIMURA, H., AND MATSUI, N. On the fundamental properties of fully quaternionic hopfield network. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)* (Brisbane, Australia, June 2012), pp. 1–4.
- [23] ISOKAWA, T., NISHIMURA, H., AND MATSUI, N. Quaternionic Neural Networks for Associative Memories. In *Complex-Valued Neural Networks*, A. Hirose, Ed. Wiley-IEEE Press, 2013, pp. 103–131.
- [24] JANKOWSKI, S., LOZOWSKI, A., AND ZURADA, J. Complex-Valued Multi-State Neural Associative Memory. *IEEE Transactions on Neural Networks* 7 (1996), 1491–1496.

- [25] KOHONEN, T. Correlation Matrix Memory. *IEEE Transactions on Computers C-21* (1972), 353–359.
- [26] KOHONEN, T. *Self-organization and associative memory*, 2nd edition ed. Springer-Verlag New York, Inc., New York, NY, USA, 1987.
- [27] KONG, S.-G., AND KOSKO, B. Adaptive Fuzzy Systems for Backing up a Truck-and-Trailer. *IEEE Transactions on Neural Networks* 3, 2 (1992), 211–223.
- [28] KOSKO, B. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, NJ, 1992.
- [29] KUROE, Y., AND TANIGUCHI, Y. Models of self-correlation type complex-valued associative memories and their dynamics. In *Artificial Neural Networks: Biological Inspirations – ICANN 2005*, W. Duch, J. Kacprzyk, E. Oja, and S. Zadrozny, Eds., vol. 3696 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2005, pp. 185–192.
- [30] KUROE, Y., YOSHIDA, M., AND MORI, T. On activation functions for complex-valued neural networks – existence of energy functions –. In *Artificial Neural Networks and Neural Information Processing – ICANN/ICONIP 2003*, O. Kaynak, E. Alpaydin, E. Oja, and L. Xu, Eds., vol. 2714 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2003, pp. 985–992.
- [31] KUSAMICHI, H., ISOKAWA, T., MATSU, N., OGAWA, Y., AND MAEDA, K. A New Scheme for Color Night Vision by Quaternion Neural Network. In *Proceedings of the 2nd International Conference on Autonomous Robots and Agents (ICARA 2004)* (Dec. 2004), pp. 101–106.
- [32] LEE, D.-L. Improvements of complex-valued Hopfield associative memory by using generalized projection rules. *IEEE Transactions on Neural Networks* 17, 5 (September 2006), 1341–1347.
- [33] LIU, Q., AND WANG, J. A One-Layer Recurrent Neural Network for Constrained Nonsmooth Optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 41, 5 (Oct. 2011), 1323–1333.
- [34] MANDIC, D., JAHANCHAH, C., AND TOOK, C. A Quaternion Gradient Operator and Its Applications. *Signal Processing Letters, IEEE* 18, 1 (Jan 2011), 47–50.
- [35] MATSUI, N., ISOKAWA, T., KUSAMICHI, H., PEPPER, F., AND NISHIMURA, H. Quaternion neural network with geometrical operators. *Journal of Intelligent and Fuzzy Systems* 15, 3 (Jan. 2004), 149–164.
- [36] MCELIECE, R. J., POSNER, E. C., RODEMICH, E. R., AND VENKATESH, S. The capacity of the Hopfield associative memory. *IEEE Transactions on Information Theory* 1 (1987), 33–45.

- [37] MÜEZZINOĞLU, M., GÜZELİŞ, C., AND ZURADA, J. A New Design Method for the Complex-Valued Multistate Hopfield Associative Memory. *IEEE Transactions on Neural Networks* 14, 4 (July 2003), 891–899.
- [38] NAKANO, K. Associatron: A Model of Associative Memory. *IEEE Trans. on Systems, Man, Cybernetics SMC-2* (1972), 380–388.
- [39] NITTA, T. *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*. Igi Global, Hershey - PA, USA, Feb. 2009.
- [40] OSANA, Y. Chaotic Quaternionic Associative Memory. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)* (Brisbane, Australia, June 2012), pp. 1–8.
- [41] PAJARES, G. A Hopfield Neural Network for Image Change Detection. *IEEE Transaction on Neural Networks* 17, 5 (Sept. 2006), 1250–1264.
- [42] SERPEN, G. Hopfield Network as Static Optimizer: Learning the Weights and Eliminating the Guesswork. *Neural Processing Letters* 27, 1 (2008), 1–15.
- [43] SMITH, K., PALANISWAMI, M., AND KRISHNAMOORTHY, M. Neural Techniques for Combinatorial Optimization with Applications. *IEEE Transactions on Neural Networks* 9, 6 (Nov. 1998), 1301–1318.
- [44] STEINBRUCH, K. Die Lernmatrix. *Kybernetick 1* (1961), 36–45.
- [45] SUSSNER, P., ESMI, E. L., VILLAVERDE, I., AND GRAÑA, M. The Kosko Subsethood Fuzzy Associative Memory (KS-FAM): Mathematical Background and Applications in Computer Vision. *Journal of Mathematical Imaging and Vision* 42, 2–3 (Feb. 2012), 134–149.
- [46] SUSSNER, P., MIYASAKI, R., AND VALLE, M. E. An Introduction to Parameterized IFAM Models with Applications in Prediction. In *Proceedings of the 2009 IFSA World Congress and 2009 EUSFLAT Conference* (Lisbon, Portugal, July 2009), pp. 247–252.
- [47] SUSSNER, P., AND VALLE, M. E. Grayscale Morphological Associative Memories. *IEEE Transactions on Neural Networks* 17, 3 (May 2006), 559–570.
- [48] SUSSNER, P., AND VALLE, M. E. Implicative Fuzzy Associative Memories. *IEEE Transactions on Fuzzy Systems* 14, 6 (Dec. 2006), 793–807.
- [49] SUSSNER, P., AND VALLE, M. E. Recall of Patterns Using Morphological and Certain Fuzzy Morphological Associative Memories. In *Proceedings of the IEEE World Conference on Computational Intelligence 2006* (Vancouver, Canada, 2006), pp. 209–216.

- [50] SUSSNER, P., AND VALLE, M. E. Morphological and Certain Fuzzy Morphological Associative Memories for Classification and Prediction. In *Computational Intelligence Based on Lattice Theory*, V. Kaburlasos and G. Ritter, Eds. Springer-Verlag, Heidelberg, Germany, 2007, ch. 8, pp. 149–172.
- [51] SUZUKI, Y., KITAHARA, M., AND KOBAYASHI, M. Dynamic complex-valued associative memory with strong bias terms. In *Neural Information Processing*, B.-L. Lu, L. Zhang, and J. Kwok, Eds., vol. 7062 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 509–518.
- [52] TANAKA, G., AND AIHARA, K. Complex-Valued Multistate Associative Memory With Nonlinear Multilevel Functions for Gray-Level Image Reconstruction. *IEEE Transactions on Neural Networks* 20, 9 (Sept. 2009), 1463–1473.
- [53] TERABAYSASHI, K., AND HIROSE, A. Ultra-short acoustic imaging using complex-valued spatio-temporal neural-network for null-steering: experimental results. In *Proceedings of the IEEE World Conference on Computational Intelligence 2014 (WCCI 2014)* (Beijing, China, July 2014).
- [54] VALLE, M. Complex-valued recurrent correlation neural networks. *IEEE Transactions on Neural Networks and Learning Systems* (2014). Accepted for publication.
- [55] VALLE, M. E. A Novel Continuous-Valued Quaternionic Hopfield Neural Network. In *Proceedings of the Brazilian Conference on Intelligent Systems 2014 (BRACIS 2014)* (São Carlos, Brazil, October 2014).
- [56] VALLE, M. E. An Introduction to Complex-Valued Recurrent Correlation Neural Networks. In *Proceedings of the IEEE World Conference on Computational Intelligence 2014 (WCCI 2014)* (Beijing, China, July 2014).
- [57] ZHANG, B.-L., ZHANG, H., AND GE, S. S. Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory. *IEEE Transactions on Neural Networks* 15, 1 (Jan. 2004), 166–177.
- [58] ZHANG, D., AND ZUO, W. Computational Intelligence-Based Biometric Technologies. *IEEE Computational Intelligence Magazine* 2, 2 (May 2007), 26–36.
- [59] ZHANG, H., HUANG, W., HUANG, Z., AND ZHANG, B. A Kernel Autoassociator Approach to Pattern Classification. *IEEE Transactions on Systems, Man and Cybernetics, Part B* 35, 3 (June 2005), 593–606.