



Universidade Estadual de Campinas

Instituto de Matemática, Estatística e Computação Científica

Departamento de Matemática Aplicada

Um Estudo sobre o Treinamento de Redes Neurais com Valores Hipercomplexos

A Study on Training Hypercomplex-Valued Neural Networks

Marcos Eduardo Ribeiro do Valle Mesquita

E-mail: valle@ime.unicamp.br

Resumo

Redes neurais artificiais (RNAs) são modelos adaptativos inspirados na capacidade do cérebro humano de aprender e resolver tarefas complexas. Estudos recentes mostraram que a substituição de parâmetros com valores reais por números hipercomplexos pode resultar melhoras significativas no desempenho da RNA devido, em partes, à uma compactação de operações e um melhor tratamento de dados multidimensionais. Exemplos de álgebras de números hipercomplexos usadas na construção de redes neurais incluem números complexos e hiperbólicos, quatérnios, octônios, álgebras de Cayley-Dickson e álgebras de Clifford. Este projeto tem como objetivo investigar redes neurais com valores hipercomplexos em diversas álgebras a fim de explorar características destes modelos em contraste com modelos reais consolidados. Especificamente, temos como objetivo estudar e desenvolver regras de treinamento para redes neurais artificiais com valores hipercomplexos com ênfase nas redes de Hopfield e máquinas de aprendizado extremo.

Palavras-chave: Inteligência computacional, rede neural artificial, rede de Hopfield, máquina de aprendizado extremo.

Keywords: Computational intelligence, artificial neural network, hypercomplex numbers, Hopfield network, extreme learning machine.

Campinas, 8 de setembro de 2021.

1 Introdução Geral

Redes neurais artificiais (RNAs) são modelos matemáticos inspirados no sistema nervoso humano. De modo geral, RNAs tentam emular o processo de sinapse, responsável por conduzir informações e gerar respostas adequadas a um estímulo recebido. RNAs são frequentemente empregadas em problemas de classificação, regressão, previsão, controle, entre outras [15, 16, 18].

Redes neurais com valores hipercomplexos são RNAs cujos parâmetros e operações são definidas usando uma álgebra hipercomplexa, tal qual os números complexos, os números hiperbólicos [2, 51], os quatérnios [5], ou os octônios [6]. Destacamos que redes neurais com valores complexos foram introduzidas na década de 70 por Aizenberg et al. e constituem ainda um tópico de pesquisa [1, 3, 12, 56]. Outro exemplo, a Q-MLP desenvolvida por Arena et al. é uma extensão do perceptron de múltiplas camadas real (MLP, do inglês *multilayer perceptron*) para parâmetros quaterniônicos [4]. A Q-MLP possui muitas características atraentes incluindo a simplificação de operações, a diminuição no número total de parâmetros e a condensação de múltiplas informações referentes a um mesmo objeto numa única entidade. Diversos trabalhos demonstram o desempenho da Q-MLP em problemas de classificação e regressão [8, 32, 42] enquanto outros trabalhos desenvolvem o aspecto teórico desta rede neural, como a demonstração do teorema da aproximação universal [4], a proposição de algoritmos alternativos de treinamento [57] e a introdução de derivadas generalizadas de Hamilton-Riemann para ajuste de parâmetros [46]. Aspectos geométricos também foram consideradas no desenvolvimento de RNAs com valores hipercomplexos. Por exemplo, Isokawa et al. desenvolveram uma rede neural de múltiplas camadas que considera as rotações realizada na multiplicação por quatérnios [27]. De um modo geral, diferente dos modelos reais, uma rede neural com valores hipercomplexos processa dados multi-dimensionais como uma única entidade [13, 19, 50]. Aplicações de redes neurais baseadas em números hipercomplexos inclui controle [13, 4], previsão [64, 46] e processamento de imagens e sinais [12, 42, 60, 63].

Nos anos 2000, devido ao aumento significativo no volume dos dados a serem tratados, abordagens que tem em vista a complexidade computacional dos algoritmos surgiram com maior expressividade. Nesse contexto, Huang et al. propuseram as máquinas de aprendizado extremo (ELMs, do inglês *extreme learning machines*) [24]. Resumidamente, uma ELM é uma rede neural no qual os parâmetros da camada intermediária são determinados aleatoriamente enquanto que os parâmetros da camada de saída são ajustados através de um problema de quadrados mínimos regularizado. A substituição de um método iterativo de ajuste por um método direto como o método dos quadrados mínimos resulta uma drástica diminuição no número de operações na fase de treinamento. Sobretudo, propriedades de aproximação também são válidas para as ELMs [10]. Em 2017, Minemoto et al. introduziram um modelo quaterniônico da ELM e exploraram sua aplicação em um problema de auto-codificação (do inglês *autoencoding*) de imagens coloridas [48]. Como mencionado anteriormente, RNAs com valores hipercomplexos são adequadas ao tratamento de dados multidimensionais, incluindo imagens coloridas, multi-bandas

ou hiper-espectrais. Com efeito, conforme destacado por Minemoto et al., a ELM quaterniônica apresentou um desempenho superior à rede real correspondente, corroborando a adequação destes modelos ao tratamento de dados multidimensionais como imagens e vídeos, entre outros [48]. Recentemente, estendemos a ELM para as álgebras de Cayley-Dickson generalizadas e, num estudo preliminar, obtivemos resultados melhores que ambas as redes com valores reais e quaterniônicos num experimento de auto-codificação. Em vista desses resultados preliminares, propomos nesse projeto de pesquisa investigar detalhadamente modelos e aplicações das ELM com valores hipercomplexos.

Diferentes das redes progressivas, a rede de Hopfield é um exemplo importante de RNA recursiva que pode ser usada para implementar uma memória associativa e resolver problemas de otimização [20, 21]. Especificamente, a rede de Hopfield é implementada por uma rede neural recursiva de camada única constituída por neurônios de McCulloch e Pitts [18]. 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 [54] e otimização [21, 45, 59, 62]. Apesar de suas características atraentes, a rede de Hopfield é projetada para armazenar e recordar padrões binários. Todavia, muitas aplicações, incluindo a recuperação de imagens coloridas, requerem o armazenamento de vetores com valores hipercomplexos [19]. Com efeito, há um interesse crescente no desenvolvimento de redes de Hopfield com valores hipercomplexos, conforme descrito no próximo parágrafo.

A pesquisa em redes de Hopfield com valores hipercomplexos data do final dos anos 80 [52, 53]. Em 1996, Jankowski et al. [31] propuseram uma versão da rede de Hopfield de múltiplos estados representados por números complexos e aprendizado de Hebb. Esse modelo corroborou com o desenvolvimento de muitos outros modelos com valores hipercomplexos. Por exemplo, regras de aprendizado aprimoradas para redes de Hopfield com múltiplos estados representados por números complexos incluem a regra de projeção proposta por Lee [43] e o treinamento proposto por Müezzinoğlu com base na solução de um conjunto de desigualdades [49]. Pesquisas recentes sobre as redes de Hopfield com valores complexos de múltiplos estados incluem regras de aprendizado iterativo [30] e redes com pesos simétricos [35]. Além dos modelos de valores complexos, Isokawa et al. desenvolveram versões de redes de Hopfield com valores nos quatérnios, usando tanto a regra do aprendizado de Hebb como por projeção [26, 29]. Isokawa e colaboradores também investigaram redes de Hopfield baseadas em quatérnios comutativos, que são isomorfos aos tessarinos [28]. A estabilidade das redes neurais de Hopfield com valores nos quatérnios unitários foi abordada por Valle e de Castro [65]. Funções de ativação de múltiplos estados para redes de Hopfield com valores nos quatérnios e tessarinos foram propostas e investigadas por Kobayashi [34, 37]. Além dos quatérnios e tessarinos, redes de Hopfield baseadas nas álgebras de Clifford [40, 66], incluindo números hiperbólicos [33, 36, 38], foram desenvolvidas na última década. Recentemente, desenvolvemos um resultado geral que assegura a estabilidade de redes de Hopfield com valores hipercomplexos. Contudo, não apresentamos uma regra efetiva de armazenamento de padrões para esses mode-

los. Dessa forma, além do estudo das ELM, nesse projeto estudaremos regras de aprendizado para a rede de Hopfield com valores hipercomplexos. Nesse contexto, destacamos que o aprendizado por projeção pode ser formulado como um problema de quadrados mínimos, tal como no treinamento da ELM. Portanto, apesar do contexto diferente, o treinamento da rede de Hopfield e da ELM com valores hipercomplexos estão intimamente relacionados.

2 Identificação e Caracterização do Problema

2.1 Conceitos Básicos sobre Números Hipercomplexos

Um número hipercomplexo sobre o corpo dos reais é escrito como

$$p = p_0 + p_1 \mathbf{i}_1 + \dots + p_n \mathbf{i}_n, \quad (1)$$

em que n é um inteiro não-negativo, p_0, p_1, \dots, p_n são números reais e os símbolos $\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_n$ são as chamadas *unidades hiper-imaginárias* [61]. O conjunto de todos os números dados por (1) será denotado por \mathbb{H} . Exemplos de números hipercomplexos incluem os números complexos, os números hiperbólicos, os quatérnios, os tessarinos e os octônios.

Um sistema de números hipercomplexos é um conjunto de números hipercomplexos com operações de soma e multiplicação (ou produto). A adição de dois números hipercomplexos $p = p_0 + p_1 \mathbf{i}_1 + \dots + p_n \mathbf{i}_n$ e $q = q_0 + q_1 \mathbf{i}_1 + \dots + q_n \mathbf{i}_n$ é definida como segue:

$$p + q = (p_0 + q_0) + (p_1 + q_1) \mathbf{i}_1 + \dots + (p_n + q_n) \mathbf{i}_n. \quad (2)$$

O produto entre p e q , denotado pela justaposição de p e q , é definido usando a distributividade e uma tabela de multiplicação. A tabela de multiplicação estabelece o produto entre quaisquer duas unidades hiper-imaginárias através da identidade

$$\mathbf{i}_\mu \mathbf{i}_\nu = a_{\mu\nu,0} + a_{\mu\nu,1} \mathbf{i}_1 + \dots + a_{\mu\nu,n} \mathbf{i}_n, \quad \forall \mu, \nu \in \{1, \dots, n\}, \quad (3)$$

Posteriormente, usando a distributividade e a tabela de multiplicação, encontramos

$$pq = \left(p_0 q_0 + \sum_{\mu, \nu=1}^n p_\mu q_\nu a_{\mu\nu,0} \right) + \left(p_0 q_1 + p_1 q_0 + \sum_{\mu, \nu=1}^n p_\mu q_\nu a_{\mu\nu,1} \right) \mathbf{i}_1 + \dots + \left(p_0 q_n + p_n q_0 + \sum_{\mu, \nu=1}^n p_\mu q_\nu a_{\mu\nu,n} \right) \mathbf{i}_n. \quad (4)$$

Além da soma e da multiplicação, definimos a parte real de um número hipercomplexo $p = p_0 + p_1 \mathbf{i}_1 + \dots + p_n \mathbf{i}_n$ como sendo o número real $\text{Re} \{p\} := p_0$.

Note que podemos identificar um número hipercomplexo $p = p_0 + p_1 \mathbf{i}_1 + \dots + p_n \mathbf{i}_n$ com

uma $(n + 1)$ -upla (p_0, p_1, \dots, p_n) de números reais. Sobretudo, o sistema de números hipercomplexos pode ser embutido num espaço vetorial de dimensão $(n + 1)$ em que a soma e adição por escalar são determinados por (2) e (4), respectivamente. Dessa forma, podemos escrever $\dim(\mathbb{H}) = n + 1$.

A seguinte definição generaliza o conceito de conjugação para números hipercomplexos e tem um papel importante na definição e análise da estabilidade das redes de Hopfield hipercomplexas:

Definição 1 (Involução Reversa [11]). *Um operador $\tau : \mathbb{H} \rightarrow \mathbb{H}$ é uma involução reversa se satisfaz as seguintes identidades para todo $p, q \in \mathbb{H}$ e $\alpha \in \mathbb{R}$:*

$$\tau(\tau(p)) = p, \quad (\text{involução}) \quad (5)$$

$$\tau(pq) = \tau(q)\tau(p), \quad (\text{anti-homomorfismo}) \quad (6)$$

$$\tau(\alpha p + q) = \alpha\tau(p) + \tau(q). \quad (\text{linear}) \quad (7)$$

O conjugado natural de um número hipercomplexo $p = p_0 + p_1\mathbf{i}_1 + \dots + p_n\mathbf{i}_n$, denotado por \bar{p} , é dado pela equação

$$\bar{p} = p_0 - p_1\mathbf{i}_1 - \dots - p_n\mathbf{i}_n. \quad (8)$$

A conjugação natural, definida por $\tau(p) = \bar{p}$, é um exemplo de involução reversa em alguns sistemas de números hipercomplexos tais como os números complexos e os quatérnios. A aplicação identidade $\tau(p) = p$, para todo $p \in \mathbb{H}$, é também uma involução reversa se a multiplicação for comutativa. Nesse caso, chamamos a identidade de involução reversa trivial. Exemplos de involuções reversas também incluem anti-involuções nos quatérnios e a conjugação de Clifford nas álgebras de Clifford [67].

Usando uma involução reversa, podemos definir uma forma bilinear simétrica $\mathcal{B} : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}$ através da seguinte equação

$$\mathcal{B}(p, q) = \text{Re} \{ \tau(p)q \}, \quad \forall p, q \in \mathbb{H}. \quad (9)$$

Intuitivamente, \mathcal{B} fornece uma medida da relação entre p e q levando em consideração propriedades da multiplicação e também da involução reversa τ . Por exemplo, a forma bilinear simétrica \mathcal{B} coincide com o produto interno usual nos sistemas de números complexos, quatérnios e octônios com a conjugação natural. A forma bilinear simétrica \mathcal{B} é usada na seguinte definição:

Definição 2 (\mathcal{B} -função [11]). *Considere um sistema de números hipercomplexos \mathbb{H} equipado com uma involução reversa τ e seja $\mathcal{B} : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}$ a forma bilinear simétrica definida por (9). Uma função com valores hipercomplexos $\phi : \mathcal{D} \subset \mathbb{H} \rightarrow \mathcal{S} \subset \mathbb{H}$ é chamada \mathcal{B} -função quando satisfaz a seguinte desigualdade:*

$$\mathcal{B}(\phi(q), q) > \mathcal{B}(s, q), \quad \forall q \in \mathcal{D}, \forall s \in \mathcal{S} \setminus \{\phi(q)\}. \quad (10)$$

Em palavras, uma \mathcal{B} -função $\phi : \mathcal{D} \rightarrow \mathcal{S}$ transforma $q \in \mathcal{D}$ num número hipercomplexo $\phi(q) \in \mathcal{S}$ que, com respeito à forma bilinear simétrica \mathcal{B} , está mais relacionado à q que qualquer outro elemento $s \in \mathcal{S}$ na imagem de ϕ . Exemplos de \mathcal{B} -funções incluem muitas funções de ativação usadas em redes de Hopfield hipercomplexas como a função sinal multi-estado [2] e a função multi-estado gêmeas introduzidas por Kobayashi [34].

Finalmente, para analisar a estabilidade de uma ampla classe de redes de Hopfield com valores hipercomplexos consideramos a definição [11]:

Definição 3 (Sistemas de Números Hipercomplexos com Parte Real Associativa). *Um sistema de números hipercomplexos \mathbb{H} equipado com uma involução reversa τ é um sistema de número hipercomplexo com parte real associativa (Re-AHNS, do inglês real-part associative hypercomplex number system) se a seguinte identidade vale para quaisquer elementos $p, q, r \in \mathbb{H}$:*

$$\text{Re} \{(pq)r - p(qr)\} = 0. \quad (11)$$

Em particular, dizemos que \mathbb{H} é um Re-AHNS definido não-negativo se a forma bilinear simétrica \mathcal{B} dada por (9) satisfaz $\mathcal{B}(p, p) \geq 0$, para todo $p \in \mathbb{H}$.

Números complexos, quatérnios e octônios, todos equipados com a conjugação natural, são exemplos de Re-AHNS definidos não-negativos.

2.2 Rede de Hopfield Discreta no Tempo Com Valores Hipercomplexos

A rede de Hopfield real (bipolar) discreta no tempo é uma rede recorrente que pode ser usada para implementar uma memória associativa [20]. Além de ser usada como memória associativa, a rede de Hopfield foi usada para resolver problemas de controle, classificação, otimização e visão computacional [14, 21, 44, 55]. Versões hipercomplexas da rede de Hopfield foram extensivamente investigadas nos últimos anos no contexto dos números complexos [31, 43], números hiperbólicos [33], tessarinos [28, 37], quatérnios [26, 65], octônios [9, 41], entre outros [40, 66]. Nessa seção revisaremos uma ampla classe de redes de Hopfield hipercomplexas que abrange muitos modelos citados nesse parágrafo.

Seja \mathbb{H} um Re-AHNS e $\phi : \mathcal{D} \rightarrow \mathcal{S}$ uma \mathcal{B} -função em que \mathcal{S} é um conjunto compacto de \mathbb{H} . Em analogia à rede de Hopfield tradicional, seja $w_{ij} \in \mathbb{H}$ o peso da j -ésima conexão sináptica do i -ésimo neurônio de uma rede com N neurônios. Vamos denotar o estado da rede de Hopfield hipercomplexa no instante de tempo t por um vetor $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T \in \mathcal{S}^N$, em que $x_i(t) = x_{i0}(t) + x_{i1}(t)\mathbf{i}_1 + \dots + x_{in}(t)\mathbf{i}_n$ representa o estado do i -ésimo neurônio no tempo t . Dado um vetor de estado inicial $\mathbf{x}(0)$, a rede de Hopfield discreta no tempo evolui de acordo com a equação:

$$x_i(t + \Delta t) = \begin{cases} \phi(h_i(t)), & h_i(t) \in \mathcal{D}, \\ x_i(t), & \text{caso contrário,} \end{cases} \quad (12)$$

em que

$$h_i(t) = \sum_{j=1}^N w_{ij} x_j(t), \quad (13)$$

denota o potencial de ativação do i -ésimo neurônio no tempo t . Os estados da rede são geralmente atualizados de forma síncrona ou assíncrona. No modo síncrono, também chamado paralelo, todos os neurônios são atualizados simultaneamente. No modo de atualização assíncrona, também chamado serial, um único neurônio é atualizado a cada instante de tempo. Em analogia à rede de Hopfield tradicional, demonstramos o seguinte teorema com respeito à dinâmica da rede de Hopfield hipercomplexa [11]:

Teorema 4 (Dinâmica da Rede de Hopfield Hipercomplexa). *A sequência de estados de uma rede de Hopfield discreta no tempo com valores hipercomplexos definida por (12) converge no modo assíncrono se $w_{ij} = \tau(w_{ji})$ e um dos seguintes casos é verdadeiro:*

- (a) $w_{ii} = 0$ para todo $i \in \{1, \dots, N\}$.
- (b) $w_{ii} \geq 0$ é um número real não-negativo para qualquer $i \in \{1, \dots, N\}$ e \mathbb{H} é um Re-AHNS definido semi-negativo.

O resultado descrito pelo Teorema 4 é particularmente importante na implementação de memórias associativas bem como para a resolução de problemas de otimização [17, 20, 21]. Além da análise da estabilidade, aplicações efetivas da rede de Hopfield requerem regras apropriadas para determinar os pesos sinápticos, chamada regra de treinamento.

No contexto de memórias associativas, desejamos determinar w_{ij} de modo que certos vetores $\mathbf{u}^1, \dots, \mathbf{u}^P$ sejam estados estacionários atratores da rede de Hopfield. O aprendizado de Hebb e o armazenamento por projeção são exemplos de regras de treinamento para certas versões da rede de Hopfield com valores hipercomplexos [31, 43, 29, 47]. O aprendizado de Hebb, apesar de sua simplicidade, resulta em uma baixa capacidade de armazenamento devido à interferência cruzada [20, 29]. O armazenamento por projeção, apesar de aumentar significativamente a capacidade de armazenamento, não pode ser estendido diretamente para algumas versões hipercomplexas da rede de Hopfield como a versão baseada em números hiperbólicos [38]. Em vista dessa observação, nesse projeto de pesquisa iremos investigar regras de armazenamento para as redes de Hopfield com valores hipercomplexos.

2.3 Máquinas de Aprendizado Extremo com Valores Hipercomplexos

Máquinas de aprendizado extremo (ELM, do inglês *extreme learning machines*) são redes progressivas conhecidas pelo baixo custo computacional na fase de treinamento [22, 23, 24, 25]. O baixo custo computacional ocorre porque os parâmetros das camadas intermediárias são gerados aleatoriamente enquanto que os parâmetros da camada de saída são ajustados resolvendo um problema de quadrados mínimos. Ilustramos a seguir como uma ELM com uma única camada oculta pode ser definida usando um sistema de números hipercomplexos.

Primeiro, seja $\mathbf{W} \in \mathbb{H}^{L \times N}$ a matriz dos pesos sinápticos da camada oculta com L neurônios com valores hipercomplexos. Dado um vetor de entrada $\mathbf{x} \in \mathbb{H}^N$, a saída da camada oculta é

$$\mathbf{h} = f(\mathbf{W}\mathbf{x}) \in \mathbb{H}^L, \quad (14)$$

em que $f : \mathbb{H} \rightarrow \mathbb{H}$ é a função de ativação definida componente-a-componente para vetores hipercomplexos. Por exemplo, a função tangente hiperbólica *split* $\tanh : \mathbb{H} \rightarrow \mathbb{H}$ é definida como segue para qualquer $p = p_0 + p_1\mathbf{i}_1 + \dots + p_n\mathbf{i}_n \in \mathbb{H}$:

$$\tanh(p) = \tanh(p_0) + \tanh(p_1)\mathbf{i}_1 + \dots + \tanh(p_n)\mathbf{i}_n. \quad (15)$$

De um modo similar, os parâmetros da camada de saída são representados por uma matriz $\mathbf{M} \in \mathbb{H}^{M \times L}$. A camada de saída da ELM não possui função de ativação. Portanto, a saída da ELM hipercomplexa é descrita pela equação

$$\mathbf{y} = \mathbf{M}\mathbf{h} \in \mathbb{H}^M. \quad (16)$$

Em analogia ao modelo real, a matriz \mathbf{W} da ELM hipercomplexa é gerada aleatoriamente enquanto que \mathbf{M} é determinada resolvendo um problema de quadrados mínimos. Especificamente, dado um conjunto de treinamento

$$\mathcal{T} = \{(\mathbf{x}_i, \mathbf{t}_i) : i = 1, \dots, P\} \subset \mathbb{H}^N \times \mathbb{H}^M, \quad (17)$$

primeiro organizamos as P entradas e saídas desejadas em matrizes $\mathbf{X} \in \mathbb{H}^{N \times P}$ e $\mathbf{T} \in \mathbb{H}^{M \times P}$, respectivamente. Os neurônios da camada oculta são determinados de forma aleatória, por exemplo, usando a equação

$$w_{ij} = \alpha(\text{randn} + \text{randn}\mathbf{i}_1 + \dots + \text{randn}\mathbf{i}_n), \quad (18)$$

em que α é um fator de escala e randn fornece uma amostra de uma distribuição normal com média 0 e variância 1. O fator de escala deve ser determinado de forma a evitar que os valores de \mathbf{h} estejam saturados, ou seja, as componentes de \mathbf{h} estejam relativamente bem distribuídos na imagem de f . Finalmente, na forma mais simples, os parâmetros da camada de saída são determinados resolvendo o problema quadrados mínimos

$$\min\{\|\mathbf{T} - \mathbf{M}\mathbf{H}\|_F : \mathbf{M} \in \mathbb{H}^{M \times L}\}, \quad (19)$$

em que $\mathbf{H} = f(\mathbf{W}\mathbf{X})$ é a matriz com a saída da camada oculta avaliada em \mathbf{X} e $\|\cdot\|_F$ denota a norma de Frobenius. A norma de Frobenius é definida como segue para uma matriz $\mathbf{A} \in \mathbb{H}^{M \times N}$

cujas entradas são $a_{ij} = (a_{ij})_0 + (a_{ij})_1 \mathbf{i}_1 + \dots + (a_{ij})_n \mathbf{i}_n$:

$$\|A\|_F = \sqrt{\sum_{i=1}^M \sum_{j=1}^N ((a_{ij})_0^2 + (a_{ij})_1^2 + \dots + (a_{ij})_n^2)}. \quad (20)$$

Recentemente, mostramos que (19) pode ser formulado como um problema de quadrados mínimos real para as álgebras de Cayley-Dickson generalizadas [7, 58]. Nesse projeto de pesquisa, propomos estender esse resultado para sistemas de números hipercomplexos mais gerais.

2.3.1 Experimento Computacional

Um experimento computacional para auto-codificação de imagens coloridas foi realizado para avaliar o desempenho da ELM em álgebras de Cayley-Dickson generalizadas. Num problema de auto-codificação, o treinamento da rede neural é realizado considerando o mesmo item como entrada e saída desejada. Porém, o número de neurônios na camada oculta é menor que a dimensão dos dados que serão codificados. Dessa forma, a rede neural é forçada a extrair as características mais relevantes para posteriormente reconstruir o objeto. Em vista dessa propriedade, auto-codificadores são poderosos extratores de características e, conseqüentemente, possuem papel importante na teoria da informação. Auto-codificadores também podem ser usados como modelos geradores [15].

Nesse experimento computacional preliminar, consideramos uma ELM real e quatro ELM definidas em álgebras de Cayley-Dickson, que são sistemas de números hipercomplexos de dimensão 4. Adotamos a função tangente hiperbólica para a ELM real e a tangente hiperbólica *split* definida por (15) para a ELM hipercomplexa. Os pesos sinápticos da camada oculta das ELM reais e hipercomplexas foram gerados com distribuição normal. Sobretudo, os modelos foram treinados usando 10.000 imagens coloridas da base CIFAR-10 e avaliadas em outras 10.000 imagens diferentes da mesma base de dados [39]. A ELM real possui 600 neurônios na camada oculta enquanto que as ELM hipercomplexas possuem 450 neurônios na camada oculta. Dessa forma, ambas as redes possuem o número total de parâmetros, isto é, 3.686.400 parâmetros que são representados por pontos flutuantes no computador. Detalhes sobre o experimento computacional pode ser obtido no artigo "*Extreme Learning Machines on Cayley-Dickson Algebra Applied for Color Image Auto-Encoding*" que foi apresentado recentemente no *IEEE World Congress on Computational Intelligence (WCCI) 2020*.

A Figura 1 mostra exemplos de imagens codificadas e decodificadas pelas ELMs real e hipercomplexas. Especificamente, a primeira linha apresenta uma imagem do conjunto de treinamento enquanto que a segunda linha mostra uma imagem do conjunto de teste, isto é, que não foi usada para determinar os pesos sinápticos da camada de saída da ELM. Nessa figura, as colunas marcadas com \mathbb{R} , \mathbb{Q} , \mathbb{H}_1 , \mathbb{H}_2 e \mathbb{H}_3 representam respectivamente os modelos real, quaterniônico e baseados em três álgebras de Cayley-Dickson não usuais. Visualmente, todos os modelos apresentaram resultados satisfatórios na decodificação das imagens coloridas, tanto

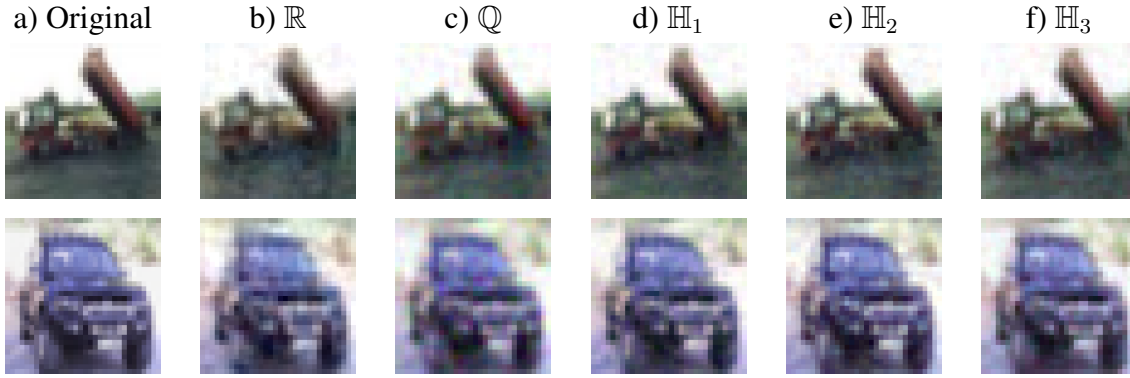


Figura 1: Imagem colorida original e as imagens decodificadas pelas ELM real (\mathbb{R}), quaterniônica (\mathbb{Q}) e baseadas em três álgebras de Cayley-Dickson não usuais (\mathbb{H}_1 , \mathbb{H}_2 e \mathbb{H}_3). Primeira linha: imagem do conjunto de treinamento. Segunda linha: imagem de teste.

Tabela 1: Média do PSNR produzido pelas ELM real e hipercomplexas no problema de auto-codificação de imagens coloridas.

Álgebra	Treinamento	Teste
	PSNR	PSNR
\mathbb{R}	27.3 ± 2.4	26.8 ± 2.7
\mathbb{Q}	28.9 ± 2.5	28.5 ± 2.7
\mathbb{H}_1	31.0 ± 2.5	30.5 ± 2.7
\mathbb{H}_2	31.2 ± 2.5	30.7 ± 2.7
\mathbb{H}_3	30.9 ± 2.5	30.4 ± 2.7

no conjunto de treinamento como de teste. Para uma melhor comparação do desempenho das ELM real e hipercomplexas, calculamos o *peak signal-to-noise ratio* entre a imagem original e a imagem decodificada por uma ELM. A Tabela 1 apresenta a média obtida por cada ELM nos conjuntos de treinamento e teste. Observe que não há uma discrepância significativa entre os valores de treino e teste. Portanto, tanto a ELM real como as redes hipercomplexas apresentaram boa capacidade de generalização. Contudo, observa-se que as versões hipercomplexas da ELM baseadas em álgebras de Cayley-Dickson não usuais apresentaram resultados melhores que a ELM real e quaterniônica. Essa observação é confirmada no *boxplot* apresentado na Figura 2. Essa observação nos motivou a investigar melhor modelos de ELM em álgebras com valores hipercomplexos.

3 Objetivos, Plano de Trabalho e Cronograma

Nesse projeto de pesquisa investigaremos regras de treinamento para redes neurais artificiais baseadas em números hipercomplexos, com ênfase no treinamento da rede de Hopfield e ELM

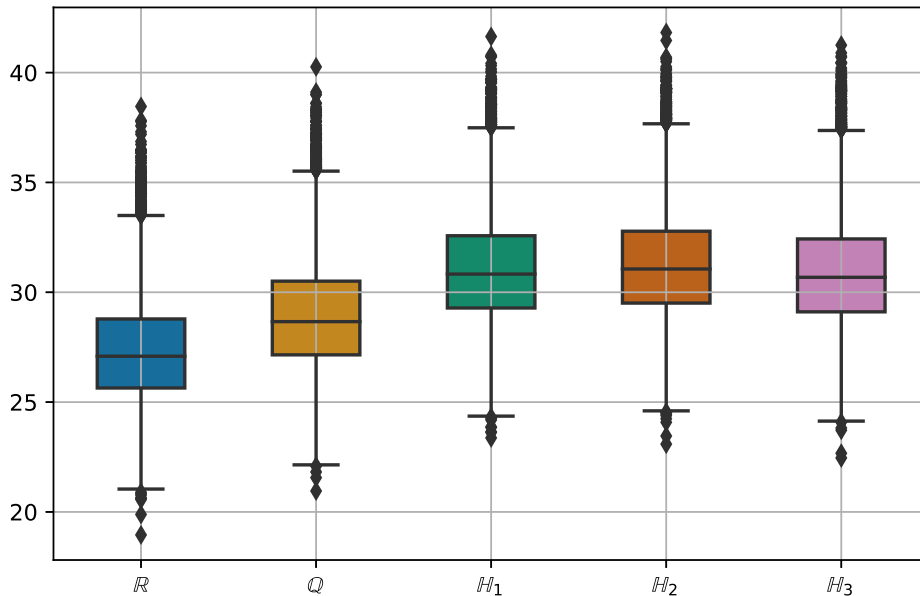


Figura 2: PSNR produzido pelas ELM real e com valores hipercomplexos no conjunto de teste.

com valores hipercomplexos. Para tanto, serão executadas as seguintes atividades:

1. Atividade 1 – Revisão Bibliográfica:

Existe na literatura muitos trabalhos sobre redes neurais artificiais usando números hipercomplexos. No primeiro semestre faremos uma revisão bibliográfica atualizada de generalizações de rede neurais artificiais usando números hipercomplexos. Também faremos uma revisão bibliográfica sobre as redes de Hopfield, as máquinas de aprendizado extremo (ELM) e a aplicação desses modelos. Os modelos estudados serão implementados em softwares como `python` e `julia`.

2. Atividade 2 – Rede de Hopfield com Valores Hipercomplexos:

No segundo e terceiro semestres investigaremos regras de treinamento para versões hipercomplexas da rede de Hopfield. Realizaremos experimentos computacionais para avaliar a capacidade de armazenamento e a tolerância à ruído desses modelos. Também investigaremos a aplicação em problemas de recuperação de imagens e classificação de padrões.

3. Atividade 3 – Máquinas de Aprendizado Extremo com Valores Hipercomplexos:

No quarto e quinto semestres estenderemos as máquinas de aprendizado extremo reais para sistemas de números hipercomplexos. Tal como na rede de Hopfield, realizaremos experimentos computacionais para avaliar o desempenho dos novos modelos hipercomplexos, por exemplo, usando problema de auto-codificação. Em particular, iremos comparar as redes neurais hipercomplexas com modelos tradicionais.

4. Atividade 4 – Desenvolvimento de Outros Modelos Hipercomplexos:

No sexto semestre estudaremos outras arquiteturas de redes neurais hipercomplexas. Por exemplo, iremos investigar generalizações da rede neural quaterniônica com operadores geométricos proposta por Isokawa e colaboradores usando álgebra geométrica [28].

As quatro atividades listadas acima seguirão o seguinte cronograma:

Mar/2022 a Ago/2022	Set/2022 a Ago/2023	Set/2023 a Ago/2024	Set/2024 a Fev/2025
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 regras de treinamento para redes neurais artificiais com valores hipercomplexos. Em particular, investigaremos regras de treinamento para a rede de Hopfield e as máquinas de aprendizado extremo, ambas definidas usando operações de um sistema de números hipercomplexos. Conforme mencionado na introdução, a rede de Hopfield pode ser usada para implementar memórias associativas e resolver problemas de otimização. As máquinas de aprendizado extremo foram usadas com sucesso em problemas de classificação, previsão, reconhecimento de padrões e auto-codificação. Portanto, esse projeto de pesquisa deve contribuir nessas aplicações.

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.

5 Resumo das Atividades de Pesquisas do Proponente

Segue abaixo um quadro resumo da produção científica e orientações do proponente:

Tipo	Quantidade
Publicações em Periódicos com Política Editorial	33
Capítulos de Livro	4
Edição de Livro	1
Orientações Concluídas	Mestrado: 9 Doutorado: 5
Citações na Literatura Científica	Publons: 583 Scopus: 751 Google Scholar: 1373
Índice h	Publons: 13 Scopus: 15 Google Scholar: 18

Produção Científica Relevante

Abaixo uma lista cronológica contendo 10 artigos publicados em revistas internacionais:

1. P. Sussner and M.E. Valle: Gray-Scale Morphological Associative Memories, IEEE Transactions on Neural Networks, Volume 17, Issue 3, May 2006. Page(s): 559 – 570.

Citações – Google Scholar: 169, Publons 96

2. P. Sussner and M.E. Valle: Implicative Fuzzy Associative Memories, IEEE Transactions on Fuzzy Systems, Volume 14, Issue 6, December 2006. Page(s): 793 - 807.

Citações – Google Scholar: 131, Publons 62

3. M.E. Valle and P. Sussner: A general framework for fuzzy morphological associative memories, Fuzzy Sets and Systems, Volume 159, Issue 7, April 2008. Page(s): 747 - 768.

Citações – Google Scholar: 137, Publons 50

4. P. Sussner and M.E. Valle: Classification of Fuzzy Mathematical Morphologies Based on Concepts of Inclusion Measure and Duality, Journal of Mathematical Imaging and Vision, Volume 32, Issue 2, October 2008. Page(s): 139 - 159.

Citações – Google Scholar: 101, Publons 53

5. M.E. Valle and P. Sussner: Storage and Recall Capabilities of Fuzzy Morphological Associative Memories with Adjunction-Based Learning, *Neural Networks*, Volume 24, Issue 1, January 2011. Page(s): 75-90.

Citações – Google Scholar: 57, Plubons 30

6. M.E. Valle: Complex-Valued Recurrent Correlation Neural Networks, *IEEE Transactions on Neural Networks and Learning Systems*, Volume 25, Issue 09, September 2014. Page(s): 1600 - 1612.

Citações – Google Scholar: 52, Plubons 39

7. M.E. Valle and F.Z. Castro: On the Dynamics of Hopfield Neural Networks on Unit Quaternions, *IEEE Transactions on Neural Networks and Learning Systems*, Volume 29, Issue 6, June 2018. Pages(s): 2464 - 2471.

Citações – Google Scholar: 42, Plubons 33

8. F.Z. Castro and M.E. Valle: A broad class of discrete-time hypercomplex-valued Hopfield neural networks, *Neural Networks*, Volume 122, February 2020. Page(s): 54-67.

Citações – Google Scholar: 24, Plubons 14

9. M.E. Valle: Reduced Dilation-Erosion Perceptron for Binary Classification, *Mathematics*, Volume 8, Issue 4, April 2020.

Citações – Google Scholar: 7, Plubons 2

10. M.E. Valle, R.A. Lobo: Hypercomplex-valued recurrent correlation neural networks, *Neurocomputing*, Volume 432, April 2021. Pages(s): 111 - 123.

Citações – Google Scholar: 3, Plubons 2

Cabe destacar que não foram listados todos os artigos publicados pelo proponente. A lista completa das publicações podem ser encontradas no *Currículo Lattes* ou na página pessoal do proponente, em www.ime.unicamp.br/~valle.

Formação de Recursos Humanos e Divulgação Científica

Com respeito à formação de recursos humanos, o proponente já concluiu 9 (nove) orientações de mestrado e 5 (cinco) orientações de doutorado. Atualmente, orienta 1 (um) aluno de

pós-doutorado, 5 (cinco) alunos de doutorado e 2 (dois) alunos de mestrado. Informações detalhadas sobre as orientações concluídas e em andamento encontram-se no *Currículo Lattes* do proponente.

Referente a divulgação científica e outras informações relevantes, destacam-se o seguinte:

1. O proponente é editor associado das revistas *Computational and Applied Mathematics* e *Journal of Computational Mathematics and Data Science*.
2. O proponente ministrou a palestra intitulada “*Fuzzy Associative Memories: Theory and Applications*”, no *IEEE World Congress on Computational Intelligence (WCCI 2018)*, o principal evento internacional da área de inteligência computacional. Detalhes sobre a palestra estão disponíveis em <http://www.ecomp.poli.br/~wcci2018/speakers/>.
3. O proponente, junto com os pesquisadores Akira Hirose (University of Tokyo, Tokyo, Japan), Danilo Mandic (Imperial College, London, U.K.) e Igor Aizenberg (Manhattan College, New York City, USA), organizaram a sessão especial sobre redes neurais com valores complexos e quaterniônicos no *2018 International Joint Congress on Neural Networks (IJCNN 2018)*. Detalhes sobre a sessão especial estão disponíveis em <http://www.ecomp.poli.br/~wcci2018/ijcnn-sessions/#ijcnn3>
4. O proponente, como membro da Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC), participou do comitê científico do XXXVI Congresso Nacional de Matemática Aplicada e Computacional (CNMAC 2017) e da comissão de organização do XXXVII Congresso Nacional de Matemática Aplicada e Computacional (CNMAC 2018).
5. O proponente é coordenador do comitê científico do VI Congresso Brasileiro de Sistemas Fuzzy (VI CBSF), que será realizado em Novembro de 2021. Detalhes sobre o evento estão disponíveis em <https://sites.google.com/unesp.br/vicbsf>.
6. O proponente organizou o IV Congresso Brasileiro de Sistemas Fuzzy (IV CBSF), que ocorreu em Novembro de 2016. Detalhes sobre o evento estão disponíveis em www.ime.unicamp.br/~cbsf4.
7. O proponente editou o livro: Valle, M.E., Dimuro, G.P., Santiago, R.H.N., Esmi, E. *Recentes Avanços em Sistemas Fuzzy*. 1. ed. São Carlos: Sociedade Brasileira de Matemática Aplicada e Computacional, 2016. v. 1. 548p.

Referências

- [1] AIZENBERG, I., AND GONZALEZ, A. Image Recognition using MLMVN and Frequency Domain Features. In *2018 International Joint Conference on Neural Networks (IJCNN)* (2018), pp. 1–8.
- [2] AIZENBERG, I. N. *Complex-Valued Neural Networks with Multi-Valued Neurons*, vol. 353 of *Studies in Computational Intelligence*. Springer, 2011.
- [3] AIZENBERG, N. N., IVASKIV, Y. L., AND POSPELOV, D. A. A certain generalization of threshold functions. *Doklady Akademii Nauk SSSR* 196 (1971), 1287–1290.
- [4] ARENA, P., FORTUNA, L., MUSCATO, G., AND XIBILIA, M. G. Multilayer Perceptrons to Approximate Quaternion Valued Functions. *Neural Networks* 10, 2 (1997), 335–342.
- [5] ARENA, P., FORTUNA, L., MUSCATO, G., AND XIBILIA, M. G. Quaternion algebra. In *Neural Networks in Multidimensional Domains*, vol. 234 of *Lecture Notes in Control and Information Sciences*. Springer London, 1998, pp. 43–47.
- [6] BAEZ, J. C. The octonions. *Bulletin of the American Mathematical Society* 39 (2002), 145–205.
- [7] BROWN, R. B. On generalized Cayley-Dickson algebras. *Pacific Journal of Mathematics* 20, 3 (1967), 415–422.
- [8] BUCHHOLZ, S., AND LE BIHAN, N. Polarized signal classification by complex and quaternionic multi-layer perceptrons. In *International Journal of Neural Systems* (4 2008), vol. 18, pp. 75–85.
- [9] CASTRO, F. Z., AND VALLE, M. E. Some Remarks on the Stability of Discrete-Time Complex-Valued Multistate Hopfield Neural Networks. In *Proceedings Series of the Brazilian Society of Computational and Applied Mathematics. Sociedade Brasileira de Matemática Aplicada e Computacional*. (Campinas – Brazil, 2018), pp. 10327–10328.
- [10] CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems* 2, 4 (12 1989), 303–314.
- [11] DE CASTRO, F. Z., AND VALLE, M. E. A broad class of discrete-time hypercomplex-valued Hopfield neural networks. *Neural Networks* 122 (2020), 54–67.
- [12] DING, T., AND HIROSE, A. Online Regularization of Complex-Valued Neural Networks for Structure Optimization in Wireless-Communication Channel Prediction. *IEEE Access* 8 (2020), 143706–143722.

- [13] FORTUNA, L., MUSCATO, G., AND XIBILIA, M. G. An hypercomplex neural network platform for robot positioning. In *Circuits and Systems, 1996. ISCAS '96., Connecting the World., 1996 IEEE International Symposium on* (1996), vol. 3, pp. 609–612 vol.3.
- [14] GAN, J. Discrete Hopfield neural network approach for crane safety evaluation. In *2017 International Conference on Mechanical, System and Control Engineering (ICMSC)* (2017), pp. 40–43.
- [15] GÉRON, A. *Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Incorporated, 2017.
- [16] GOODFELLOW, I., BENGIO, Y., AND COURVILLE, A. *Deep Learning*. MIT Press, 2016.
- [17] 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.
- [18] HAYKIN, S. *Neural Networks and Learning Machines*, 3rd ed. Prentice-Hall, Upper Saddle River, NJ, 2009.
- [19] HIROSE, A. *Complex-Valued Neural Networks*, 2nd editio ed. Studies in Computational Intelligence. Springer, Heidelberg, Germany, 2012.
- [20] HOPFIELD, J. J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences* 79, 8 (1982), 2554–2558.
- [21] HOPFIELD, J. J., AND TANK, D. W. Neural computation of decisions in optimization problems. *Biological Cybernetics* 52 (1985), 141–152.
- [22] HUANG, G.-B., WANG, D., AND LAN, Y. Extreme learning machines: a survey. *Int. J. Machine Learning & Cybernetics* 2, 2 (2011), 107–122.
- [23] HUANG, G.-B., ZHOU, H., DING, X., AND ZHANG, R. Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42, 2 (2011), 513–529.
- [24] HUANG, G.-B., ZHU, Q.-Y., AND SIEW, C.-K. Extreme learning machine: theory and applications. *Neurocomputing* 70, 1-3 (2006), 489–501.
- [25] HUANG, G.-B., ZHU, Q.-Y., SIEW, C.-K., AND OTHERS. Extreme learning machine: a new learning scheme of feedforward neural networks. *Neural networks* 2 (2004), 985–990.
- [26] 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.

- [27] ISOKAWA, T., KUSAKABE, T., MATSUI, N., AND PEPPER, F. Quaternion neural network and its application. In *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)* (2003), vol. 2774 PART 2.
- [28] ISOKAWA, T., NISHIMURA, H., AND MATSUI, N. Commutative quaternion and multistate Hopfield neural networks. In *The 2010 International Joint Conference on Neural Networks (IJCNN)* (2010), pp. 1–6.
- [29] 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.
- [30] ISOKAWA, T., YAMAMOTO, H., NISHIMURA, H., YUMOTO, T., KAMIURA, N., AND MATSUI, N. Complex-Valued Associative Memories with Projection and Iterative Learning Rules. *Journal of Artificial Intelligence and Soft Computing Research* 8, 3 (2018), 237–249.
- [31] JANKOWSKI, S., LOZOWSKI, A., AND ZURADA, J. M. Complex-Valued Multi-State Neural Associative Memory. *IEEE Transactions on Neural Networks* 7 (1996), 1491–1496.
- [32] KIM, T., AND ADALI, T. Approximation by fully complex multilayer perceptrons. *Neural Computation* 15, 7 (2003).
- [33] KOBAYASHI, M. Hyperbolic Hopfield Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* 24, 2 (2013).
- [34] KOBAYASHI, M. Quaternionic Hopfield neural networks with twin-multistate activation function. *Neurocomputing* 267 (2017), 304–310.
- [35] KOBAYASHI, M. Symmetric Complex-Valued Hopfield Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* 28, 4 (2017), 1011–1015.
- [36] KOBAYASHI, M. Hyperbolic Hopfield neural networks with directional multistate activation function. *Neurocomputing* 275 (2018), 2217–2226.
- [37] KOBAYASHI, M. Twin-multistate commutative quaternion Hopfield neural networks. *Neurocomputing* 320 (2018), 150–156.
- [38] KOBAYASHI, M. Noise Robust Projection Rule for Hyperbolic Hopfield Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* (2019), 1–5.
- [39] KRIZHEVSKY, A. Learning multiple layers of features from tiny images. Tech. rep., University of Toronto, 2009.

- [40] KUROE, Y. Models of Recurrent Clifford Neural Networks and Their Dynamics. In *Complex-Valued Neural Networks*. Wiley-Blackwell, 2013, ch. 6, pp. 133–151.
- [41] KUROE, Y., AND IIMA, H. A model of Hopfield-type octonion neural networks and existing conditions of energy functions. In *2016 International Joint Conference on Neural Networks (IJCNN)* (2016), pp. 4426–4430.
- [42] KUSAMICHI, H., ISOKAWA, T., MATSIU, 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)* (2004), pp. 101–106.
- [43] LEE, D.-L. Improvements of complex-valued Hopfield associative memory by using generalized projection rules. *IEEE Transactions on Neural Networks* 17, 5 (2006), 1341–1347.
- [44] LI, J., LI, X., HUANG, B., AND ZHAO, L. Hopfield Neural Network Approach for Supervised Nonlinear Spectral Unmixing. *IEEE Geoscience and Remote Sensing Letters* 13, 7 (2016), 1002–1006.
- [45] 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 (2011), 1323–1333.
- [46] MANDIC, D. P., JAHANCHAH, C., AND TOOK, C. C. A Quaternion Gradient Operator and Its Applications. *Signal Processing Letters, IEEE* 18, 1 (1 2011), 47–50.
- [47] MINEMOTO, T., ISOKAWA, T., NISHIMURA, H., AND MATSUI, N. Quaternionic multistate Hopfield neural network with extended projection rule. *Artificial Life and Robotics* 21, 1 (3 2016), 106–111.
- [48] MINEMOTO, T., ISOKAWA, T., NISHIMURA, H., AND MATSUI, N. Feed forward neural network with random quaternionic neurons. *Signal Processing* 136 (2017), 59–68.
- [49] MÜEZZINOĞLU, M. K., GÜZELİŞ, C., AND ZURADA, J. M. A new design method for the complex-valued multistate hopfield associative memory. *IEEE Transactions on Neural Networks* 14, 4 (7 2003), 891–899.
- [50] NITTA, T. *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*. Igi Global, Hershey - PA, USA, 2009.
- [51] NITTA, T., AND KUROE, Y. Hyperbolic gradient operator and hyperbolic backpropagation learning algorithms. *IEEE Transactions on Neural Networks and Learning Systems* 29, 5 (2018), 1689–1702.

- [52] NOEST, A. J. Associative Memory in Sparse Phasor Neural Networks. *EPL (Europhysics Letters)* 6, 5 (1988), 469.
- [53] NOEST, A. J. Discrete-state phasor neural networks. *Physical Review A* 38, 4 (1988), 2196–2199.
- [54] PAJARES, G. A Hopfield Neural Network for Image Change Detection. *IEEE Transaction on Neural Networks* 17, 5 (2006), 1250–1264.
- [55] PAJARES, G., GUIJARRO, M., AND RIBEIRO, A. A Hopfield Neural Network for Combining Classifiers Applied to Textured Images. *Neural Networks* 23, 1 (2010), 144–153.
- [56] POPA, C. A. Complex-Valued Deep Boltzmann Machines. In *Proceedings of the International Joint Conference on Neural Networks* (10 2018), vol. 2018-July, Institute of Electrical and Electronics Engineers Inc.
- [57] POPA, C. A. Learning Algorithms for Quaternion-Valued Neural Networks. *Neural Processing Letters* 47, 3 (2018).
- [58] SCHAFFER, R. D. On the Algebras Formed by the Cayley-Dickson Process. *American Journal of Mathematics* 76, 2 (1954), 435–446.
- [59] SERPEN, G. Hopfield Network as Static Optimizer: Learning the Weights and Eliminating the Guesswork. *Neural Processing Letters* 27, 1 (2008), 1–15.
- [60] SHANG, F., AND HIROSE, A. Quaternion Neural-Network-Based PolSAR Land Classification in Poincare-Sphere-Parameter Space. *IEEE Transactions on Geoscience and Remote Sensing* 52 (2014), 5693–5703.
- [61] SHENITZER, A., KANTOR, I. L., AND SOLODOVNIKOV, A. S. *Hypercomplex Numbers: An Elementary Introduction to Algebras*. Springer New York, 1989.
- [62] SMITH, K., PALANISWAMI, M., AND KRISHNAMOORTHY, M. Neural Techniques for Combinatorial Optimization with Applications. *IEEE Transactions on Neural Networks* 9, 6 (1998), 1301–1318.
- [63] 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, 2014).
- [64] UJANG, B. C., TOOK, C. C., AND MANDIC, D. P. Quaternion-valued nonlinear adaptive filtering. *IEEE Transactions on Neural Networks* 22, 8 (2011).

- [65] VALLE, M. E., AND CASTRO, F. Z. On the Dynamics of Hopfield Neural Networks on Unit Quaternions. *IEEE Transactions on Neural Networks and Learning Systems* 29, 6 (2018), 2464–2471.
- [66] VALLEJO, J. R., AND BAYRO-CORROCHANO, E. Clifford Hopfield Neural Networks. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)* (2008), pp. 3609–3612.
- [67] VAZ, J., AND DA ROCHA, R. *An Introduction to Clifford Algebras and Spinors*. Oxford University Press, 2016.