

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E
ENGENHARIA DE MATERIAIS

APLICAÇÃO DE REDES NEURAIS PARA CLASSIFICAÇÃO DE
MICROINCLUSÕES DE SULFETO DE MANGANÊS EM AÇOS

Marcos Fernando de Oliveira Filho

São Carlos-SP
2023

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E
ENGENHARIA DE MATERIAIS**

**APLICAÇÃO DE REDES NEURAIS PARA CLASSIFICAÇÃO DE
MICROINCLUSÕES DE SULFETO DE MANGANÊS EM AÇOS**

Marcos Fernando de Oliveira Filho

Dissertação apresentada ao
Programa de Pós-Graduação em Ciência e
Engenharia de Materiais como requisito parcial
à obtenção do título de MESTRE EM CIÊNCIA
E ENGENHARIA DE MATERIAIS.

Orientador: Dr. José Eduardo Spinelli

Coorientador: Prof. Tomaz Toshimi Ishikawa

São Carlos-SP

2023

DEDICATÓRIA

Dedico este trabalho ao meu Avô José Custódio da Silva (*in memoriam*), que me ensinou o valor do trabalho e do estudo, sendo a minha referência de vida me inspirou durante todo o tempo que estivemos juntos, me apoiando e patrocinando as minhas decisões.

VITAE

Bacharel em Engenharia de Produção pela Universidade Metodista de Piracicaba (UNIMEP) Câmpus Santa Bárbara D'Oeste (2016).



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência e Engenharia de Materiais

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Marcos Fernando de Oliveira Filho, realizada em 24/02/2023.

Comissão Julgadora:

Prof. Dr. José Eduardo Spinelli (UFSCar)

Prof. Dr. Conrado Ramos Moreira Afonso (UFSCar)

Prof. Dr. Rodrigo André Valenzuela Reyes (IFSP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência e Engenharia de Materiais.

“Cada adversidade, cada fracasso, cada dor de cabeça carrega consigo
a semente de um benefício igual ou maior.”
(Napoleon Hill)

AGRADECIMENTOS

Em primeiro lugar gostaria de agradecer a Deus, pelo dom da vida que me concedeu. Agradeço também pelo dom da sabedoria que me levou a realização do Mestrado e me sustentou em todos os momentos de dificuldades.

Agradeço à minha família por todo o amor e apoio que sempre me deram. Vocês sempre foram a minha estrutura e sem vocês eu jamais teria conseguido. Agradeço também a minha esposa Gabriela Gomes Oliveira por toda paciência que precisou ter nesse período, além disso a compreensão de perder finais de semana comigo para estudar, escrever e ler artigos. Com certeza você tem toda a minha gratidão.

Agradeço aos meus professores e orientadores José Eduardo Spinelli e Tomaz Toshimi Ishikawa que demonstraram paciência em momentos de instabilidade. Tenho certeza de que fiz a escolha certa quando solicitei que vocês fossem meus orientadores. Spinelli: te agradeço imensamente por acreditar no potencial desse trabalho e, por muitas vezes, me motivar com a sua visão de realidade. Tomaz: te agradeço pela paciência e por ter sido a primeira pessoa a acreditar em mim e nesse trabalho. Seu perfil calmo muitas vezes me deu força para superar os desafios que esse trabalho nos trouxe.

À Empresa Villares Metals S/A que contribuiu para a realização desse trabalho através de recursos, materiais e disponibilidade para a realização.

Agradeço aos colegas Pierre D Amelio Briquet Caradec e Rafael Sola de Paula de Ângelo Calsaverini que apresentaram ao mundo do “*machine learning*”, sendo fundamentais no desenvolvimento desse trabalho.

Agradeço aos meus amigos Alexsandro Brandão, Ana Paula Medeiros, Ericsson Rodrigo S. Lima e Jairo Almeida de Melo do laboratório da Villares Metals que me ajudaram na preparação e avaliação das amostras utilizadas nesse trabalho. O meu muito obrigado ao meu antigo supervisor Vinicius Ottani, que sempre me apoiou e que muitas vezes me ajudou.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001

RESUMO

Os fatores que mensuram a influência das inclusões nas propriedades dos aços são a distribuição, morfologia e tamanho das inclusões. A norma utilizada como referência é a ASTM E45, a qual explica o método de varredura da amostra no intuito de analisar e classificar as inclusões em aços. Os métodos manuais são os mais utilizados nos laboratórios e empresas metalúrgicas por apresentarem baixo custo. No entanto os automáticos se caracterizam pelo alto custo de operação, o que dificulta seu emprego nas indústrias. Os modelos de redes neurais fazem parte das novas tecnologias sendo extremamente vantajosos para diversas aplicações. O presente trabalho de Mestrado teve como motivação o emprego de uma nova metodologia de classificação de inclusões em aços utilizando modelos de redes neurais. O intuito foi alcançar a maior acurácia possível na classificação de severidades de inclusões de MnS utilizando imagens capturadas a partir de corpos de prova de diferentes aços processados em um laboratório de uma indústria do setor metalúrgico. O processo de classificação por redes neurais foi validado por meio de comparação com resultados obtidos manualmente, apresentando alta acurácia e rapidez na tomada de decisão. Para entender a efetividade da rede neural foi aplicado o conceito de matriz de confusão, que são tabelas comparativas do número de imagens que a rede neural trouxe como predição de severidades em relação aos valores reais classificados manualmente. Os primeiros resultados dos testes não foram positivos, sendo necessária uma reclassificação de todas as imagens do banco de dados para reduzir a confusão da rede neural. Após reclassificação e aplicação da técnica “*Dropout*”, os resultados dos testes foram superiores aos anteriores. Em conclusão, os resultados de acurácia do treinamento e validação melhoraram, sendo possível realizar a comparação entre a classificação manual e a automática. De forma geral a rede neural representou rapidez na tomada de decisão se mostrando ser uma ferramenta em potencial para a classificação de inclusões em aços.

Palavras-chave: Inclusões; redes neurais; banco de dados; Aços; MnS; Análise de imagem.

ABSTRACT

APPLICATION OF NEURAL NETWORKS FOR CLASSIFICATION OF MICROINCLUSIONS OF MANGANESE SULFIDE IN STEEL

The factors that measure the influence of inclusions on steel properties are the distribution, morphology and size of inclusions. The standard used as a reference is ASTM E45, which explains the sample scanning method in order to analyze and classify inclusions in steels. Manual methods are the most used in laboratories and metallurgical companies because of their low cost, while automatic methods are characterized by high operating costs, which makes their use in industries difficult. Neural network models, on the other hand, are part of new technologies and are extremely advantageous for several applications. This Master's work was motivated by the use of a new methodology for classifying inclusions in steels using neural network models. The aim was to achieve the highest possible accuracy in classifying the severities of MnS inclusions using images captured from specimens processed in a laboratory of a metallurgical industry. The classification process by neural networks was validated through comparison with results obtained manually, showing higher accuracy and speed in decision making. The results also showed that the classification of severities with a smaller number of images in the database presented a lower accuracy. To understand the effectiveness of the neural network, the concept of confusion matrix was applied, which are comparative tables of the number of images that the neural network brought as a prediction of severities in relation to the actual manually classified values. The first test results were not positive, requiring a reclassification of all images in the database to reduce confusion in the neural network. After reclassification and application of the Dropout technique, the test results were superior to the previous ones. In conclusion, the training and validation accuracy results improved, making it possible to compare manual and automatic classification. In general, the neural network represented speed in decision making, proving to be a potential tool for the classification of inclusions.

Keywords: Non-metallic inclusion; neural network; database; Steel; MnS; Image analysis.

ÍNDICE DE ASSUNTOS

	Pág.
FOLHA DE APROVAÇÃO.....	i
AGRADECIMENTOS	iii
RESUMO.....	v
ABSTRACT	vii
ÍNDICE DE ASSUNTOS	ix
ÍNDICE DE TABELAS	xi
ÍNDICE DE FIGURAS	xiii
SÍMBOLOS E ABREVIATURAS.....	xvii
1 INTRODUÇÃO.....	1
1.1 OBJETIVOS	4
2 REVISÃO BIBLIOGRÁFICA	5
2.1 A importância e classificação dos aços	5
2.2 Impacto das inclusões no processamento e nas propriedades dos aços.....	10
2.3 Aços resulfurados.....	13
2.4 Avaliação tradicional de inclusões em aços.....	15
2.5 Classificação manual, redes neurais e métodos automatizados de classificação.....	16
2.5.1 Classificação manual	16
2.5.2 Métodos automatizados de classificação.....	17
2.5.3 Redes Neurais	19
2.6 Redes Neurais aplicado em estudos de Engenharia de Materiais.....	25
3 MATERIAIS E MÉTODOS	29
3.1 Seleção dos materiais.....	30
3.2 Preparação metalográfica para a análise de inclusões.....	31
3.3 Preparação dos corpos de prova.....	34
3.4 Métodos de classificação de inclusões não metálicas conforme ASTM E45.	37
3.5 Criação do banco de dados	40
3.6 Classificação das imagens do banco de dados.....	42
3.7 Rede neural.....	45

3.8	Testes e validações na rede neural.....	46
3.9	Fluxograma de aplicação da rede neural	47
4	RESULTADOS E DISCUSSÃO	51
4.1	Resultados utilizando imagens computadorizadas	51
4.2	Resultados utilizando o banco de dados com imagens reais.	53
5	CONCLUSÕES.....	65
6	SUGESTÕES PARA TRABALHOS FUTUROS	67
7	REFERÊNCIAS BIBLIOGRÁFICAS	69
	APÊNDICE A – Imagens que compõe o banco de dados.....	75
	APÊNDICE B – Códigos de programação utilizados.....	77

ÍNDICE DE TABELAS

	Pág.
Tabela 1 - Tabela de classificação de aços e principais aplicações. Fonte: Villares Metals.....	7
Tabela 2 - Classificação dos aços de acordo com as normas SAE.[22]	9
Tabela 3 - Resultados dos ensaios mecânicos obtidos por Tervo [2].	11
Tabela 4 - Parâmetros de largura e diâmetro para realizar a classificação de inclusões em aços. Fonte: Norma ASTM E45.[8].....	16
Tabela 5 - Composição química dos aços utilizados nas análises. Fonte: Villares Metals.....	31
Tabela 6 - Valores mínimos para a classificação da severidade. (Métodos A, D e E) [8].....	43
Tabela 7 - Número de imagens utilizada no primeiro treinamento da rede neural.	44
Tabela 8 – Número de imagens utilizadas no segundo treinamento da rede neural.	44
Tabela 9 – Número de imagens utilizadas no treinamento final da rede neural.	45
Tabela 10 - Número de imagens por severidade utilizadas no primeiro teste. .	54
Tabela 11 - Quantidade de imagens por severidade utilizadas no segundo teste.	56
Tabela 12 - Quantidade de imagens no banco de dados por severidade após a reclassificação.....	59

ÍNDICE DE FIGURAS

	Pág.
Figura 1 - Relação volume de produção de aço x país. Fonte: <i>World Steel Association</i> . [19]	6
Figura 2 - Classificação geral de metais ferrosos.....	8
Figura 3 - MnS e TiN em aço duplex[3].....	13
Figura 4 - Amostra do aço AISI 430F preparada para análise de inclusão. As linhas cinza escuras são os sulfetos de manganês (MnS). Fonte: Villares Metals.	15
Figura 5 - Método de varredura manual no corpo de prova para classificação de inclusões. Fonte: Norma ASTM E45. [8]	17
Figura 6 - Imagem representativa de diferentes planos das microinclusões em uma amostra representativa.[30].....	19
Figura 7 - Representação esquemática de um neurônio com seus parâmetros de entrada x , k , pesos w , k , "bias" b , função de ativação f e saída y . [32]	20
Figura 8 - Gráfico da função sigmoideal, onde c_1 e c_2 são constantes.[33]	20
Figura 9 - Diagrama de grafos de uma rede neural com sua camada de entrada, camada oculta, camada de saída e suas sinapses indicadas por setas.[34] ...	21
Figura 10 - Ilustração do método de funcionamento da camada convolucional, onde o núcleo é 3x3 e varre a imagem gerando um valor para cada passo.[36]	22
Figura 11 - Ilustração da camada " <i>max pooling</i> " onde a camada de entrada é separada em um grid 2x2 e os maiores valores dentro da região são selecionados.[37]	22
Figura 12 - Esquema da operação " <i>Dropout</i> ", sendo à esquerda uma rede neural com camadas densas e a direita após aplicar a operação.[38].....	23
Figura 13 - R-CNN: Exemplo de arquitetura e características CNN em redes neurais.[40].....	24
Figura 14 - Exemplo da leitura e interpretação da matriz de confusão.	28
Figura 15 - Esquema ilustrativo de retirada de amostra para preparação e face de análise.....	32

Figura 16 - Equipamento cut-off utilizado para realizar o corte dos corpos de prova de aço.....	34
Figura 17 - Lixadeira tipo correia com lixa 100 utilizada para lixar os corpos de prova de aço.....	35
Figura 18 - Embutidora da fabricante Buehler utilizada para embutir os corpos de prova de aço.....	35
Figura 19 - Máquina de polimento automático utilizada para realizar o polimento dos corpos de prova de aço.	36
Figura 20 - Microscópio analisador de imagem utilizado para a criação do banco de dados de imagens dos aços.....	37
Figura 21 - Reticulado utilizado para classificação das inclusões. Fonte: Norma ASTM E45.[8].....	38
Figura 22 - Reticulado utilizado para a classificação de inclusões no método B. Fonte: Norma ASTM E45.[8].....	38
Figura 23 - Chart utilizado para a subclassificação das inclusões no método do comprimento. Fonte: Norma ASTM E45.[8].....	39
Figura 24 – Fluxograma da metodologia e etapas desse trabalho.....	42
Figura 25 - Exemplo de imagem que foi excluída pela presença de riscos, os quais podem confundir a rede neural.	45
Figura 26 - Esquema das camadas da rede neural utilizada.	46
Figura 27 – Fluxograma de aplicação da rede neural.	49
Figura 28 - Exemplos de imagens com elipses com diferentes razões de aspecto de 1, 3, 5 e 7.	52
Figura 29 - Resultados de aprendizado da rede neural para o caso de elipses, com a acurácia e a função perda para treinamento e para validação.	53
Figura 30 - Resultados de aprendizado da rede neural para o caso de fotos reais, com a acurácia e a função perda para o banco de imagens para treinamento e para validação utilizando 40 épocas.	54
Figura 31 - Resultados de aprendizado da rede neural para o caso de fotos reais, com a acurácia e a função perda para o banco de imagens para treinamento e para validação utilizando 100 épocas.	55

Figura 32 - Matriz de confusão após o primeiro treinamento, comparando classificação manual e classificação automática (100 épocas).....	56
Figura 33 - Resultados de aprendizado da rede neural no segundo teste após ampliar o banco de dados, treinamento e validação, utilizando 40 épocas.....	57
Figura 34 - Matriz de confusão após o segundo treinamento, comparando entre classificação manual e classificação automática (40 épocas).....	58
Figura 35 - Resultados de aprendizado da rede neural no terceiro teste após reclassificação do banco de dados, treinamento e validação utilizando 200 épocas.....	59
Figura 36 - Matriz de confusão após o terceiro treinamento, comparando classificação manual e classificação automática (200 épocas).....	60
Figura 37 - Resultados de aprendizado da rede neural no quarto teste após aplicar a técnica de Dropout, treinamento e validação utilizando 200 épocas.	61
Figura 38 - Matriz de confusão após o quarto treinamento, comparando classificação manual e classificação automática (200 épocas).....	62
Figura 39 - Gráfico comparativo da evolução dos resultados dos treinamentos da rede neural	63

SÍMBOLOS E ABREVIATURAS

AISI: American Iron and Steel Institute

ASTM: American Standard for Testing of Materials

CNN: Convolutional Neural Network

DIN50602: Metallographic Test Methods - Microscopic Examination of Stainless Steels for Non-Metallic Inclusions with Image Series

Fe: Ferro

FeS: Sulfeto de Ferro (II)

ML: Machine Learning

MnS: Sulfeto de Manganês

MTBF: Mean Time Between Failures

MTTR: Mean Time to Repair.

R-CNN: Region-Based Convolutional Neural Network

SAE: Society of Automotive Engineers

SPARK-DAT: Software de espectrometria de emissão óptica.

V430FG ®: Aço resulfurado produzido pela Villares Metals S/A

VAR: Refusão à vácuo

VIM: Fusão por indução à vácuo

b: "Bias"

ϵ : Constante Arbitrária para Estabilidade Numérica

R: Tamanho de Raio Definido

μ : Média dos Sinais de Entrada

w_k : Pesos

x : Valor do Sinal de Entrada

x_k : Valores de Entrada

x_0 : Posição Horizontal do Centro

y : Valor de Saída

y_0 : Posição Vertical do Centro

σ : Desvio Padrão

1 INTRODUÇÃO

Os processos produtivos têm influência direta nas propriedades dos aços, se tornando um indicativo essencial para controle, medição, desempenho e qualidade. Como resultado do processo de solidificação do aço, por exemplo, é inevitável a formação de inclusões não metálicas, que são partículas tais como óxidos, silicatos e sulfetos. Deste modo, todos os aços apresentam inclusões não metálicas, podendo influenciar negativamente ou positivamente a aplicação do material, de acordo com as suas morfologias, composições, frações e distribuições [1]. Neste contexto, o estudo e controle de microinclusões é fundamental na produção de um aço, logo que as inclusões podem fragilizar o material a ponto de reduzir drasticamente suas propriedades mecânicas e a resistência a corrosão [2, 3, 4]. De forma contrária, em alguns casos, podem ter efeito benéfico, como, por exemplo, os aços resulfurados, aos quais intencionalmente são adicionadas inclusões de MnS, o que além de propiciar melhor rendimento metalúrgico, colabora na usinabilidade, permitindo aumentar a velocidade de corte, a vida da ferramenta, a segurança e produtividade reduzindo o custo de produção [5, 6, 7].

O estudo e conhecimento das inclusões é um tema relevante dentro da metalurgia e conhecer os fatores que podem contribuir para um melhor desempenho se torna vantajoso no ponto de vista financeiro, de qualidade e da otimização dos processos.

Embora uma variedade de processos de fabricação de aço de alta qualidade, incluindo agitação eletromagnética, borbulhamento e filtragem, sejam usados em operações de grande escala, esses processos são ineficazes na remoção de inclusões menores que 20 μm [4]. Até certo ponto pode-se perceber que as inclusões são uma característica indissociável dos aços e seu controle torna-se fundamental neste contexto. As primeiras formas de classificação foram propostas no trabalho de Sims e Dahle, que estabeleceram que a morfologia do MnS pode ser dividida em três categorias, Tipo 1: sulfetos globulares distribuídos aleatoriamente; Tipos 2 e 3: sulfetos de contorno de grão e sulfetos angulares, respectivamente. A partir desse trabalho muitos avanços foram feitos nas classificações das inclusões, culminando na norma ASTM E45 – 18ª edição [8], aprovada em junho de 2018 e publicada em setembro de 2018 pela ASTM

International. Esta norma introduz o método como um procedimento para determinar as inclusões não metálicas presentes nos aços.

Entretanto, o trabalho de classificação e estudo das inclusões é no mínimo muito trabalhoso ou extremamente oneroso e caro quando se consideram equipamentos de última geração como (ASPEX – microscópio automático de análise e Clemex – Software integrado de microscopia) que são capazes de classificar e contar as inclusões presentes no aço. Desconsiderando esses equipamentos que, atualmente não são a realidade de muitos produtores de aço, o trabalho persiste de forma custosa e lenta. Neste contexto, operadores experientes analisam inúmeros campos de imagens à procura das inclusões com maior nível de severidade, desconsiderando os artefatos como riscos, sujeiras e gotas, quando possível, sendo que muitas vezes esses artefatos podem confundir até o mais experiente dos operadores. De fato, pessoas menos experientes naturalmente encontram mais dificuldade de classificar ou desconsiderar esses artefatos que influenciam na análise final do operador, necessitando muitas vezes repreparar a amostra a ser classificada, gerando retrabalho e perdendo um recurso escasso: o tempo. Vale lembrar que a eficiência e a precisão da categorização de defeitos são críticas para o controle de qualidade do produto siderúrgico [9, 10].

Levando em consideração as oportunidades, dentre as demandas do setor metalúrgico, foi identificada a possibilidade de utilizar a tecnologia para otimizar o processo de classificação de microinclusões em aços e ligas. Desta forma, a utilização de redes neurais para a classificação das inclusões pode ser considerada uma alternativa, para avaliar os corpos de prova conforme a norma ASTM E45 [8], especificamente voltada para as inclusões de sulfeto de manganês, por se tratar de uma inclusão bem característica em aços.

Dentre todas as tecnologias utilizadas recentemente na metalurgia, o uso de redes neurais tem se destacado pelo potencial de correlação e prospecção que elas oferecem. As redes neurais são combinações de nós interligados que emulam comportamento semelhante aos neurônios do cérebro humano, e fazem parte do grupo de técnicas para aprendizado de máquina (“*Machine Learning*” – ML). Na atualidade, é comum encontrar tecnologias que utilizam as redes neurais como: carros autônomos, recomendações em plataformas *streamers*, sites de pesquisa,

classificação de imagens, entre outros. Recentemente, alguns estudos têm utilizado redes neurais para problemas na área metalúrgica, mostrando que existem oportunidades de aplicação a serem exploradas [11, 12, 13].

Espera-se que o emprego de metodologia de análise de inclusões em aços utilizando modelos de rede neural possa trazer ganhos em agilidade, assertividade e a capacidade de correlacionar os resultados com os demais processos produtivos, além de se tornar possível a análise e classificação por pessoas com menos experiência na área de metalografia.

O estudo e aplicação de redes neurais vem se disseminando por diversos segmentos de forma exponencial, e uma das justificativas é o aumento do “*turnover*” – indicador que mensura a quantidade de funcionários que deixam as empresas ou os seus cargos. No entanto, quanto maior o índice de turnover maior o número de pessoas inexperientes tomando decisões importantes dentro das empresas/negócios [14, 15, 11]. Conseqüentemente é muito vantajoso aplicar o uso de redes neurais e *machine learning* nesses cenários, pois de certa forma o conhecimento está sendo conservado como um patrimônio da empresa/negócio, no qual contribui o desempenho da empresa e do empregado, gerando rápido resultado, engajamento por parte do empregado e fácil aprendizado.

Considerando que as redes neurais são ferramentas exploradas e desenvolvidas de acordo com a demanda do mercado, os desafios são inúmeros. Inclusive no contexto de classificação de imagens por redes neurais convolucionais, os desafios começam na programação computacional e chegam até no tratamento das imagens utilizadas no banco de dados.

Vale lembrar que o emprego de redes neurais requer a realização de várias tarefas, incluindo a criação e preparação de bancos de dados de treinamento e controle, a seleção do projeto da rede neural, a otimização dos parâmetros operacionais e a verificação dos resultados adquiridos.

Os desafios mencionados valorizam ainda mais a importância do presente mestrado. Nas iniciativas de implementação de redes neurais há necessidade de criação de bancos de imagens cada vez mais robustos [16], sendo essenciais aspectos como a quantidade de imagens, forma como as imagens são geradas, padronizando os equipamentos de captura e todos os parâmetros de trabalho,

qualidade das imagens excluindo possíveis artefatos ou riscos, tratamento das imagens após a captura para padronização, divisão do número de imagens para cada treinamento específico ou para cada severidade (no caso de inclusões).

Os aspectos mencionados influenciam no resultado final da acurácia. Algumas questões desafiadoras podem ser colocadas: como é possível gerar um banco de dados de imagens ópticas com níveis de inclusões/severidades que se apliquem à realidade? Como criar um banco de dados com centenas de imagens diferentes, se um determinado aço é produzido esporadicamente de acordo com a demanda do mercado? Ademais, não basta ter inúmeras imagens para compor o banco de dados, estas precisam ter um padrão de qualidade para não influenciar na aprendizagem da rede neural, muitas vezes sendo necessário a aplicação de filtros para ajuste de cor e resolução.

1.1 OBJETIVOS

Os objetivos do presente trabalho são:

- Aplicar o modelo de redes neurais em imagens geradas por computador para verificar a possibilidade de reconhecimento de padrões desta rede.
- Desenvolver o modelo de redes neurais com base em um banco de imagens de inclusões de MnS de diferentes aços com diferentes graus de severidade segundo ASTM.
- Comparar o modelo de rede neural com a identificação de inclusões MnS feitas manualmente por um metalógrafo.

2 REVISÃO BIBLIOGRÁFICA

2.1 A importância e classificação dos aços

O aço é um material muito importante na vida moderna. Automóveis, aviões, navios, linhas de transmissão de energia elétrica, tubulações de água, redes integradas de telefonia, etc... são feitos de aço. Nas casas, o aço está presente em larga escala, dos talheres às panelas, passando pelos vergalhões que garantem a estabilidade das construções. Ainda, além da presença direta nos bens duráveis, o aço é vital na construção das máquinas e equipamentos que tornam possível a humanidade usufruir dos benefícios e facilidades conferidos pelos bens de consumo modernos. Desta maneira, fica inconcebível qualquer tentativa de imaginar o mundo moderno sem a presença do aço. Pode-se dizer que o poderio econômico de uma nação está direta e intimamente ligado com o consumo per capita de aço [17].

O volume de aço produzido mundialmente tem aumentado relevantemente, o que traz investimentos para o segmento. Segundo o site *World steel Association* o volume de aço produzido mundialmente em 2019 foi de 1.828,4 milhões de toneladas, e em 2021 foi de 1.951,9 milhões de toneladas [18]. O Brasil aparece na nona colocação de maior produtor de aço mundial com um volume de 36,1 milhões de toneladas, onde na Figura 1 é possível observar a relação volume x país, ficando nítida a força de produção da China e da Índia.

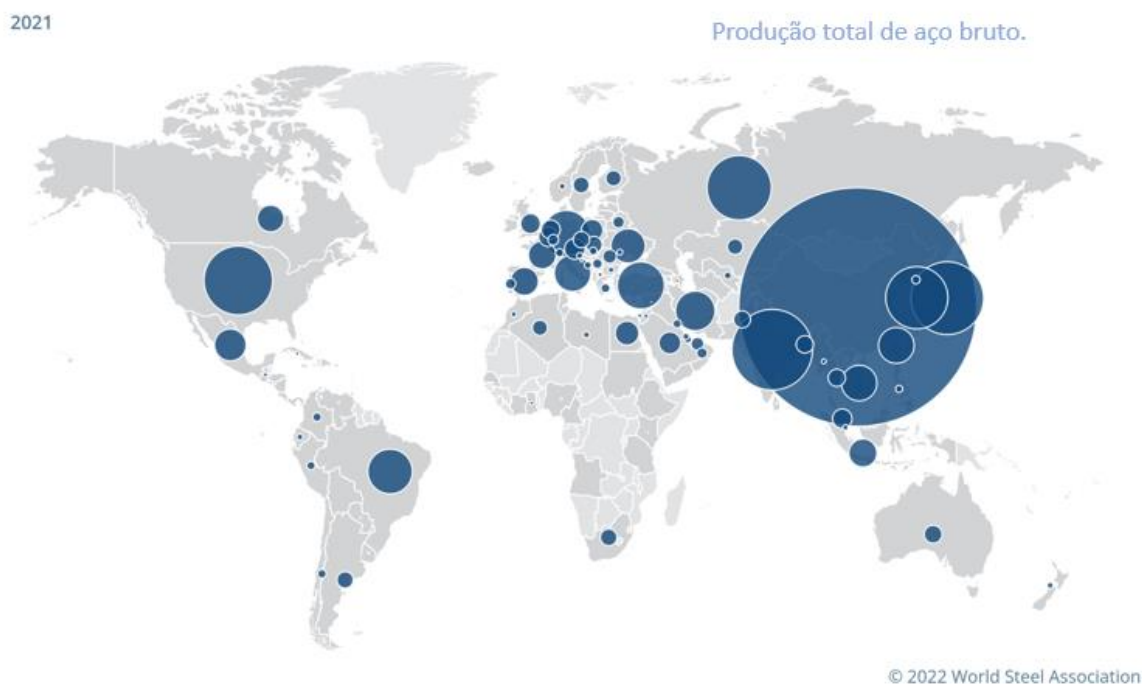


Figura 1 - Relação volume de produção de aço x país. Fonte: *World Steel Association*. [19]

A denominação aço é aplicada a uma grande variedade de ligas em que o principal componente é o elemento ferro (símbolo: Fe; número atômico: 26; peso atômico: 55,847; raio atômico: 126pm). O carbono (símbolo: C; número atômico: 6; peso atômico: 12,011; raio atômico: 70pm) é considerado o mais importante dos outros constituintes do aço, e seu teor varia de aproximadamente, 0,03 a 1,5% em peso. [20]

De uma maneira esquemática, a Figura 2 demonstra a classificação de metais ferrosos incluindo dois tipos comerciais de aço: os aços baixa liga (*low-alloy steels*) e os aços alta liga (*high-alloy steels*) [21]. Além da classificação entre baixa liga e alta liga os aços são sub-classificados em baixo carbono, médio carbono e alto carbono, os quais segundo o “*American Iron and Steel Institute*” são definidos como os aços que não contêm ou não atendem os requisitos mínimos para determinados elementos de liga como: cromo, cobalto, nióbio, molibdênio, níquel, titânio, tungstênio, vanádio ou zircônio; ou qualquer outro elemento adicionado para obter a liga desejada. Os aços carbono são classificados de acordo com o teor de carbono presente em sua composição química. Entre essas classificações, destacam-se: aço baixo carbono, aço de

médio carbono, aço de alto carbono e aço de Ultra alto carbono (menos usual), como citado na Tabela 1 abaixo.

Tabela 1 - Tabela de classificação de aços e principais aplicações. Fonte: Villares Metals

Faixa Teor Carbono	Faixa Teor manganês	Classificação do aço	Principais Aplicações
<0,3%	-----	Aço Baixo Carbono	Chapas automotivas, perfis estruturais, construção civil e latas de folhas de flandres.
0,3% - 0,6%	0,60% – 1,65%	Aço Médio Carbono	Rodas, engrenagens, virabrequins e equipamentos ferroviários.
0,60% – 1,00%	0,30% - 0,90%	Aço Alto Carbono	Talhadeiras, folhas de serrote, martelos e facas.
1,25% - 2,0%	-----	Aço Ultra Alto Carbono	Facas, eixos e punções.

Os aços baixa liga e alta resistência são aços micro ligados, ou seja, apresentam elementos de liga em sua composição. Esses aços são desenvolvidos para fornecer melhores propriedades mecânicas. Os aços baixa liga contêm baixo teor de carbono na faixa de 0,05% a 0,25%, e manganês com o limite máximo de 2%, formando uma classe de materiais que apresentam propriedades melhores do que os aços carbono, devido a adição de elementos como níquel, cromo e molibdênio.

Finalmente, os aços alta liga (dos quais o tipo mais importante é talvez o aço inoxidável – stainless steel) contêm mais de 5% em teores de elementos como: o cromo, níquel, o manganês, molibdênio ou o tungstênio [20].

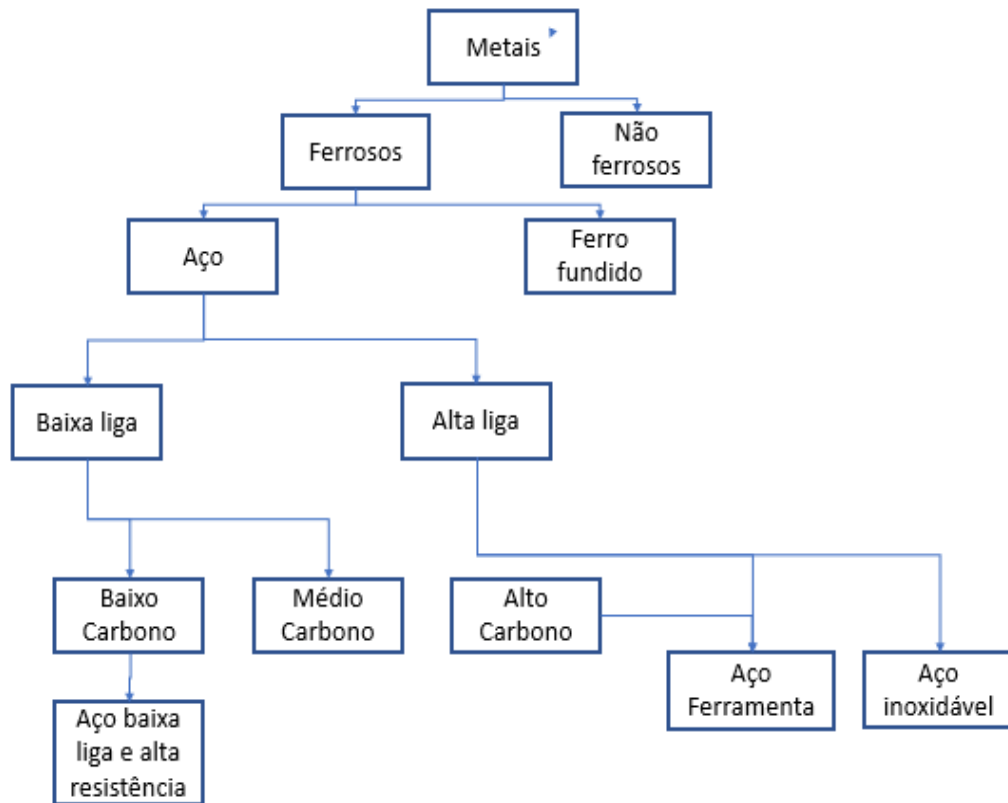


Figura 2 - Classificação geral de metais ferrosos.

Além dessa classificação esquemática e sucinta, existe a classificação normativa de órgãos normativos como: SAE, AISI e etc.; que tem como objetivo padronizar a nomenclatura do aço de acordo com a sua composição química, conforme mostra a Tabela 2.

Tabela 2 - Classificação dos aços de acordo com as normas SAE.[22]

Designação	Tipos de aço
SAE	
10xx	Aços-carbono comuns
11xx	Aços de usinagem fácil, com alto Enxofre
12xx	Aços de usinagem fácil, com alto Fósforo e Enxofre
15xx	Aços-Manganês com manganês acima de 1%
13xx	Aços-Manganês com 1,75% de manganês médio
40xx	Aços-Manganês com 0,25% de molibdênio médio
41xx	Aços-Cromo-Molibdênio com 0,4 a 1,1% de Cromo e 0,08 a 0,35% de Molibdênio
43xx	Aços-Níquel-Cromo-Molibdênio com 1,65 a 2% de Níquel, 0,4 a 0,9% de Cromo e 0,2 a 0,3% de Molibdênio
46xx	Aços-Níquel-Molibdênio com 0,7 a 2% de Níquel e 0,15 a 0,3% de Molibdênio
47xx	Aços-Níquel-Cromo-Molibdênio com 1,05% de Níquel, 0,45% de Cromo e 0,2% de Molibdênio
48xx	Aços-Níquel-Molibdênio com 3,25 a 3,75% de Níquel e 0,2 a 0,3% de Molibdênio.
51xx	Aços-Cromo com 0,7 a 1,1% de Cromo
E51100	Aços-Cromo (forno elétrico) com 1% de Cromo
E52100	Aços-Cromo (forno elétrico) com 1,45% de Cromo
61xx	Aços-Cromo-Vanádio com 0,6 ou 0,95% de Cromo e 0,1 ou 0,15% de Vanádio no mínimo.
86xx	Aços-Níquel-Cromo-Molibdênio com 0,55 de Níquel, 0,5% de Cromo e 0,2% de Molibdênio.
87xx	Aços-Níquel-Cromo-Molibdênio com 0,55 de Níquel, 0,5% de Cromo e 0,25% de Molibdênio.
88xx	Aços-Níquel-Cromo-Molibdênio com 0,55 de Níquel, 0,5% de Cromo e 0,3% a 0,4% de Molibdênio.
9260	Aços-Sílicio com 1,8% a 2,2% de Sílicio.
50Bxx	Aços-Cromo com 0,2 a 0,6% de Cromo e 0,0005 a 0,003% de boro.
51B60	Aços-Cromo com 0,8% de Cromo e 0,0005 a 0,003 de boro
81B45	Aços-Níquel-Cromo-Molibdênio com 0,3% de níquel, 0,45% de Cromo, 0,12% Molibdênio e 0,0005 a 0,003% de boro
94Bxx	Aços-Níquel-Cromo-Molibdênio com 0,45% de níquel, 0,40% de Cromo, 0,12% Molibdênio e 0,0005 a 0,003% de boro

2.2 Impacto das inclusões no processamento e nas propriedades dos aços.

As inclusões não metálicas são estudadas há muitos anos, e suas formações e influências fazem parte da maioria dos processos produtivos de aços e ligas especiais, inclusive as classes citadas na seção anterior. A morfologia e a distribuição dessas partículas influenciam diretamente as propriedades mecânicas, processamento, a resistência a corrosão e a usinabilidade dos materiais.

Como descrito na literatura [23] a presença de inclusões afeta a tenacidade e a ductilidade dos aços. Foi feita uma comparação entre duas condições diferentes de processamento do aço 13-8, uma condição refundida utilizando a rota VIM (“*Vacuum induction melting*”) + VAR (“*Vacuum arc remelting*”) e a segunda condição utilizando a dupla refusão através da rota VIM + ESR (Refusão sob Escória Eletrocondutora) + VAR. Como resultado dessa comparação, a condição que foi duplamente refundida em VIM + ESR + VAR apresentou um nível de enxofre bem inferior ao outro processo, conseqüentemente apresentando melhores resultados de redução de área e alongamento no ensaio de tração. Isso porque a fração de inclusões não metálicas diminuiu em relação a amostra da primeira rota analisada.

A literatura [23] também cita que as inclusões além da influência na tenacidade e na ductilidade dos aços, apresentam os efeitos nocivos nas propriedades dos materiais. Portanto, estes efeitos vão muito além da redução da tenacidade e da ductilidade,

Isso fica evidente no texto da literatura [2], onde são comparadas as propriedades de dois aços com teores de impurezas distintos. O referido estudo classifica esses aços como Aço A com 91 ppm de nível de impureza (IL – *impurity level*) e aço B com 55 ppm de IL. As impurezas se resumem à soma de enxofre + oxigênio + nitrogênio.

Os aços utilizados e citados na literatura [2] para a realização desse estudo apresentam a mesma composição química e foram expostos ao mesmo processo de fabricação. O processo refere-se à laminação a quente partindo do diâmetro de 200 mm até o diâmetro de 12 mm com uma temperatura de laminação de 930°C e têmpera direta após a laminação.

Na Tabela 3 abaixo é possível observar os resultados obtidos por Tervo [2] após a realização dos ensaios de tração e impacto nos sentidos transversal e longitudinal das amostras obtidas.

Tabela 3 - Resultados dos ensaios mecânicos obtidos por Tervo [2].

Aço	Direção do teste	Rm [MPa]	Rp0.2 [MPa]	RA [%]	A [%]	CVN a - 40°C [J]
A	Longitudinal	1324	1032	63,8	13,8	29
	Transversal	1348	1050	54,1	11,6	24
B	Longitudinal	1281	1000	65,3	14,0	40
	Transversal	1290	1005	63,9	12,8	39

Após obtenção dos resultados, a conclusão da literatura [2] é que o nível de impureza afetou diretamente as propriedades de tenacidade do material. Isso fica evidente quando são comparados os resultados de RA “redução de área”, A “alongamento” e CVN “Charpy V” dos aços A e B. O aço A apresenta maior resistência e conseqüentemente menor tenacidade, já o aço B apresenta um valor menor de resistência e maior tenacidade, com valores apropriados de CVN e resistência. Nitidamente foi constatado que o aço com maior IL também apresenta maior anisotropia. [23, 3]

Segundo a literatura [24] as inclusões também influenciam a superfície final do produto, as transformações de fases e a resistência à corrosão dos aços.

As inclusões não metálicas em aços para moldes plásticos são deletérias para o acabamento superficial do componente produzido, o que impactaria diretamente na peça injetada. Além disso, a literatura [24] explica a influência das inclusões na transformação de fases em alguns aços baixo carbono, onde as inclusões desempenham um papel crítico na nucleação de ferrita acicular fina em juntas soldadas, sendo essa fase fundamental para o alcance de propriedades mecânicas satisfatórias.

São apresentadas as duas principais falhas com influência direta das inclusões em aços, quais sejam: 1. trincas por hidrogênio: as quais ocorrem devido à presença de regiões microsegregadas e conseqüentemente com presença de sulfetos de manganês, o que aumenta a solubilidade temporária do hidrogênio naquela região, resultando no aprisionamento do H e formação de trincas; 2. Nucleação de “pittings”: ocorre em torno das inclusões, ocasionando a queda da

resistência a corrosão. Isso ocorre devido à presença de imperfeições, e inclusões na camada passiva do aço, sendo essa uma fina película aderida na superfície do material que possui espessura da ordem de nanômetros e tem como objetivo preservar a resistência a corrosão. A presença de inclusões permite ainda a penetração de íons agressivos que irão interagir com os íons metálicos do material, resultando no dano por “pitting”. [24]

Essa influência fica nítida no trabalho [3] quando é citado que a corrosão por pitting nas superfícies dos metais são iniciadas via um ponto fissurado do filme de óxido passivo da superfície do metal, e alega-se que certas inclusões parecem ser mais susceptíveis à corrosão do que outras. E essa distinção está associada com a presença de um enxofre fino disperso que contém precipitados em sua superfície, que agem como uma fonte de íons de sulfeto e por sua vez, catalisa reações anódicas e catódicas. E dessa forma o trabalho desenvolvido por Rieders [3] cita que as inclusões ricas em sulfeto são mais favoráveis a ação corrosiva por *pitting* do que as inclusões ricas em Ferro. [3]

As inclusões de MnS se configuram nas seguintes morfologias: poligonal, circular e alongadas. E preferencialmente, a corrosão por *pitting* se intensifica nas regiões com maior concentração de enxofre e conseqüentemente as regiões com maior número de inclusões alongadas. [4]

E como citado na literatura [3] em alguns trabalhos da referência [25, 3], as inclusões de sulfeto de manganês (MnS) presentes nos aços como é mostrado na Figura 3 apresentam mecanismos diferentes em relação ao seu comportamento de corrosão por *pitting* em aço carbono e aço inoxidável. Nos casos de aço inoxidável as inclusões de sulfeto são a causa para a reação anódica, enquanto no caso do aço carbono, as inclusões são as responsáveis pela reação catódica. Em outra alegação é demonstrado que o tamanho e a forma das inclusões afetam a cinética de dissolução. [3]

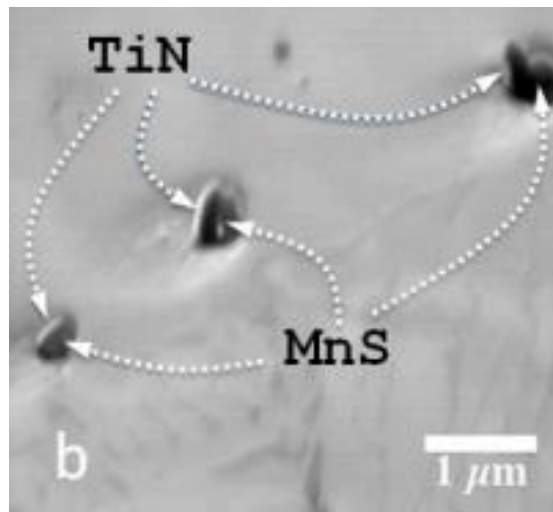


Figura 3 - MnS e TiN em aço duplex[3].

As inclusões quando não controladas podem ser extremamente nocivas para a vida útil do aço. Na referência [26] foi avaliado a relação entre as falhas por fadiga de contato e as inclusões não metálicas presentes no aço. E o trabalho deixa claro que é fundamental entender os principais fatores que regem a vida em fadiga de contato, e para isso foram utilizadas diversas amostras de aços com diferentes composições químicas, como por exemplo o aço SAE 52100, variando a sua composição e conseqüentemente as características das inclusões, sendo assim possível avaliar a vida útil. Este estudo concluiu que a vida em fadiga de contato aumentava com a redução do tamanho dos óxidos, variando de acordo com a composição dos óxidos encontrados. Dessa forma, o controle e estudo de inclusões não metálicas é bastante relevante no segmento metalúrgico. [26]

2.3 Aços resulfurados

Os aços resulfurados ou aços de usinagem fácil tem como característica o alto teor de manganês e enxofre intensificando a formação da inclusão não metálica sulfeto de manganês (MnS), o que resulta em sua principal propriedade: excelente usinabilidade. Essa propriedade faz com que os aços resulfurados sejam usados em diversas aplicações na indústria. A formação do MnS previne a formação de FeS, controla o crescimento de grão e devido a fácil usinabilidade, resulta no aumento da vida útil das ferramentas de cortes, algo extremamente interessante para empresas do setor de usinagem. O FeS é indesejável em aços uma vez que

forma fase líquida nas operações de conformação à quente. As inclusões MnS provocam a redução na ductilidade e na tenacidade do material, podendo ser consideradas concentradores de tensão, nucleando trincas durante a deformação.

[5]

O mercado do setor de aços e ligas especiais tem procurado cada vez mais os aços resulfurados para que sejam atendidos níveis de produtividade e de qualidade de forma geral, onde o segmento tem desenvolvido processos afins da redução de custo, aumento da vida útil de ferramentas e a redução do tempo gasto com as manutenções preventivas e corretivas das máquinas utilizadas e consequentemente o aumento do indicador MTBF (*“Mean time between failures”*). Este indicador, citado na literatura [27], tem como objetivo medir o tempo médio entre as falhas dos equipamentos e a redução do MTTR (*“Mean time to repair”*) que é o tempo médio gasto para realizar o reparo.

A usinabilidade dos aços resulfurados está diretamente relacionada com a adição de enxofre para a formação dos sulfetos de manganês (MnS), pois estas inclusões facilitam a quebra do cavaco em operações de usinagem. Porém, as inclusões também estão relacionadas a algumas propriedades do material, e os fatores que influenciam essas propriedades são a morfologia, tamanho e a distribuição dos sulfetos de manganês. Portanto se torna essencial o estudo e a classificação das inclusões dos aços resulfurados. [5]. Abaixo, na Figura 4, é mostrada uma micrografia do aço AISI 430F preparada para análise de microinclusões.



Figura 4 - Amostra do aço AISI 430F preparada para análise de inclusão. As linhas cinza escuras são os sulfetos de manganês (MnS). Fonte: Villares Metals.

2.4 Avaliação tradicional de inclusões em aços

Todos os ensaios metalográficos são baseados e desenvolvidos com referências oficiais, a fim de padronizar e facilitar o entendimento dos resultados.

As principais normas utilizadas para a classificação de micro inclusões são: ASTM E45 e DIN50602 [8, 28].

Segundo a norma ASTM E45 [8] existem 5 sistemas (A-E) aceitáveis de análise de inclusões através de imagens obtidas por microscopia óptica. E, contudo, as análises são atribuídas as categorias baseadas na morfologia e não necessariamente em sua composição química.

Alguns aços ou ligas possibilitam a utilização de mais de um método para a classificação ou também é possível utilizar a combinação de métodos.

De acordo com a norma [8], o método microscópico é usado para caracterizar o tamanho, distribuição, número e o tipo de inclusão na superfície da amostra polida, com o objetivo de padronizar “*charts*”, que descrevem uma série de configurações de inclusões típicas que foram criadas para direcionar a comparação com os campos vistos no microscópio.

As vantagens e desvantagens do método microscópico são: inclusões podem ser caracterizadas de acordo com o seu tamanho, tipo e número. Além de possibilitar que inclusões extremamente pequenas possam ser relevadas na classificação. A desvantagem do método microscópico se resume ao campo de classificação individual de 0,5mm². O campo é considerado muito pequeno, o que

limita o tamanho do corpo de prova, pois uma amostra muito grande em relação a este campo, levaria muito tempo para ser analisada completamente.

No método microscópico, as inclusões são classificadas dentro de 4 categorias (chamada de tipo), baseadas em sua morfologia e 2 subcategorias baseadas em sua largura ou diâmetro. As categorias são: A – sulfeto, B – alumina, C – silicato e D – óxido globular, e o que as define como fina e grossa é a espessura/diâmetro da inclusão.

Após a classificação do tipo, as inclusões devem ser classificadas como fina ou grossa, através da Tabela 4 a seguir.

Tabela 4 - Parâmetros de largura e diâmetro para realizar a classificação de inclusões em aços. Fonte: Norma ASTM E45.[8]

Tipos de inclusões	Séries Finas		Séries Grossas	
	Largura mínima μm (in.)	Largura Máxima μm (in.)	Largura mínima μm (in.)	Largura Máxima μm (in.)
A	2 (.00008)	(.00016)	>4 (.00016)	12 (.0005)
B	2 (.00008)	(.00035)	>9 (.00035)	15 (.0006)
C	2 (.00008)	(.0002)	>5 (.0002)	12 (.0005)
D	2 (.00008)	(.0003)	>8 (.0003)	13 (.0005)
Qualquer inclusão com dimensões maiores que o máximo para a série analisada deve ser relatada como 'oversized' e adicionar as suas dimensões.				

2.5 Classificação manual, redes neurais e métodos automatizados de classificação

2.5.1 Classificação manual

O procedimento manual de classificação das inclusões não metálicas acontece através da microscopia óptica. Segundo Mansur [29] a microscopia óptica é uma técnica de exames com ampliações das dimensões físicas que vão de 10 a cerca de 1500 vezes. Nesta técnica, o contraste da imagem é resultado da diferença de refletividade da luz nas diversas regiões da amostra. Sendo utilizada a técnica de microscopia óptica, o próximo passo é a análise de inclusões através

do auxílio de “charts” padrão já descritos na norma ASTM E45, que por meio de desenhos representativos de microtexturas, indicam os tipos, formas, tamanhos, quantidades e distribuição das mesmas. [29].

O procedimento manual é descrito pela ASTM E45 da seguinte forma: Existem 2 técnicas que podem ser empregadas para atingir um campo de análise $0,5 \text{ mm}^2$. A primeira técnica é projetar a imagem do microscópio com um aumento de 100x, dentro de uma tela de visualização que contém um quadrado de dimensões de 71 mm de lado. A segunda técnica é utilizar o reticulado próprio para o microscópio, no qual é possível sobrepor o quadrado dentro do campo de visão [8].

A análise começa no canto da amostra onde se inicia a varredura conforme mostra a Figura 5 e a comparação com os “charts” de acordo com o tipo e a espessura de cada inclusão. A área mínima requerida para a análise é de 160 mm^2 .

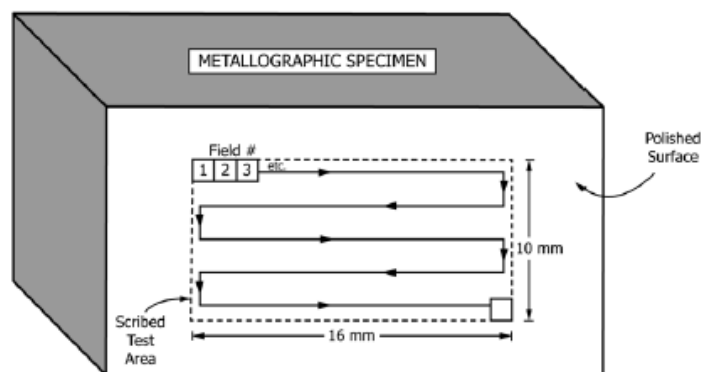


Figura 5 - Método de varredura manual no corpo de prova para classificação de inclusões.

Fonte: Norma ASTM E45. [8]

Após toda varredura os campos são classificados e descritos conforme o método escolhido, conforme previsto na norma ASTM E45 de A-E. Maiores detalhamentos dos métodos serão descritos no capítulo de Materiais e Métodos.

2.5.2 Métodos automatizados de classificação

Segundo a literatura [30], a espectrometria de emissão óptica é muito utilizada pelas indústrias para realizar análise de amostras sólidas com o objetivo de controlar a composição química de seus materiais. A técnica se baseia em uma centelha que é direcionada a amostra resultando na excitação de elétrons dos

elementos presentes na mesma, e seguida pela emissão de ondas radioativas que são convertidas em concentração química. Uma moderna evolução na área analítica foi o desenvolvimento do sistema *Spark-DAT*, que tende a ser uma ferramenta complementar a espectrometria de emissão óptica, no qual permite que ela conceda a composição química da amostra e a fração volumétrica de inclusões não metálicas. [30]

A primeira etapa do trabalho [30] consistiu na contagem e caracterização das inclusões encontradas nas amostras com o auxílio das técnicas de microscopia. Com base nas informações obtidas nestes equipamentos, foi desenvolvido um método de análise no espectrômetro de emissão óptica acoplado com o *Spark-DAT* para permitir a detecção das mesmas inclusões. Uma vez identificados os componentes das inclusões avaliadas pela microscopia, foi desenvolvido um programa de análise no software do espectrômetro óptico para possibilitar a avaliação dos mesmos com o *Spark-DAT*. O programa utilizou os algoritmos desenvolvidos pelo fabricante para coletar os sinais de cada elemento e gerar informações sobre as inclusões encontradas.[30]

O método através do *Spark-DAT* tem algumas vantagens, quando comparado com as metodologias de microscopia. Em primeiro lugar, a possibilidade de fornecer a análise química e informações sobre o conteúdo de inclusões com um único instrumento, reduzindo investimento e custos operacionais. Isso também simplifica e encurta o processo analítico, tornando-o adequado para o controle de produção.[30]

O processo de análise através de redes neurais é uma automatização da microscopia óptica através do estudo de imagens, ou seja, a análise por redes neurais não elimina a preparação metalográfica e todo o trabalho manual nela envolvido, mas otimiza e fornece confiabilidade na hora da tomada de decisão.

As redes neurais oferecem diversas vantagens na análise e classificação de imagens sobre alguns outros métodos. E como citado na literatura [31] a maior vantagem da rede neural é a capacidade de ajustar modelos não lineares complexos, além de permitir uma visão diferente de um conjunto de dados ou processos interligados, possibilitando um melhor entendimento e correlação de todos os processos produtivos até as características do produto final. [31]

Comparando os métodos, observa-se que o método de classificação através do *Spark-DAT* é um método que tem como foco a análise química e consegue realizar uma análise muito mais representativa da inclusão pois leva em consideração o volume e área da amostra, não somente o plano como os outros métodos, de acordo com representação na Figura 6. Outra característica desse método é o alto investimento nos equipamentos capacitados a realizar esse estudo. Por consequência desse alto investimento nem todas as empresas têm recursos suficientes para tal aquisição. Por outro lado, o método de classificação através de redes neurais não confrontaria o método *Spark-DAT* pois as aplicações têm diferentes objetivos, ou seja, o melhor cenário seria aplicar os dois métodos de classificação, pois teria a assertividade da espectrometria óptica e toda otimização do processo de classificação por microscopia óptica através das redes neurais.[29]

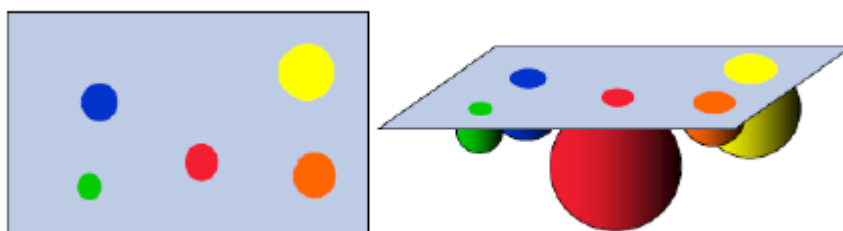


Figura 6 - Imagem representativa de diferentes planos das microinclusões em uma amostra representativa.[30]

2.5.3 Redes Neurais

Uma rede neural é definida como um processador maciçamente paralelamente distribuído, constituído de unidades de processamento simples, sendo estas unidades denominadas neurônios [32]. Um neurônio é modelado conforme Figura 7, sendo que primeiro recebe valores de entrada x_k , que são multiplicados cada um com um peso w_k , e em seguida somados os valores resultantes em uma somatória (Eq. 1). A esta somatória é adicionado um valor b (“Bias”), e o resultado é o argumento de uma função f . [32]

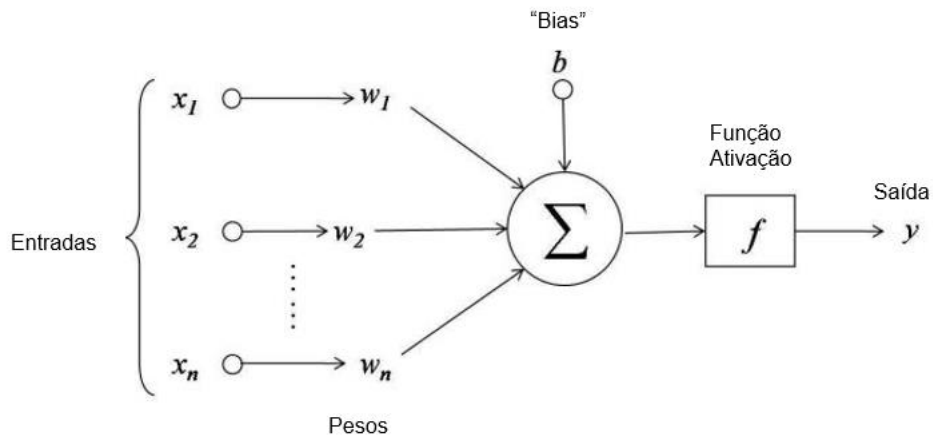


Figura 7 - Representação esquemática de um neurônio com seus parâmetros de entrada x , k , pesos w , k , "bias" b , função de ativação f e saída y . [32]

$$y = f\left(\sum_k w_k x_k + b\right) \quad \text{Eq. 2.1}$$

Há inúmeras funções que podem ser definidas para a f . Uma das funções mais comuns é a função Sigmoidal representada pela Figura 8, cuja forma se assemelha a um sinal. O valor de saída y é o valor de entrada para outro neurônio e assim sucessivamente. Numa rede neural, são definidas diferentes camadas de neurônios interconectadas (Figura 9), sendo sempre uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída.

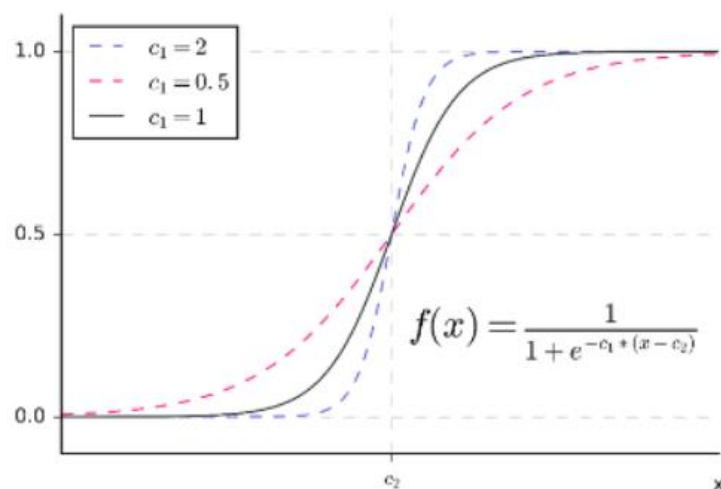


Figura 8 - Gráfico da função sigmoideal, onde c_1 e c_2 são constantes. [33]

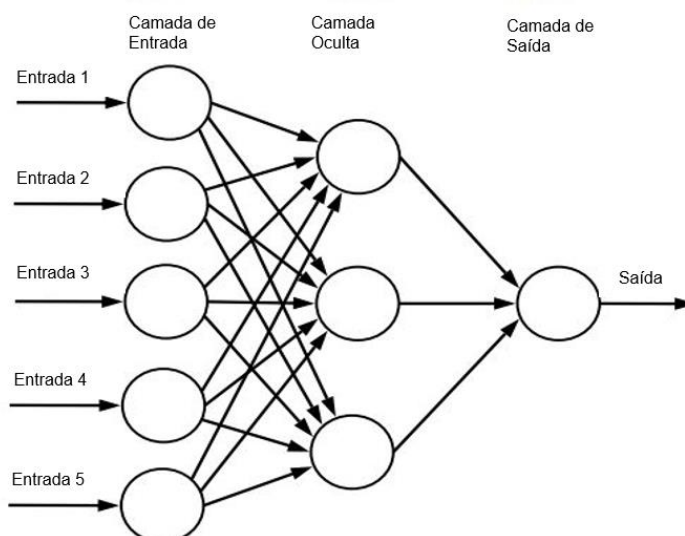


Figura 9 - Diagrama de grafos de uma rede neural com sua camada de entrada, camada oculta, camada de saída e suas sinapses indicadas por setas.[34]

Cada camada oculta apresenta uma estrutura pré-definida de suas sinapses (conexão entre neurônios, onde são definidos os pesos). As estruturas que são utilizadas neste trabalho são: Densa, Convolutiva, “*Max Pooling 2D*”, “*Dropout*”, “*Batch Normalization*”; e estas serão apresentadas nos próximos itens.

A arquitetura das redes neurais dentro de um modelo de rede configura-se de acordo com a forma ou a topologia no qual os neurônios estão estruturados. As arquiteturas de redes neurais artificiais são: memória de curto-prazo, auto - enconder, auto - enconder variacional, perceptron de multicamadas e convolucionais. [35]

A primeira estrutura de rede neural a ser apresentada é a densa, cujas sinapses entre os neurônios e os neurônios da camada anterior e a camada posterior são completamente preenchidas entre si. Como exemplo, se a camada anterior apresenta 5 neurônios, e a camada oculta apresenta 4 neurônios, haverá 20 sinapses entre estas duas camadas conectando cada um dos neurônios. Este modelo de rede neural é o próprio apresentado na Figura 9.

Uma rede convolutiva é um *perceptron* de múltiplas camadas projetado especificamente para reconhecer formas bidimensionais com um alto grau de invariância quanto a translação, escalamento, inclinação e outras formas de distorção[32]. Uma camada convolucional secciona a sua entrada em pequenos

“grids”, e gera uma resposta para cada espaço no grid. Os parâmetros que definem uma camada convolucional são: filtros (“*filters*”), passo (“*strider*”), núcleo (“*kernel*”), preenchimento (“*padding*”) e sua função (pesos e “*bias*”). [32]

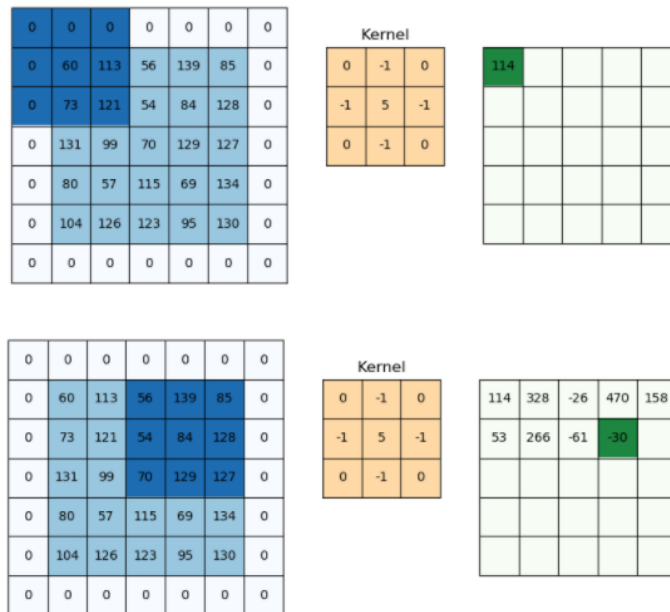


Figura 10 - Ilustração do método de funcionamento da camada convolucional, onde o núcleo é 3x3 e varre a imagem gerando um valor para cada passo.[36]

A camada “*max pooling*” subdivide o domínio de entrada em matrizes 2x2 (este parâmetro pode ser modificado) e seleciona o valor máximo deste grid. A operação pode ser vista na figura abaixo (Figura 11):

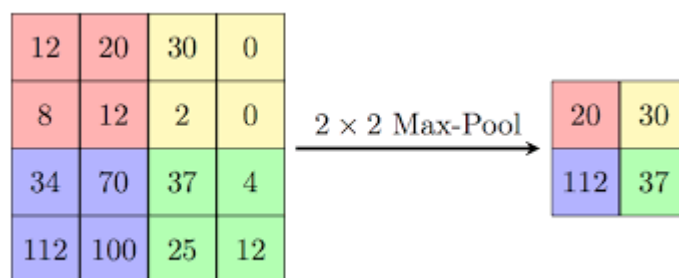


Figura 11 - Ilustração da camada “*max pooling*” onde a camada de entrada é separada em um grid 2x2 e os maiores valores dentro da região são selecionados.[37]

O “*Dropout*” é aplicado à camada anterior de neurônios, e esta função elimina alguns neurônios da camada anterior (Figura 12). Esta eliminação é feita aleatoriamente e com uma taxa definida.

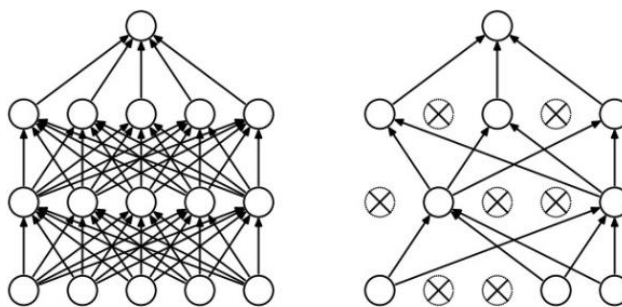


Figura 12 - Esquema da operação "Dropout", sendo à esquerda uma rede neural com camadas densas e a direita após aplicar a operação.[38]

A camada denominada "*batch normalization*" é definida por uma operação de normalização de seus sinais de entrada. Nesta operação de normalização, é calculada a média de seus sinais de entrada e desvio padrão, e posteriormente, é utilizada a expressão:

$$x_i^* = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \quad \text{Eq. 1.2}$$

onde x é o valor do sinal de entrada, x^* é o valor normalizado, μ é a média dos sinais de entrada, σ o desvio-padrão e ϵ uma constante arbitrária para estabilidade numérica. Desta forma, x^* estará contido em um domínio mais próximo da região de maior derivada da função sinal sigmoidal, apresentando assim maior capacidade de aprendizagem.

As imagens utilizadas no banco de dados de treinamento, validação e teste do presente trabalho de Mestrado foram classificadas através da "*Region-based convolutional neural network*" ou R-CNN para realizar a segmentação das imagens com o propósito de selecionar regiões da imagem e checar se essas regiões contém o objeto solicitado, nesse caso as inclusões MnS.

Segundo o trabalho [39], antes da análise, estruturas de interesse devem ser detectadas e definidas de acordo com uma representação adequada para quantificação pelo computador. Isto é alcançado através da segmentação, o processo de fracionamento de uma imagem em várias regiões ou segmentos homogêneos. A segmentação constitui uma transição importante no fluxo de análise de imagem, substituindo os valores de intensidade por rótulos de região. A

maioria dos métodos tradicionais de segmentação são baseados na intensidade e relações espaciais de pixels. [39]

Para segmentar os dados da imagem de entrada, o computador transforma o problema de segmentação em um problema de classificação de pixels em que cada pixel pode ser como pertencente a um segmento específico ou classe. Um conjunto de pixels de entrada que foi rotulado é representado no espaço de recursos e, em seguida, usado como o conjunto de treinamento. Uma vez que o classificador é treinado, ele pode ser usado para classificar o resto dos pixels de entrada ou dados de imagem completamente novos. [39]

Um objeto é formado por 4 diferentes tipos de características como escala variável, cores, textura e bordas de partículas. R-CNN usa a procura seletiva para identificar os padrões com base nas diferentes regiões do objeto que são solicitadas. A Figura 13 mostra a arquitetura típica R-CNN. O primeiro passo da arquitetura seleciona um conjunto de regiões propostas e extrai características de cada região usando CNN como exemplificado na Figura. E finalmente cada característica extraída será classificada. [40]

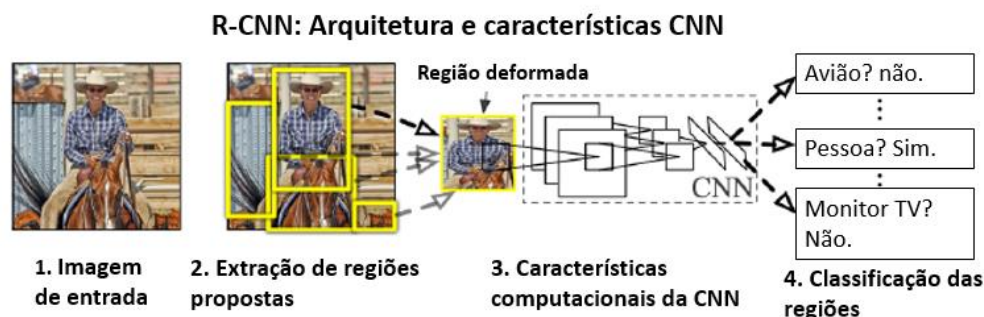


Figura 13 - R-CNN: Exemplo de arquitetura e características CNN em redes neurais.[40]

A classificação e análise da imagem é feita em cada pixel, ou seja, o valor de cada pixel, após o escalonamento, corresponde a uma entrada da rede neural. Após a etapa de treinamento, a imagem é processada por inteiro, fazendo-se uma varredura, gerando como saída uma imagem temática na qual cada tema representa um dos rótulos existentes na imagem original. A classificação é uma etapa posterior, onde a cada rótulo associa-se uma classe ou nível. [13]

2.6 Redes Neurais aplicado em estudos de Engenharia de Materiais

As redes neurais estão sendo aplicadas em diversos estudos, inclusive no segmento metalúrgico, onde temos diversas oportunidades para a aplicação do *machine learning*. Diversas ferramentas e modelos de *machine learning* tem sido estudado, como no trabalho [15], que utilizou métodos como *Logistic Regression*, *K-Nearest Neighbors*, *Support Vector Classifier (linear and RBF kernels)*, *Random Forests*, *AdaBoost*, *Gradient Boosting* e *Artificial Neural Networks*, com o objetivo de encontrar a melhor e mais precisa forma de avaliar o nível de inclusão em fios de aço para reforços em pneus. O treinamento, validação e teste foram conduzidos usando resultados experimentais obtidos através de um programa de controle de qualidade da fábrica que consistiu em 855 peças fundidas onde foram monitorados 140 parâmetros diferentes e 7,25% dessas peças foram reprovadas por consequência do número e das propriedades de inclusões não metálicas presentes. Dentre todas as ferramentas e modelos utilizados o mais preciso foi o Random Forests que apresentou 92% de acurácia contra 88% de acurácia da rede neural.[15]

Atualmente vários pesquisadores têm publicado estudos com redes neurais conforme demonstrado na literatura [11], no qual os autores analisaram a previsão da rugosidade da superfície de torneamento CNC de um aço inoxidável austenítico AISI 304, levando em consideração parâmetros de máquina como: velocidade de corte, avanço da ferramenta e profundidade do corte, alguns testes foram realizados tendo como dados de entrada os parâmetros de máquina e dados de saída a rugosidade superficial do material. No qual alcançaram um valor de regressão $R = 0,91$ (onde 1 seria a correlação perfeita), valor extremamente relevante [11].

Os trabalhos anteriores utilizaram ferramentas de machine learning e modelos de redes neurais, acionado como dados de entrada parâmetros de processo para realizar a predição através dos dados de saída. Outra aplicação que atualmente está sendo muito estudada é a classificação de imagens mediante das redes neurais. No estudo da literatura [12] foi explorado a segmentação e classificação de defeitos em imagens STEM de aço utilizando “*deep learning*”, também é citado que os defeitos foram segmentados em 3 classificações; linhas de

discordâncias, precipitados e vazios. O modelo de rede neural utilizado analisou as imagens com aumento de 250000x, e o maior desafio do trabalho foi realizar a segmentação em escala nanométrica em micrografias eletrônicas, o que requer um mecanismo complexo de contraste. Porém independentemente dos desafios encontrados no trabalho, os resultados do treinamento mostraram alta precisão em cada pixel em todas as 3 classificações de defeitos: 91,60 \pm 1,77% em discordâncias, 93,39 \pm 1,00% em precipitados e 98,85 \pm 0,56% em vazios, ou seja, um resultado relevante. Este mesmo trabalho também mostra a comparação dos resultados obtidos através da rede neural com os resultados obtidos pela segmentação de pessoas experientes no assunto, e concluiu que a aplicação da rede neural é superior na acurácia dos resultados e além disso no tempo gasto para realizar a classificação.

De forma geral as redes neurais têm sido aplicadas em vários segmentos, inclusive como apoio estratégico na tomada de decisão de grandes empresas como foi citado na literatura [41], sendo citado sobre a dificuldade de realizar previsões de demanda para conseguir gerenciar o estoque de forma equilibrada e efetiva em uma grande empresa metalúrgica do cenário nacional. Em seu trabalho a autora compara a previsão de demanda entre um modelo desenvolvido em regressão multilinear e um código computacional baseado em redes neurais artificiais. Para realizar essa comparação os autores levaram em consideração os seguintes dados: Consumo específico, *lead time*, produção mensal e anual da aciaria e estoque de segurança, criando um banco de dados no período de janeiro/2015 a dezembro/2016.

De acordo com a autora [41] o método de regressão multilinear é atualmente utilizado pela empresa em questão e foi implementado em 2011, ocasionando uma mudança sistêmica para essa nova metodologia e com isso um ganho de R\$ 2,5 milhões em redução dos estoques de ferroligas. Tendo isso como um grande sucesso, as redes neurais terão que apresentar um excelente resultado para justificar a troca da metodologia.

O trabalho [41] descreve como objetivos: atingir um nível mais elevado de precisão, trazendo algumas vantagens como a eficácia na tomada de decisões estratégicas da empresa, redução do valor imobilizado do estoque e garantir

também um melhor atendimento aos clientes internos. Além disso, espera-se obter ganhos utilizando as Redes neurais artificiais, devido a um maior controle e previsibilidade da compra das matérias primas. [41]

Para realizar essa comparação só foram utilizados 3 materiais selecionados de maior impacto quanto ao valor de estoque e quanto a incerteza de demanda, sendo eles: Ferro manganês, Alumínio Gotão e Ferro Nióbio.

A conclusão do trabalho foi que o modelo de Redes neurais artificiais foi o mais apropriado do que a regressão multilinear para cálculos de previsão de demanda. O modelo de rede neural apresentou um erro médio de 0,52% contra 2,65% da regressão multilinear, onde os resultados do modelo de rede neural chegaram mais próximos aos valores encontrados no banco de dados de janeiro/2015 a dezembro/2016. Além disso ainda é citado que existe a oportunidade de se obter um ganho de R\$ 2 milhões em redução de estoque imobilizado se aplicar a nova metodologia. [41]

É possível encontrar vários trabalhos atuais aplicando modelos de redes neurais, como é visto no trabalho [35] realizado em 2021, onde se aplicou redes neurais para reconhecimento de descontinuidades em cordões de solda. Nesse trabalho foi utilizado a arquitetura convolucional para reconhecimento e classificação de imagens, no qual o autor utilizou 10 imagens de um cordão de solda no qual foram extraídas 96 fatias que foram as entradas do treinamento da rede. O trabalho mostra um nível de complexidade inferior ao que será empregado nesse mestrado, pois existem dois resultados potenciais: Com descontinuidade e sem descontinuidade, o que resulta em alta acurácia como demonstrado nos resultados. A rede neural alcançou 95% de acurácia no teste, ou seja, acertando 95 % das imagens classificadas [35].

Atualmente muitos trabalhos com o [39] utilizam ferramentas específicas para analisar de forma rápida o desempenho da rede neural e uma dessas ferramentas é a matriz de confusão, na qual consiste no fornecimento dos valores do conjunto de teste comparando sua predição com a classe correta de cada segmento. Os valores representam a quantidade de imagens classificadas onde são divididas entre classificação manual e automática (predição), Se a classificação manual for igual a classificação automática isso significa que a rede neural acertou

em cheio a classificação e quão maior forem os acertos maior será a tendência de formar uma linha na diagonal da matriz como demonstrado na Figura 14 através da linha vermelha. As linhas brancas demonstram as classificações vizinhas. Essas não são consideradas erros uma vez que podem apresentar pequena diferença em comparação com a severidade da classificação manual.

Portanto, a matriz de confusão proporciona o fácil entendimento de quantas imagens a rede neural consegue acertar e quais são as tendências de acertos e erros, contribuindo para ajustes e avaliações.

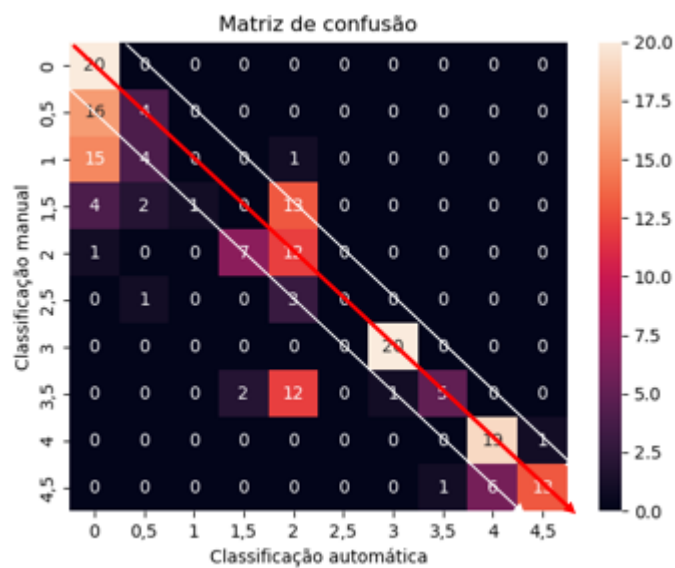


Figura 14 - Exemplo da leitura e interpretação da matriz de confusão.

3 MATERIAIS E MÉTODOS

Na realização deste Mestrado foi utilizada uma sequência metodológica, de acordo com as seguintes etapas:

a. Seleção dos aços

- i. Análise prévia dos aços para selecionar as amostras que constituíram o banco de dados;

b. Preparação metalográfica para a análise de inclusões / preparação dos corpos de prova;

- ii. Preparação das amostras selecionadas na etapa anterior.
- iii. Análise crítica da preparação verificando a qualidade do polimento, lavagem e secagem.

c. Classificação de inclusões não metálicas de forma manual conforme ASTM E45.

- iv. Estudo dos métodos e da norma de referência.
- v. Seleção do método utilizado.
- vi. Criação de um “*chart*” comparativo utilizando os aços selecionados para facilitar a classificação dos analistas.

d. Criação do banco de dados

- vii. Registro das imagens ópticas.
- viii. Análise crítica das imagens capturadas
- ix. Exclusão das imagens com defeitos de preparação.

e. Classificação das imagens computadorizadas

- x. Criação das imagens computadorizadas
- xi. Treinamento com banco de dados das imagens computadorizadas
- xii. Análise dos resultados

f. Classificação das imagens do banco de dados

- xiii. Classificação das imagens de acordo com o método A da norma ASTM E45
 - xiv. Segregação das imagens por nível de severidade
 - xv. Contagem da quantidade de imagens em cada severidade.
- g. Treinamento da rede neural*
- xvi. Realização do treinamento da rede neural com todas as imagens selecionadas para compor o banco de dados
 - xvii. Análise do resultado do treinamento
- h. Testes e validações na rede neural*
- xviii. Início da etapa de validação.
 - xix. Análise dos resultados da validação do treinamento
 - xx. Realização dos testes.
- i. Resultados*
- xxi. Análise dos resultados obtidos após todas as etapas de treinamento / validação e teste da rede neural. Aplicação de matrizes de confusão e gráficos de acurácia.
 - xxii. Análise dos resultados por nível de severidade.
 - xxiii. Avaliação da necessidade de ampliar o banco de dados em algum nível de severidade específico de acordo com o resultado da acurácia.
 - xxiv. Se ampliado o banco de dados é necessário voltar a etapa “g” de treinamento e seguir os passos lá estabelecidos.

3.1 Seleção dos materiais

Algumas considerações foram avaliadas para realizar a seleção dos aços utilizados nesse trabalho.

A primeira consideração foi o fato de que a rede neural será treinada apenas com inclusões do tipo A – Sulfeto. Assim sendo, aços com a presença de outros tipos de inclusões podem confundir o treinamento da rede neural e resultar em uma baixa acurácia.

A segunda consideração que foi avaliada está relacionada ao aspecto visual das inclusões na hora da análise microscópica. Contudo, inclusões mais nítidas ou em maiores quantidades facilitaram o treinamento da rede neural.

Após considerar essas avaliações, foram selecionados aços que apresentam preferencialmente inclusões do tipo A – Sulfeto em grandes quantidades e de forma nítida, o que aumentou a probabilidade de se obter boas imagens para a criação do banco de dados. Os aços utilizados para o treinamento dessa rede neural foram: 430F, 304L, 1050 e 8630. As composições químicas dos aços utilizados conforme Tabela 5. Todos os aços foram analisados depois do processo de conformação, forjamento ou laminação.

Tabela 5 - Composição química dos aços utilizados nas análises. Fonte: Villares Metals.

Aços	%C	%Mn	%P	%S
430F	0,045 – 0,060	0,80 – 0,90	0,03 Max	0,28 - 030
304L	0,03 Max	2,0 Max	0,040 Max	0,030 Max
1050	0.48 – 0,55	0,60 – 0,90	0,040 Max	0,050 Max
8630	0,27 – 0,33	0,65 – 1,05	0,015 Max	0,015 Max

3.2 Preparação metalográfica para a análise de inclusões

Para a realização da análise através da microscopia óptica é necessária uma preparação prévia do corpo de prova com alguns cuidados a serem tomados como: a planicidade da amostra e, em especial, evitar possível arredondamento das superfícies durante o polimento, e não menos importante, um polimento de alta qualidade para não influenciar na tomada de decisão e na classificação. [42]

A preparação metalográfica das amostras para a realização de uma análise micrográfica em microscópio óptico é dividida nas seguintes fases:

- 1 – Escolha da localização da seção a ser estudada.
- 2 –Obtenção de uma superfície plana e polida no local escolhido para a classificação.
- 3 – Exame no microscópio óptico para a observação das características visíveis sem ataque químico que são as inclusões, polibilidade e os defeitos de preparação incluindo os riscos e deformação na superfície

4 – Registro através de fotografia se necessário.

A escolha da seção é uma das principais e mais relevante técnica a ser aplicada na preparação metalográfica, pois a maioria dos ensaios metalográficos dependem e variam de seção para seção, pois existem alguns fatores como: nível de segregação, encruamento, sentido de deformação etc...., os quais podem alterar diretamente na análise dos resultados. [43]

A classificação da microinclusão é um ótimo exemplo para transparecer a importância dessa técnica. O ensaio de microinclusão deve ser realizado no sentido de deformação do aço mais especificamente na seção longitudinal do aço, levando em consideração que as inclusões se formam com o aço ainda no estado líquido, ou seja, antes de qualquer esforço de deformação em operações de conformação. Dessa forma, é necessário analisar as inclusões no sentido de deformação, pois elas também serão deformadas, podendo resultar em morfologias totalmente diferentes da inicial. Como todos os ensaios metalográficos são baseados em normas, o ensaio de classificação de microinclusão não é diferente, e a norma referente a esse ensaio (ASTM E45) cita que as inclusões devem ser classificadas de acordo com a sua morfologia. Por esse motivo, a importância da escolha e localização da seção estudada por metalografia, a retirada de amostra deve ser realizada no sentido longitudinal, ou seja, sentido de deformação, conforme a Figura 15. [43]

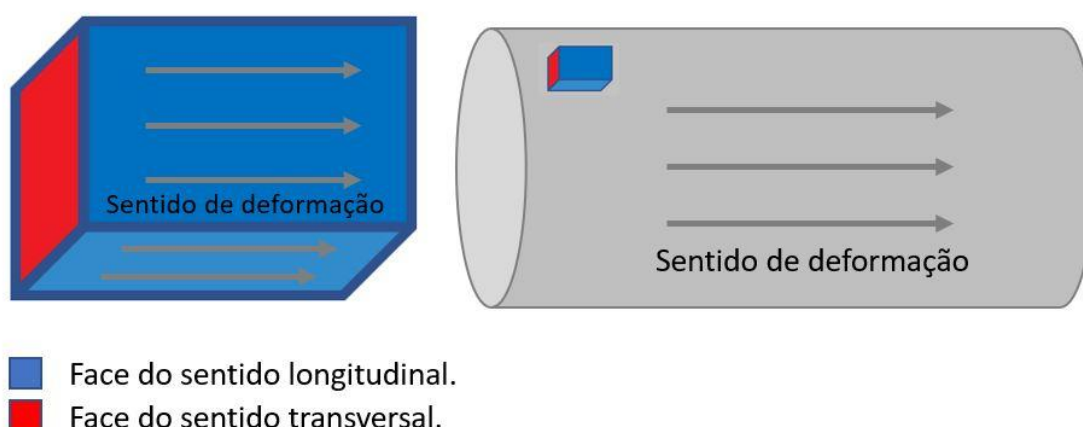


Figura 15 - Esquema ilustrativo de retirada de amostra para preparação e face de análise.

Para exemplificar um possível erro caso seja escolhido a seção incorreta de análise, comparam-se duas classificações de posições distintas (transversal/longitudinal) do mesmo material, onde observa-se que na transversal a classificação estabelecerá somente a presença de óxidos globulares, podendo ser totalmente diferente da posição longitudinal, onde pode-se encontrar (alumina, sulfeto e silicatos) características expressando deformação e alteração da morfologia das inclusões.

A obtenção de uma superfície plana e polida no local escolhido para estudo surge na preparação “clássica”, após o corte de serra, plaina, torno ou *cutoff* com discos abrasivos. Depois dessa etapa, há duas opções principais: (a) a amostra pode ser embutida em resina ou baquelite, o que facilita o manuseio do corpo de prova na etapa do lixamento / polimento, resultando em um melhor polimento superficial, preservando as arestas (superfícies); ou (b) a amostra pode ser preparada diretamente, sem a realização do embutimento.

A preparação da superfície de interesse envolve o lixamento em lixadeiras mecânicas seguindo uma sequência de papéis de lixa ou até mesmo a utilização de máquinas especializadas no polimento automático, eliminando o trabalho manual.

Após o emprego das lixas, a etapa de polimento é realizada sobre disco giratório de feltro, sobre o qual se aplica uma leve camada de um agente de polimento que é constituído por abrasivo em suspensão líquida. Os abrasivos mais comuns são: alumina e o diamante em pasta ou suspensão, porém o uso do abrasivo para o polimento geralmente varia de acordo com os ensaios a serem realizados. No ensaio de metalografia para revelar microinclusões o abrasivo mais utilizado é o diamante em suspensão com partículas de tamanho entre 6 μ m e 1 μ m.

O polimento e o lixamento podem ser realizados de forma manual ou em máquinas automatizadas, mantendo o cuidado com a superfície, porém existem cuidados adicionais no processo de polimento, quais sejam: cuidados com as condições de pressão e a rotação dos discos. O excesso de pressão ou a rotação incorreta podem causar defeitos de polimento. Os principais defeitos de polimento que podem ser encontrados são: cometas, manchas marrons, pittings e auréolas escuras e heterogêneas.

Cometas: Ocorrem devido à pressão excessiva durante o polimento ou a partículas que se destacam de inclusões não – metálicas duras e quebradiças (como alumina, por exemplo). [43]

Manchas marrons: Aparecem quando no fim do polimento, a pressão contra o abrasivo é fraca demais, e ocorrem mais frequentemente em aços com teor elevado de fósforo. [43]

Auréolas escuras e heterogêneas: Aparecem geralmente quando a lavagem do corpo de prova em água, após o polimento, é muito demorada ou não é feita a secagem imediata do corpo de prova depois de lavado. [43]

3.3 Preparação dos corpos de prova

Os corpos de provas foram preparados separadamente, porém o processo para todos os corpos de prova foi o mesmo, conforme descrito a seguir.

O processo de preparação se iniciou no corte do corpo de prova que foi realizado em uma cut-off do modelo COR-40 do fabricante Arotec, conforme a Figura 16. O corpo de prova foi cortado com refrigeração padrão da máquina e no sentido longitudinal.



Figura 16 - Equipamento cut-off utilizado para realizar o corte dos corpos de prova de aço.

Após o corte do corpo de prova, a face a ser analisada foi lixada em uma lixadeira correia do modelo LC01 do fabricante Teclago com uma lixa de granulometria #100 para remover toda e qualquer influência do disco abrasivo, conforme Figura 17.



Figura 17 - Lixadeira tipo correia com lixa 100 utilizada para lixar os corpos de prova de aço.

Em seguida a face lixada foi identificada e embutida a quente com baquelite. A embutidora utilizada foi a Simplimet 4000 Buehler, conforme fotografia da Figura 18.



Figura 18 - Embutidora da fabricante Buehler utilizada para embutir os corpos de prova de aço.

Depois do embutimento ser realizado, os baquelites com os corpos de prova foram levados para a máquina de polimento automático do modelo MAP'S 2 do fabricante Struers, Figura 19, onde são fixados em um suporte chamado “holder” para realizar toda sequência de polimento do corpo de prova.



Figura 19 - Máquina de polimento automático utilizada para realizar o polimento dos corpos de prova de aço.

A MAP'S apresenta diversas possibilidades de processos que podem ser utilizados de acordo com o material ou a análise a ser realizada, portanto, para esses corpos de prova o processo que foi utilizado seguiu os seguintes passos:

- 1- Pedra abrasiva
- 2- Prato abrasivo de polimento "Alegro" com uso de diamante em suspensão 9 μm .
- 3- Prato abrasivo de polimento "Alegro" com uso de diamante em suspensão 6 μm .
- 4- Pano de polimento "Mol" com uso de diamante em suspensão 1 μm .

Posteriormente ao processo de polimento automático, caso necessário ainda pode ser realizado um ajuste fino com polimento manual na politriz com mol e o uso de pasta de diamante de 1 μm , o que resultará em um resultado ainda melhor.

Foram examinados aproximadamente 10 corpos de prova para cada aço citado nesse trabalho, a partir dos quais foram capturadas as imagens e criados os bancos de dados. Foi utilizado o microscópio Zeiss do modelo ZEISS Axio Observer, conforme mostrado na Figura 20.



Figura 20 - Microscópio analisador de imagem utilizado para a criação do banco de dados de imagens dos aços.

3.4 Métodos de classificação de inclusões não metálicas conforme ASTM E45.

A seguir são descritos os métodos A, B, C, D e E, conforme a norma ASTM E45:

Método A “pior campo”: Este método exige uma área polida de no mínimo 160 mm^2 com um aumento de 100x. O tamanho do campo deve ser igual a área de $0,50 \text{ mm}^2$, e cada campo de $0,50 \text{ mm}^2$ deve ser comparado com o quadro e as imagens JK (Plate I-A), no qual se classifica o pior campo com a maior severidade encontrada de cada tipo de inclusão A, B, C e D para ambas subclassificações: série fina ou grossa. O nível de severidade do pior campo deve ser reportado para todos os corpos de prova examinados. [8]

Existem duas técnicas manuais que podem ser empregadas para encontrar um campo de $0,50 \text{ mm}^2$. Na primeira técnica, envolve projetar a imagem de 100x do microscópio na direção de uma máscara quadrada de 71 mm (2.79 in.). A segunda opção é utilizar um reticulado feito para o microscópio no qual irá sobrepor a máscara quadrada diretamente no campo de visão. A Figura 21 a seguir mostra o reticulado.[8]

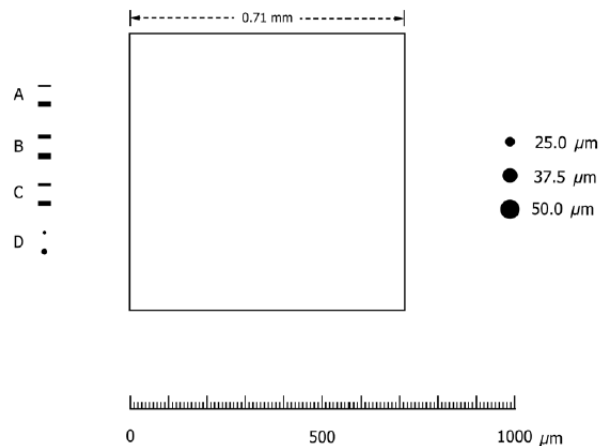


Figura 21 - Reticulado utilizado para classificação das inclusões. Fonte: Norma ASTM E45.[8]

Método B “Comprimento”: Este método exige uma área superficial polida de 160 mm², com um aumento de 100x. Neste método qualquer inclusão onde o comprimento seja de 0,127mm, ou maior, deve ser medida e individualmente computada.[8]

Este método utiliza padrões de linhas paralelas no qual a distância das linhas é equivalente a 0,127mm quando visto em 100x. O padrão pode ser desenhado ou pode ser utilizado o reticulado feito para o microscópio que irá sobrepor o padrão sobre a imagem vista pelo microscópio. O reticulado pode ser visto na Figura 22 a seguir.[8]

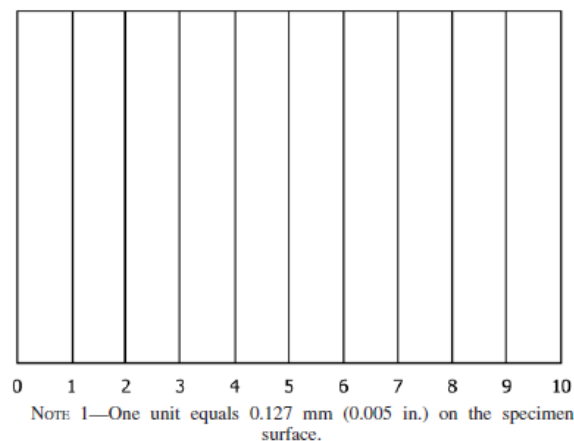


Figura 22 - Reticulado utilizado para a classificação de inclusões no método B. Fonte: Norma ASTM E45.[8]

Para iniciar o teste, deve se marcar a área de teste, e todas as inclusões dentro desse campo demarcado devem ser medidas e computadas.

Para expressar os resultados, o comprimento da maior inclusão deve ser considerado, e deve ser complementado da sua largura. As subclassificações como d (disconnected), vd (very disconnected), e g (grouped) podem ser usadas para descrever o grau de conectividade ou 'clusters', como pode ser observado na Figura 23.[8]

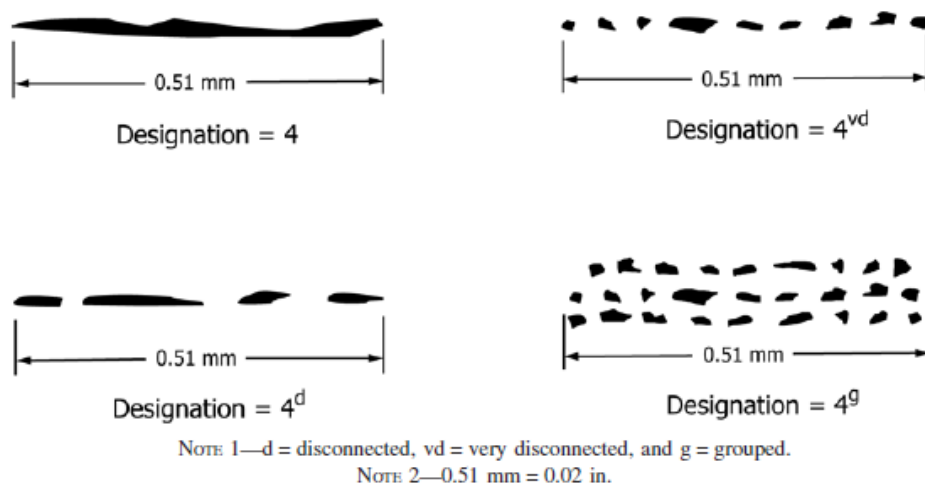


Figura 23 - Chart utilizado para a subclassificação das inclusões no método do comprimento.

Fonte: Norma ASTM E45.[8]

A média de todas as inclusões deve ser reportada como um único número, seguido pelo número de inclusões.

Para exemplificar o resultado para uma simples amostra neste método o resultado seria: 6d-2³-A. O resultado indica que a maior inclusão encontrada foi até a linha 6 e apresentava um aspecto visual semelhante a subclassificação "disconnected". Indica também que foram encontradas mais 3 inclusões, onde a média dessas inclusões se encontra na unidade 2 e todas as inclusões foram classificadas como sendo do tipo A.[8]

Método C "Óxidos e Silicatos" – O método C será introduzido no trabalho, porém não será estudado profundamente, uma vez que este método não será utilizado para a classificação e treinamento da rede neural. Este método exige uma área mínima de 160 mm² de superfície polida com o aumento de 100x. Em cada campo da amostra deve ser analisado a presença de óxidos de alumina não deformável e silicatos deformáveis, e classificar em comparação com o Plate II da norma.[8]

Método D “Conteúdo de baixa inclusão” – O método D é pretendido para a aplicação em aço com baixo nível de inclusão, onde todos os níveis de severidade devem ser relatados. A amostra deve apresentar uma área mínima de superfície polida de 160mm² e para cada campo de 0,50mm² é examinado todos os tipos de inclusões: A, B, C e D e comparado com o “chart” Plate I-A. O resultado desse método leva em consideração todos os campos analisados, ou seja, todos os campos comparados devem ser registrados e computados.[8]

Método E “SAM Rating” – Este método é usado para classificar inclusões de aço de uma maneira que reflita a severidade e a frequência de maior ocorrência nos tipos B (alumina) e D (óxido globular).[8]

3.5 Criação do banco de dados

Inicialmente, foram criados dois bancos de dados distintos para treinamento da rede neural sendo eles: banco de dados com imagens computadorizadas e o banco de dados com imagens reais.

O objetivo da criação de dois bancos de dados se fez necessário para o conhecimento e descobrimento do comportamento da rede neural analisando imagens de microinclusões de sulfetos de manganês. Por esse motivo foram criadas 800 imagens computadorizadas conforme descrito no capítulo 4 para se obter um resultado prévio e entender todas as dificuldades computacionais que teríamos. Após ótimos resultados com o banco de dados computadorizados foi iniciada a criação do banco de dados de imagens reais.

A criação do banco de dados de imagens reais é o resultado de vários processos essenciais já citados nesse trabalho como seleção do material, preparação metalográfica e classificação das microinclusões conforme norma ASTM E45 método “worst field” “A”.

Para a realização da classificação e criação do banco de dados foram registradas 3746 imagens com o aumento de 100x, sendo estas segmentadas em 3 pastas diferentes, treinamento, validação e teste. As imagens foram divididas em 80% para treinamento e 20% para validação e as imagens para teste foram escolhidas aleatoriamente. Essa proporção foi utilizada no trabalho [44] utilizado

como referência, e as imagens estão sem a presença de escala micrométrica para não influenciar no treinamento da rede neural, .

Os bancos de dados utilizados nesse trabalho foram criados e classificados especificamente para atender a demanda da Villares Metals, o que conseqüentemente dificulta o treinamento, pois não se caracteriza por ser um banco de dados padronizado como aqueles citados nos trabalhos [44] e [45]. Nestes trabalhos os bancos de dados contam com números de imagens padronizados e já classificados. Por se tratar de análise de microinclusões em aços com composições químicas e graus de deformação específicos, dificilmente seria gerado um banco de dados padronizado para essa aplicação, e dessa forma foram criados os bancos de dados representativos.

O banco de dados de treinamento foi utilizado para o treinamento e aprendizado da rede neural, já o banco de dados de validação foi utilizado para medir a acurácia do treinamento e aprendizado do treinamento, e por último o banco de dados do teste, que foi avaliado manualmente pelos especialistas para a comparação com o resultado da rede neural.

As imagens foram capturadas de forma aleatória, de diversos campos e aços diferentes. Com as fotos prontas, o próximo passo foi a classificação de acordo com a norma, e a classificação foi sustentada pela criação de “charts” comparativos dos próprios aços utilizados, como citado no capítulo anterior.

Em seguida, as imagens foram preparadas para a rede neural através de um tratamento onde foram normalizadas as cores e reduzida a resolução das imagens.

O programa utilizado para realizar esse tratamento foi o Python, utilizando o Open CV que teve a função de reduzir a resolução das imagens para 160 x 120 (pixels).

A metodologia utilizada no trabalho fica evidente no fluxograma da Figura 24, onde mostra todas as etapas do trabalho desde preparação manual até o início da criação do banco de dados.

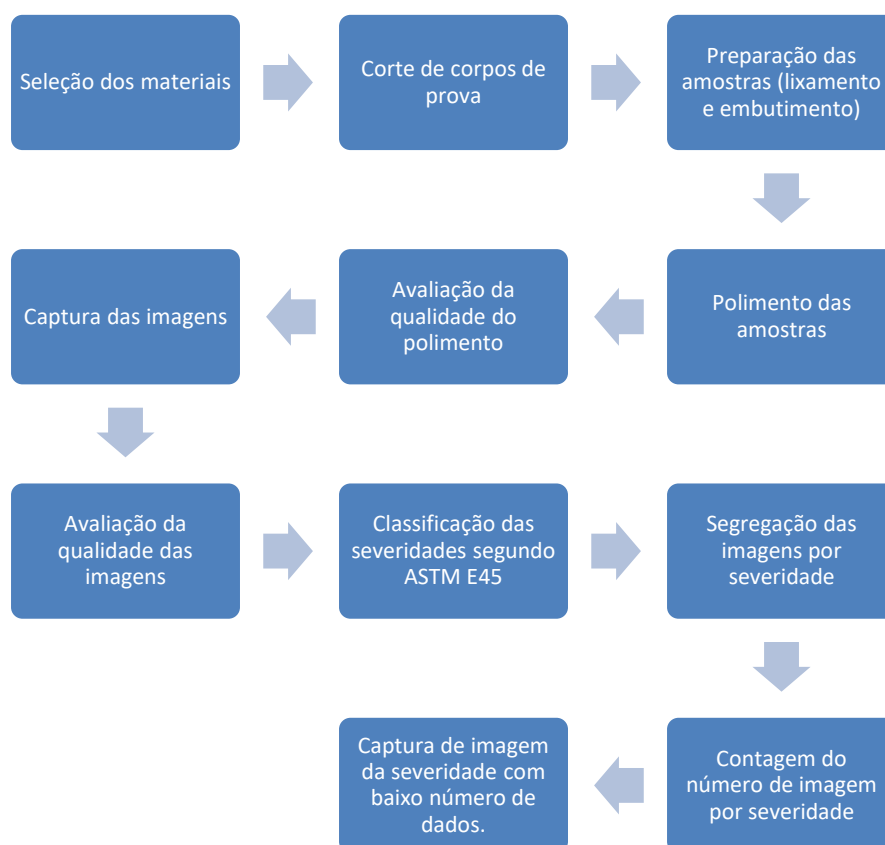


Figura 24 – Fluxograma da metodologia e etapas desse trabalho.

3.6 Classificação das imagens do banco de dados

As imagens do banco de dados foram classificadas pelos analistas com o auxílio dos charts. A criação do chart tem como objetivo facilitar a classificação e aumentar a assertividade da análise. O chart foi criado com imagens aleatórias, onde foi demarcado o campo de 0,5 mm², exigido em norma, e realizada a medição de cada inclusão que se encontrava dentro desse campo demarcado, condizendo com os parâmetros de classificação. O chart foi criado de acordo com a Tabela 6 abaixo, onde todas as imagens foram medidas e classificadas, tendo assim uma referência de severidade.

Tabela 6 - Valores mínimos para a classificação da severidade. (Métodos A, D e E) [8]

(mm(in.)100x				
Severidade	A	B	C	D
0,5	3,7(0,15)	1,7(0,07)	1,8(0,07)	1
1,0	12,7(0,50)	7,7(0,30)	7,6(0,30)	4
1,5	26,1(1,03)	18,4(0,72)	17,6(0,60)	9
2,0	43,6(1,72)	34,3(1,35)	32,0(1,26)	16
2,5	64,9(2,56)	55,5(2,19)	51,0(2,01)	25
3,0	89,8(3,54)	82,2(3,24)	74,6(2,94)	36
3,5	118,1(4,65)	114,7(4,52)	102,9(4,06)	49
4,0	149,8(5,90)	153,0(6,02)	135,9(5,35)	64
4,5	189,8(7,47)	197,3(7,77)	173,7(6,84)	81
5,0	223,0(8,78)	247,6(9,75)	216,3(8,52)	100
(µm(in.) 1x				
Severidade	A	B	C	D
0,5	37,0(.002)	17,2(.0007)	17,8(.0007)	1
1,0	127,0(.005)	76,8(.003)	75,6(.003)	4
1,5	261,0(.010)	184,2(.007)	176,0(.007)	9
2,0	436,1(.017)	342,7(.014)	320,5(.013)	16
2,5	649,0(.026)	554,7(.022)	510,3(.020)	25
3,0	898,0(.035)	822,2(.032)	746,1(.029)	36
3,5	1181,0(.047)	1147,0(.045)	1029,0(.041)	49
4,0	1498,0(.059)	1530,0(.060)	1359,0(.054)	64
4,5	1898,0(.075)	1973,0(.078)	1737,0(.068)	81
5,0	2230,0(.088)	2476,0(.098)	2163,0(.085)	100

Com o auxílio do “chart” a classificação foi realizada por 2 técnicos metalógrafos experientes que somam juntos mais de 20 anos de experiência em classificação de microinclusão. Os técnicos Aleksandro dos Santos Brandão e Marcos Fernando de Oliveira Filho foram os responsáveis pela criação e classificação do banco de dados. Após a classificação conforme a norma ASTM E45, que resultou na severidade de cada imagem, as imagens foram divididas em

pastas de acordo com a sua classificação variando entre 0 a 5,0 AF. Só foram consideradas as classificações que apresentaram um número suficiente de imagens para realizar o treinamento.

A Tabela 7 mostra o número de imagens que compõe o banco de dados. Essas imagens foram utilizadas para realizar o primeiro treinamento da rede neural com imagens reais.

Tabela 7 - Número de imagens utilizada no primeiro treinamento da rede neural.

Severidade	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
Quantidade de imagens no banco de dados	92	328	548	288	44	4	144	616	760	328	4

Após o primeiro treinamento foi verificada a necessidade de aumentar o banco de dados para as severidades 1,0 / 1,5 / 2,0 e 2,5. A Tabela 8 mostra a quantidade de imagens após a inserção de imagens adicionais.

Tabela 8 – Número de imagens utilizadas no segundo treinamento da rede neural.

Severidade	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
Quantidade de imagens no banco de dados	92	328	748	832	248	36	144	616	760	328	4

Após realizar alguns testes com o novo banco de dados foi verificado que a acurácia do treinamento e validação aumentaram, porém, os resultados dos testes não corresponderam com os resultados de treinamento e validação. Dessa forma foi necessário reclassificar todas as imagens por severidade, sendo obrigatório a exclusão de imagens com baixa qualidade ou iluminação fora do padrão, que poderiam contribuir para a confusão da rede neural, conforme mostrado na Figura 25. A Tabela 9 mostra a quantidade final de imagens após a reclassificação.



Figura 25 - Exemplo de imagem que foi excluída pela presença de riscos, os quais podem confundir a rede neural.

Tabela 9 – Número de imagens utilizadas no treinamento final da rede neural.

Severidade	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
Quantidade de imagens no banco de dados	80	108	124	528	1152	32	124	528	508	308	4

3.7 Rede neural

A construção da rede neural foi feita através da definição das camadas de entrada e saída, como também das estruturas das camadas ocultas. Neste trabalho foi utilizada a seguinte estrutura de rede neural, ilustrada na Figura abaixo (Figura 26). A camada de entrada é uma convolucional 2D, com dimensão de entrada (160, 120, 1), onde 160x120 pixels é o tamanho de cada imagem e o valor de 1 é para a escala de cinza (que apresenta apenas um valor). Os valores dos pixels variam de 0 a 255 na escala de cinza, e para entrada na rede neural, estes valores são divididos por 255 para variarem no intervalo de 0 a 1.

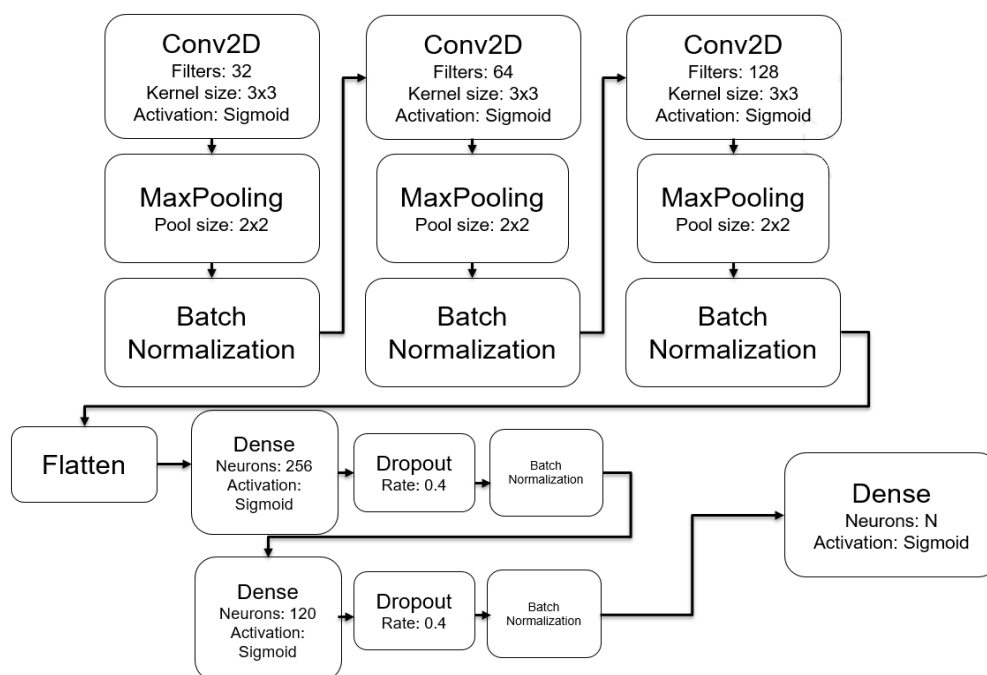


Figura 26 - Esquema das camadas da rede neural utilizada.

Esta estrutura é definida por 3 sequências de camadas convolucionais, “max pooling” e “batch normalization”, seguida por 2 sequências de camada densa, “dropout” e “batch normalization”. A camada final é densa e tem o número de neurônios igual ao número de classificações (N). Outras estruturas de redes neurais foram testadas, porém, esta estrutura apresentou maior capacidade de aprendizado do que outras. Esta rede neural tem a sua estrutura próxima à da rede neural VGG16, apenas modificando a quantidade de neurônios e a função ativação de ReLU para sigmoidal. [46]

3.8 Testes e validações na rede neural

Para construção do banco de imagens, são separados três grupos principais de imagens: banco de treinamento, banco de validação e banco de teste. No banco de treinamento estão as imagens que foram utilizadas pela rede neural para seu treinamento. No banco de validação estão as imagens que foram avaliadas durante treinamento da rede neural, porém, não entrarão para o cálculo da função perda. O banco de validação é essencial para que se observe o aprendizado da rede neural

para o problema. O banco de testes contém imagens que foram utilizadas após treinamento para verificar a acurácia do modelo com imagens reais diversas.

3.9 Fluxograma de aplicação da rede neural

O processo de aplicação de uma rede neural pode ser trabalhoso e muito longo, de acordo com o segmento a ser aplicado. Muitas vezes se trata de um processo cíclico, onde é necessário realizar diversos ajustes para encontrar o melhor resultado possível e esse processo fica evidente na Figura 27 que demonstra todo o fluxo.

Para a aplicação de CNN na classificação de microinclusões não foi diferente, foram realizados diversos ajustes visando aumentar a acurácia e melhorar o treinamento da rede neural. E como toda rede neural esse processo teve início na criação do banco de dados, sendo ela a primeira e mais importante etapa dentro do processo. Em função a relevância dessa etapa no processo muitas vezes é necessário retrabalhar o seu banco de dados para obter um bom resultado.

Após a criação de um banco de dados consistente e representativo inicia-se a separação das imagens para realizar o treinamento, a validação e o teste. A separação foi realizada de forma planejada, pois é essencial que haja imagens de todas as severidades no banco de dados de validação e teste, pois caso não haja imagens de algumas severidades dentro do banco de dados, o resultado do treinamento dessa severidade pode ser um resultado de acurácia falso positivo.

Posteriormente à divisão do banco de dados, as imagens foram avaliadas para verificar a necessidade de realizar a normalização da tonalidade, para que isso não influencie no treinamento e na classificação.

O custo computacional é um desafio conhecido dentro do machine learning. A rede neural analisa e classifica a imagem através de um processo de avaliação de pixels, e quanto maior a resolução da imagem maior a quantidade de pixel a ser avaliado. E conseqüentemente maior custo computacional na aplicação da rede neural. Por esse motivo reduzimos a resolução das imagens, com o objetivo de otimizar o processamento computacional. O processamento após a redução das imagens foi de aproximadamente 12 horas utilizando um computador com

processador AMD RYZEN 5 2600X, placa de vídeo NVIDIA GeForce RTX2060 SUPER e sistema operacional Windows 10 PRO.

Em seguida a todo o processo de preparação se dá o início do treinamento da rede neural, que tem como objetivo ensinar todo o banco de dados para a inteligência artificial. Após a conclusão do treinamento, foi realizada a avaliação da acurácia do treinamento, para avaliar se o treinamento foi eficaz ou não. Caso o resultado for negativo é necessário realizar alguns ajustes visando o aumento da acurácia no treinamento.

Os ajustes variam de acordo com os problemas encontrados, podendo ser: dados insuficientes no banco de dados, imagens com baixa qualidade, falta de representatividade, entre outros...

Seguidamente ao treinamento bem-sucedido começou o processo de validação com a finalidade de validar ou não o treinamento da rede neural. Neste processo foram classificadas 20% das imagens do banco de dados pela rede neural recém treinada. O objetivo do processo de validação é comprovar que a rede neural realmente alcançou o aprendizado.

Na eventualidade da validação comprovar que o treinamento não foi eficaz, ou seja, resultando em um valor baixo de acurácia, é necessário realizar ajustes na rede neural/banco de dados.

Em conclusão ao processo de desenvolvimento de uma rede neural, existe o último passo onde se compara o resultado encontrado pela rede neural com o padrão conhecido, sendo ele um especialista ou uma imagem padronizada. O teste é realizado com imagens aleatórias e o desempenho da rede neural deve acompanhar os resultados obtidos anteriormente no processo de treinamento e validação.

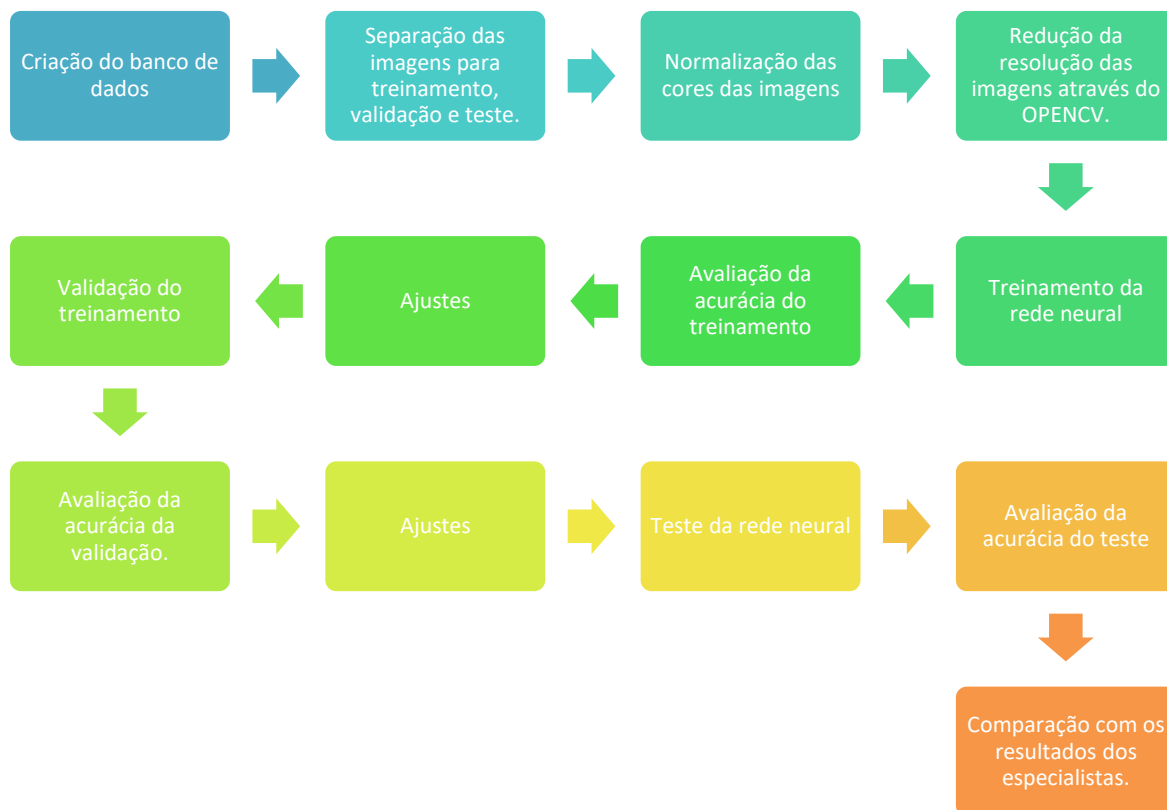


Figura 27 – Fluxograma de aplicação da rede neural.

4 RESULTADOS E DISCUSSÃO

4.1 Resultados utilizando imagens computadorizadas

Como primeiro resultado do modelo de redes neurais, foram utilizadas imagens geradas por computador para verificar em antemão como seria o treinamento para imagens reais. Para geração destas imagens, foram distribuídas 200 elipses em cada imagem definidas pela equação:

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} < R^2 \quad \text{Eq.4.1}$$

Onde x_0 é a posição horizontal do centro, y_0 é a posição vertical do centro (os quais são dados aleatoriamente), a e b são constantes, as quais $b=1$ e a é a razão de aspecto, e R um tamanho de raio definido por uma distribuição normal de média igual a 10 e desvio-padrão igual a 0,5. As imagens foram diminuídas de tamanho 400 x 400 pixels para tamanho de 160x120 pixels. Foram selecionadas quatro razões de aspecto para definir as classificações sendo elas: 1, 3, 5 e 7 (Figura 28); e geradas 200 imagens de cada classificação para treinamento, validação e teste.

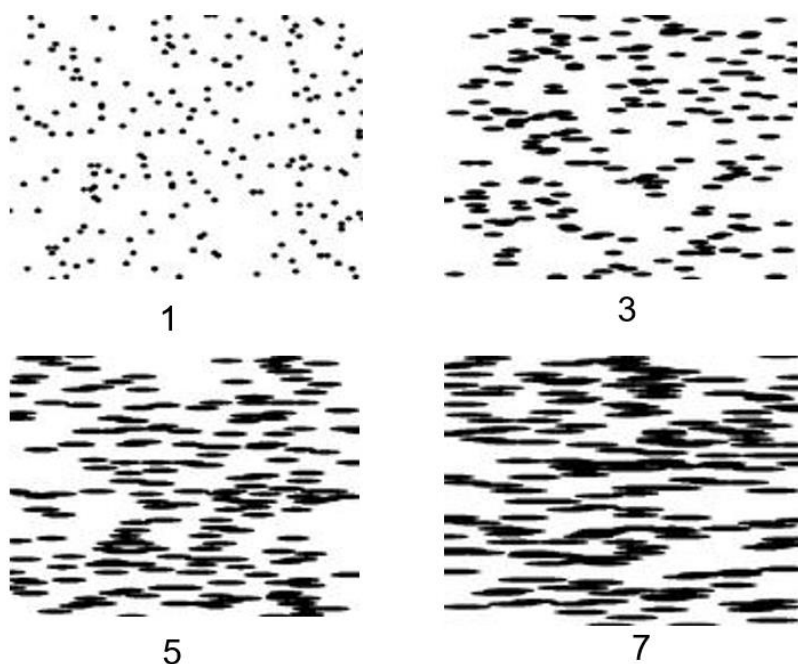


Figura 28 - Exemplos de imagens com elipses com diferentes razões de aspecto de 1, 3, 5 e 7.

Para seu treinamento foi utilizado tamanho de batelada (“Batch size”) de 400 imagens e 20 épocas. As épocas são caracterizadas como uma rodada de treinamento onde a rede neural treina com todo o banco de dados e volta realizando o ajuste dos pesos. Em cada época, computa-se a função perda e a acurácia tanto no conjunto de imagens para treinamento quanto no conjunto de imagens para validação. É possível observar na Figura 29 que a rede neural apresentou bom aprendizado, pois atingiu a acurácia de 100% no conjunto de treinamento, enquanto atingiu acurácia de 98% no conjunto de validação, ambos apresentando decréscimo da função perda.

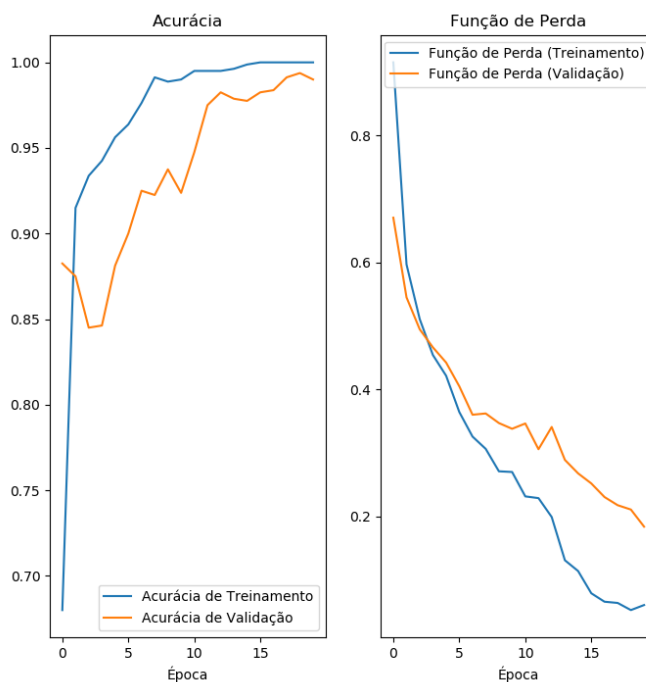


Figura 29 - Resultados de aprendizado da rede neural para o caso de elipses, com a acurácia e a função perda para treinamento e para validação.

Após os primeiros resultados utilizando imagens computadorizadas foi iniciado o treinamento utilizando imagens reais, o treinamento com imagens computadorizadas foi muito importante para ajustar a programação da rede neural e favorecer o aprendizado.

4.2 Resultados utilizando o banco de dados com imagens reais.

O treinamento com imagens reais proporcionou diversos desafios não encontrados no treinamento com imagens computadorizadas, pois no treinamento das imagens computadorizadas não existem variáveis como má preparação, artefatos, diferença de cor e diferença de resolução nas imagens compondo os bancos de dados.

Apesar das variáveis presentes no processo de treinamento com imagens reais, os resultados encontrados foram positivos para uma primeira versão.

O banco de dados do primeiro teste foi composto de diversas imagens de diferentes açoes, sendo a quantidade de imagens de cada severidade demonstrada na Tabela 10.

Tabela 10 - Número de imagens por severidade utilizadas no primeiro teste.

Severidade	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
Quantidade de imagens no banco de dados	92	328	548	288	44	4	144	616	760	328	4

Na Figura 30 é possível observar o resultado do treinamento com imagens reais utilizando 40 épocas, onde a acurácia chegou a aproximadamente 95% no treinamento e 70% na validação e a função perda para o treinamento atingiu 0,2 no treinamento e aproximadamente 0,7 na validação. Os resultados foram considerados positivos, mas não o suficiente para aplicação. Foi aumentado o número de épocas no treinamento para verificar como a rede neural responderia.

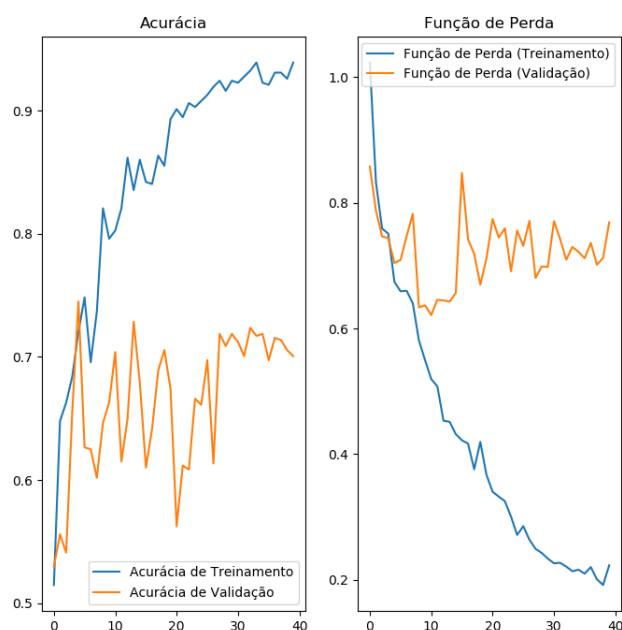


Figura 30 - Resultados de aprendizado da rede neural para o caso de fotos reais, com a acurácia e a função perda para o banco de imagens para treinamento e para validação utilizando 40 épocas.

Na Figura 31 é demonstrado que mesmo aumentando o número de épocas no treinamento a melhora na acurácia não é significativa, ou seja, a rede neural não está mais aprendendo com o treinamento. O resultado após o treinamento de 100 épocas foi aproximadamente 92% de acurácia no treinamento e 75% de

acurácia na validação e em relação à função perda os resultados foram aproximadamente 0,2 no treinamento e 0,8 na validação.

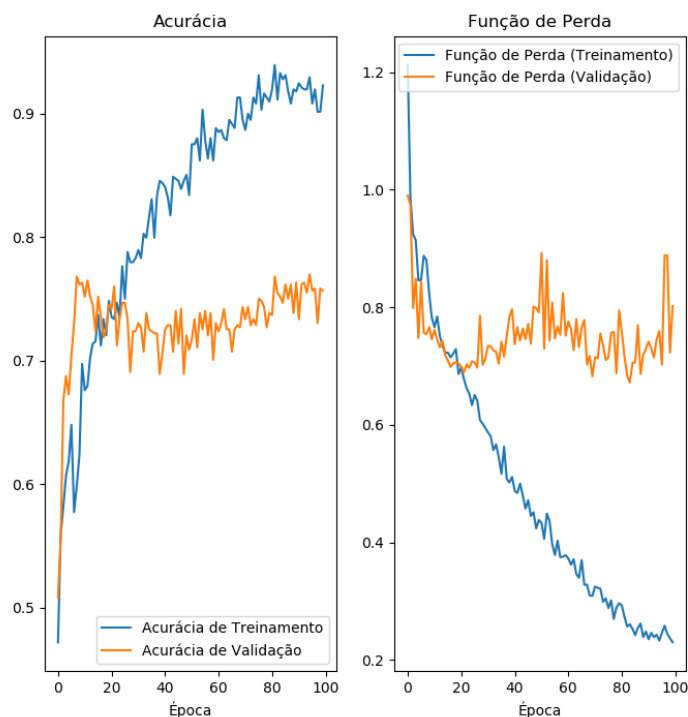


Figura 31 - Resultados de aprendizado da rede neural para o caso de fotos reais, com a acurácia e a função perda para o banco de imagens para treinamento e para validação utilizando 100 épocas.

Após análise do segundo treinamento utilizando maior número de épocas foi necessário avaliar o porquê da confusão da rede neural na validação, e para entender o que poderia estar acontecendo foi utilizada uma ferramenta chamada matriz de confusão que indica através de uma matriz quais as imagens que a rede neural está classificando de forma incorreta. A matriz de confusão não é uma ferramenta para combater o erro, mas sim mostrar o que pode estar acontecendo de errado no treinamento.

Analisando a matriz de confusão na Figura 32 é possível observar que a rede neural erra mais no caso de imagens com severidades próximas e aquelas com menor número de imagens. Isso ocorre porque a linha limite de classificação é tênue a ponto de confundir até mesmo os analistas mais experientes, sendo um reduzido número de imagens por severidade para treinamento um fator agravante.

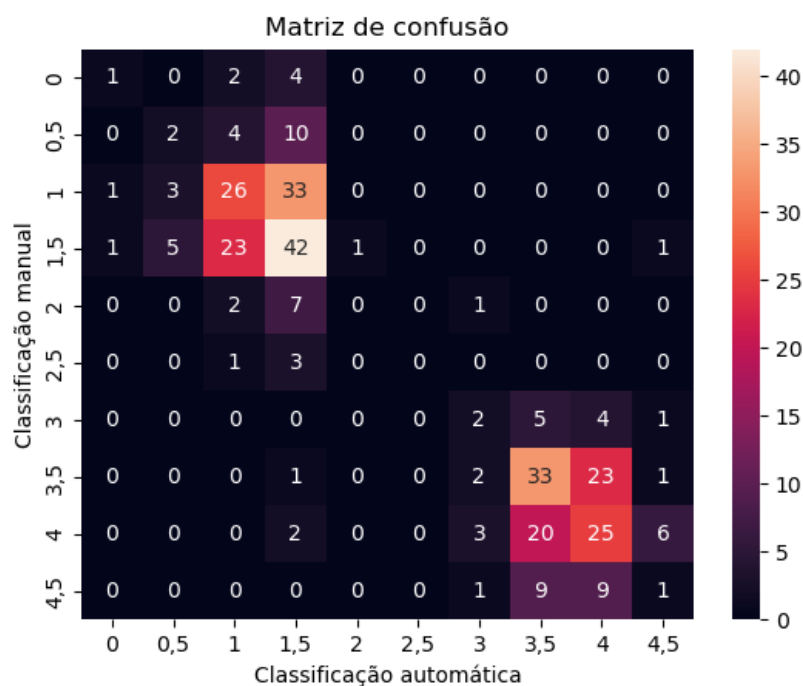


Figura 32 - Matriz de confusão após o primeiro treinamento, comparando classificação manual e classificação automática (100 épocas).

Após analisar os resultados da matriz de confusão foi observado que a porcentagem de acerto diminui proporcionalmente ao número de imagens disponíveis no banco de dados. Dessa forma, o banco de dados foi aumentado quantitativamente para as severidades 1,0; 1,5; 2,0 e 2,5. O banco de dados após a ampliação do número de imagens ficou conforme mostra a Tabela 11.

Tabela 11 - Quantidade de imagens por severidade utilizadas no segundo teste.

Severidade	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
Quantidade de imagens no banco de dados	92	328	748	832	248	36	144	616	760	328	4

Com a ampliação do banco de dados o resultado não foi como o esperado, a acurácia de validação foi menor que aquela relacionada ao primeiro treinamento, e conseqüentemente a acurácia do teste também foi menor. Na Figura 33 fica

evidente que a rede neural apresentou mais dificuldade na aprendizagem do que no primeiro teste.

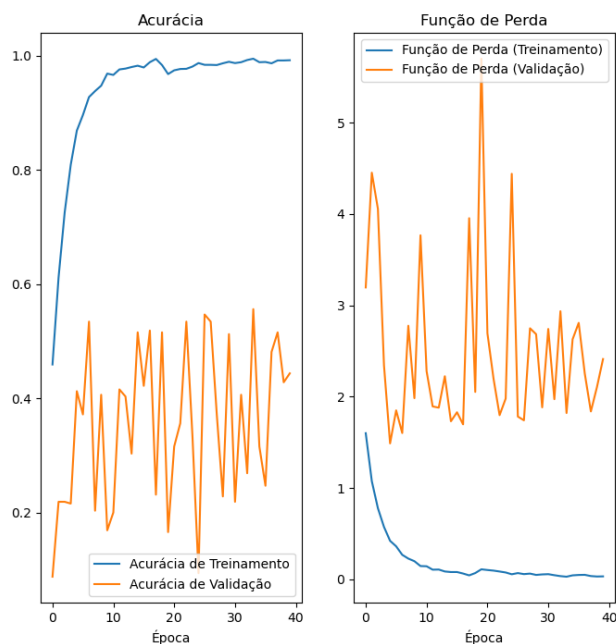


Figura 33 - Resultados de aprendizado da rede neural no segundo teste após ampliar o banco de dados, treinamento e validação, utilizando 40 épocas.

Apesar dos resultados do segundo teste apresentarem um alto valor de acurácia de treinamento, o valor de acurácia de validação deixa claro que a rede neural não aprendeu as características das imagens e isso fica mais evidente na Figura 34, onde é demonstrada a matriz de confusão do segundo teste.

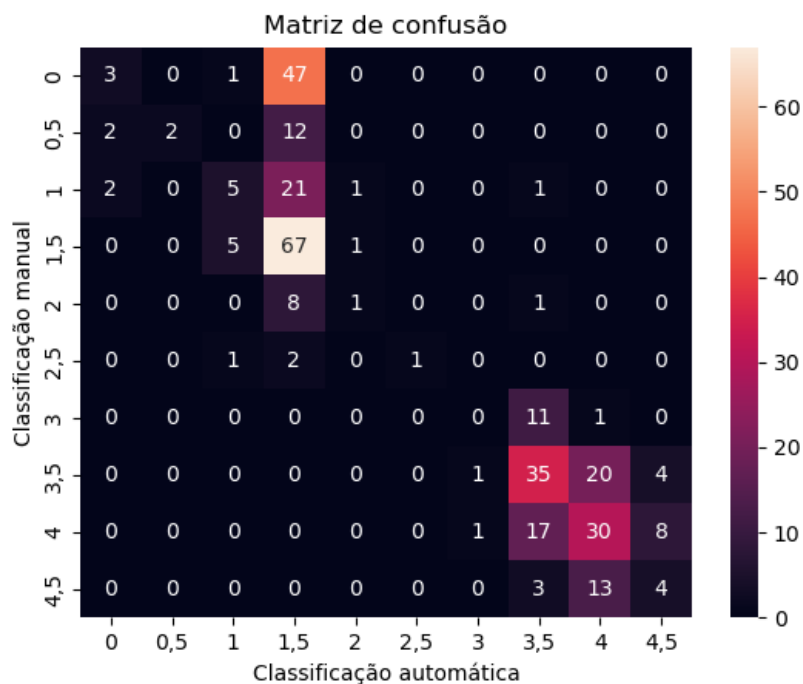


Figura 34 - Matriz de confusão após o segundo treinamento, comparando entre classificação manual e classificação automática (40 épocas).

Após analisar os resultados do primeiro e do segundo testes, observou-se que o aumento da quantidade de imagens no banco de dados não influenciou na acurácia da validação e nos resultados dos testes como era o esperado. Portanto somente a ampliação do banco de dados não foi suficiente para aumentar a acurácia e melhorar a aprendizagem.

A ampliação do banco de dados tem como objetivo aumentar e explorar a maior quantidade de inclusões e suas respectivas características. Essa técnica pode ter efeito contrário se a classificação não for realizada de forma correta, resultando em um treinamento confuso e de baixa qualidade de aprendizagem.

A rede neural apresentou muita dificuldade para classificar corretamente as imagens das severidades 0; 0,5; 1,0; 2,0; 2,5; 3,0; 4,5. No intuito de corrigir o problema, o banco de dados foi reclassificado, onde foi possível encontrar algumas imagens classificadas incorretamente. o que pode ter confundido a rede neural, resultando na baixa acurácia de validação e nos testes.

A reclassificação consistiu em acessar imagem por imagem de todas as severidades e reclassificar utilizando o “Chart” da norma ASTM E45 para auxiliar na tomada de decisão. Sendo assim foi observada a existência por exemplo de algumas imagens muito semelhantes e, conseqüentemente, indicando uma mesma severidade, porém em pastas distintas. Após a identificação dessa divergência, as imagens foram realocadas para a classificação correta e o banco de dados foi demonstrado na Tabela 12.

Tabela 12 - Quantidade de imagens no banco de dados por severidade após a reclassificação.

Severidade	0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
Quantidade de imagens no banco de dados	80	108	124	528	1152	32	124	528	508	308	4

A Figura 35 mostra os resultados obtidos após a reclassificação do banco de dados.

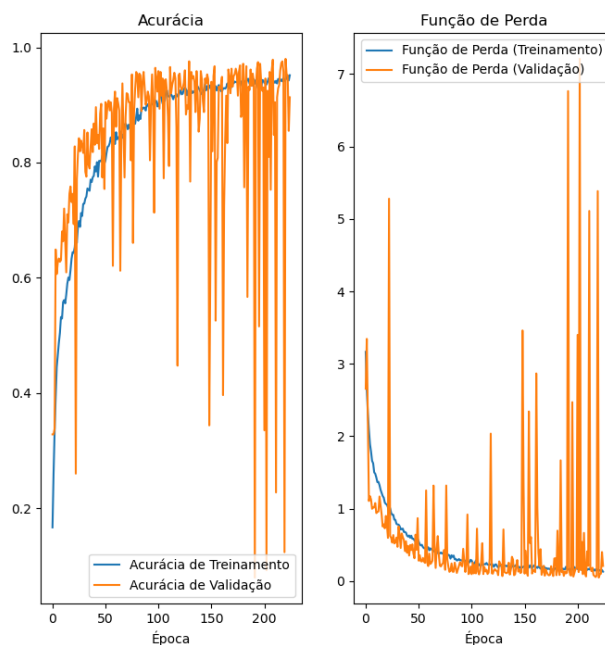


Figura 35 - Resultados de aprendizado da rede neural no terceiro teste após reclassificação do banco de dados, treinamento e validação utilizando 200 épocas.

Os resultados do terceiro teste foram positivos atingindo 90% de acurácia em treinamento e validação, resultados muito superiores aos testes anteriores. A Figura 36 indica a matriz de confusão do teste após o terceiro treinamento. Os resultados foram melhores que os treinamentos anteriores, porém a rede neural ainda apresenta uma confusão na classificação das severidades 0,5; 1,0; 1,5; 3,5 e uma tendência de classificação mais apurada para as severidades 0 e 2,0.

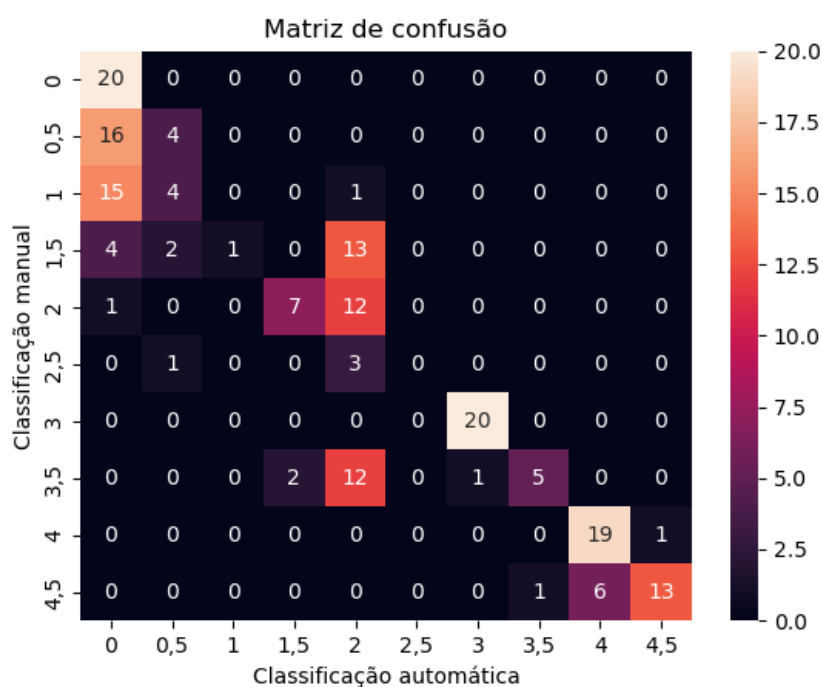


Figura 36 - Matriz de confusão após o terceiro treinamento, comparando classificação manual e classificação automática (200 épocas).

Uma vez que o banco de dados foi reclassificado de forma correta, foi desconsiderada uma possível confusão por parte de classificações imprecisas. Deste modo, foram alteradas as configurações da rede neural para que ela entenda e aprenda as características essenciais para a classificação. Objetivando diminuir a tendência de maior classificação nas severidades 0 e 2,0, aplicamos a técnica do *Dropout* já citada no capítulo anterior. O *Dropout* seleciona as características relevantes para a classificação ser precisa e combate o problema de “*overfitting*”, conforme citado no trabalho [47].

O resultado foi positivo no treinamento e na validação. Na Figura 37 podemos observar que os resultados foram superiores a todos os outros testes já realizados.

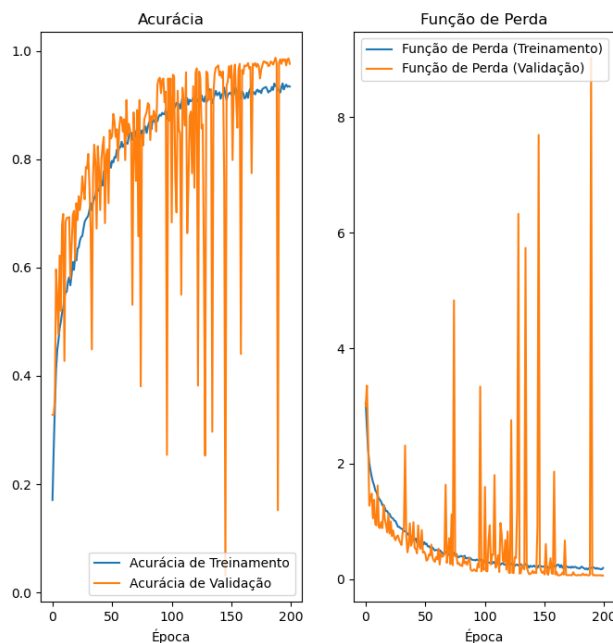


Figura 37 - Resultados de aprendizado da rede neural no quarto teste após aplicar a técnica de Dropout, treinamento e validação utilizando 200 épocas.

Os resultados após a aplicação da técnica *Dropout* trouxeram a aprendizagem das principais características das imagens e isso ficou visível no treinamento e validação, onde resultados de 93% de acurácia no treinamento e 96% de acurácia na validação foram alcançados. A Figura 38 mostra que os resultados do teste comparando a classificação automática com a manual também foram muito superiores aos testes anteriores e além disso o resultado mostrou que a tendência de maior classificação em algumas severidades foi minimizada.

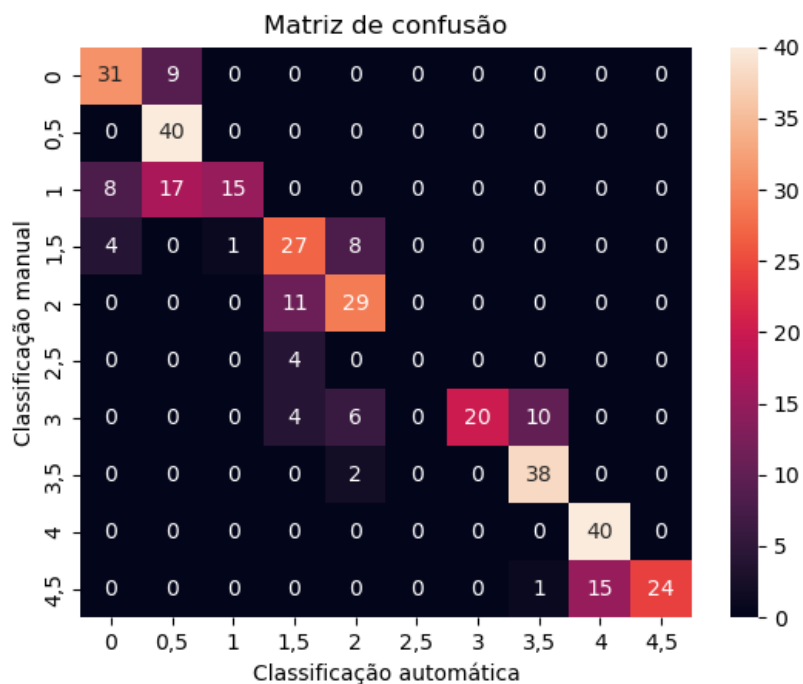


Figura 38 - Matriz de confusão após o quarto treinamento, comparando classificação manual e classificação automática (200 épocas).

Analisando o resultado do teste do quarto treinamento destacam-se os seguintes pontos:

- Os resultados são coerentes com a dificuldade encontrada pelos especialistas, onde se esbarra em uma diferença mínima na classificação de duas severidades vizinhas de inclusões.
- Os erros considerados inaceitáveis onde o resultado não condiz com a severidade da classificação manual e nem ao menos com as suas vizinhas ocorreram em menor frequência.
- Duas severidades apresentaram 100% de acurácia nos testes.
- As severidades 2,5 e 5,0 não foram consideradas no estudo por falta de um número de imagens suficiente para compor o banco de dados.

A Figura 39 mostra um gráfico comparativo da evolução dos resultados encontrados nos treinamentos da rede neural. O gráfico foi dividido em 3 categorias sendo elas: % Acertos, %Acertos parciais e % Erro. Os acertos se configuram nas

imagens que apresentaram a mesma severidade da análise manual. Já os acertos parciais configuram-se em resultados vizinhos a severidade da análise manual por exemplo: Análise manual = 2,0, as imagens classificadas a 1,5 e 2,5 são consideradas acertos parciais pela pequena diferença entre algumas imagens e a % Erro configura-se em todas as classificações que não se encaixam nas demais categorias citadas acima ou seja, severidades que não foram acertos e acertos parciais.

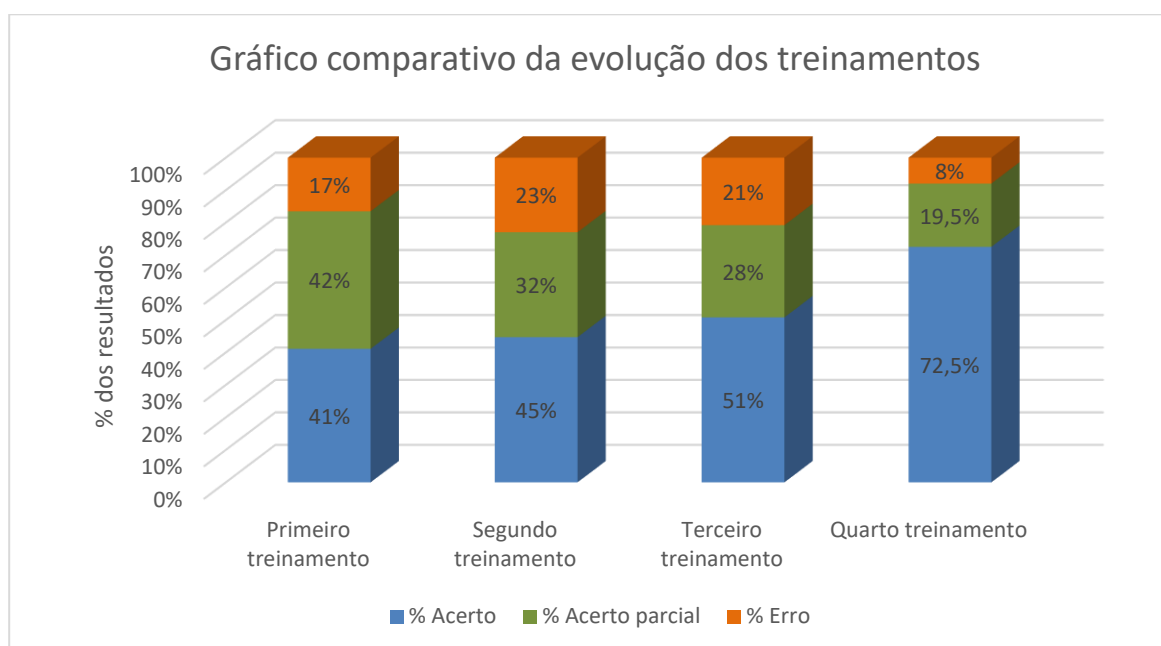


Figura 39 - Gráfico comparativo da evolução dos resultados dos treinamentos da rede neural

5 CONCLUSÕES

- A qualidade das imagens utilizadas para o treinamento da rede neural influencia diretamente no resultado encontrado, por esse motivo os retrabalhos de preparação e captura de fotos foram as atividades que mais exigiram atenção, empenho, dedicação e habilidade.
- A quantidade de fotos no banco de dados influencia no resultado do aprendizado da rede neural. Em severidades que não possuem grande quantidade de fotos, o treinamento e a aprendizagem se tornam limitados o que pode resultar na baixa acurácia de validação e teste. Já no caso de um banco de dados grande e representativo, maior será o mapeamento das diversas morfologias/aspectos de inclusões e, menor será a chance da rede neural encontrar uma característica a qual ela não foi treinada a identificar.
- A rede neural se mostra uma ferramenta em potencial para ser usada para treinamento pessoal, retenção de conhecimento, otimização do processo de classificação, estabelecimento de correlação entre processo e produto.
- Os resultados da rede neural quando comparados com os resultados dos técnicos metalógrafos se mostraram eficazes e precisos em algumas severidades, chegando a apresentar 100% de acurácia. No entanto, a linha tênue de classificação entre duas severidades vizinhas também foi um desafio para a rede neural assim como também é um desafio para os especialistas. Com isso, o resultado de acertos precisos diminuiu para algumas severidades.
- Alguns fatores que podem ter influenciado negativamente na aprendizagem da rede neural: diferente quantidade de imagens por severidade, diferentes parâmetros de iluminação na captação das imagens, classificação manual realizada por mais de uma pessoa e artefatos da preparação das amostras metalográficas.

- Os resultados deste trabalho de Mestrado abrem perspectiva para o desenvolvimento de novos estudos utilizando “*machine learning*” em processos produtivos no segmento metalúrgico. Podem ser aplicadas redes neurais para a classificação de microestruturas, fases, distribuição de carbonetos, e para a classificação de inclusões correlacionando-as com a composição química do aço.

6 SUGESTÕES PARA TRABALHOS FUTUROS

- Ampliar o estudo para diferentes tipos de inclusões como Alumina, silicato e óxido. Dessa forma, podendo aplicar o modelo de rede neural para um maior conjunto de aços.
- Comparar os resultados obtidos utilizando rede neural com os resultados de outras ferramentas de análise de imagem como o software ImageJ.
- Estudar a junção de ferramentas como o ImageJ com a rede neural, podendo unir a otimização da ferramenta com a inteligência artificial da rede neural.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] R. Kiessling, Non-metallic inclusions in steel, **London: Iron and Steel Institute**, v.1, p.1 – 497, 1968.
- [2] H. Tervo, A. Kaijalainen, T. Pikkarainen, S. Mehtonen e D. Porter, Effect of impurity level and inclusions on the ductility and toughness of an ultra-high-strength steel, **Materials Science & Engineering A**, v.697, p.184 – 193, 2017.
- [3] N. Rieders, M. Nandasiri, D. Mogk e R. Avci, “New Insights into Sulfide Inclusions in 1018 Carbon Steels,” **MDPI**,v.11, p. 428, 2021.
- [4] W. Shi, S. Yang, A. Dong e J. Li, “Understanding the corrosion mechanism of spring steel induced by MnS inclusions with different sizes,” **The minerals, metals & materials society**,v70, p. 2513-2522, 2018.
- [5] A. Bobsin, Estudo da morfologia, tamanho, distribuição e precipitação sobre óxidos de inclusões de sulfetos de manganês em aços resulfurados SAE 1141 acalmados ao, Porto Alegre, 2019.
- [6] A. E. D., Hanum, A. L. Aços de corte fácil de baixo carbono. **Metalurgia - ABM**, v. 33, n. 241, p.773-776, 1977.
- [7] B. A. L. de Brito. Aspectos metalúrgicos na avaliação da usinabilidade de aços. **Rem: Revista Escola de Minas**, v. 55, n. 2, p. 103–109, 2002.
- [8] ASTM E45-18a, Standard Test Methods for Determining the Inclusion Content of Steel, ASTM International, **West Conshohocken**, PA, 2018, www.astm.org.
- [9] N. Neogi, D. K. Mohanta e P. K. Dutta, Review of vision-based steel surface inspection systems, **J Image Video Proc**,v. 50, 2014.

[10] J. P. Yun, S. Choi, Y.-j. Jeon, D.-c. Choi e S. W. Kim, Detection of line defects in steel billets using Undecimated wavelet transform, **Pohang: International conference on control, automation and systems 2008**, p.1725-1728, 2018.

[11] U. Ç. S. Ekici, "Support vector machines models for surface roughness prediction in CNC turning of AISI 304 austenitic stainless steel," **Journal of Intelligent Manufacturing**, v. 23, p.639-650, 2010.

[12] S. Y. H. R. S. D. J. E. B. H. Graham Roberts, "Deep learning for semantic segmentation of defects in advanced STEM images of steels," **Scientific reports**, v. 9, 2019.

[13] F. Liberman, Classificação de imagens digitais por textura usando redes neurais, Porto Alegre, 1997.

[14] J. A. T. Fernandes e C. R. d. Rosa, O clima organizacional: Um conceito motivador para redução do turnover, São Paulo.

[15] M. Cuartas, Machine learning algorithms for the prediction of non-metallic inclusions in steel wires for tire reinforcement, **Journal of Intelligent Manufacturing**. v. 32, 2021.

[16] M. Peet, H. Hasan e H. Bhadeshia, **Published in the International Journal of Heat and Mass Transfer**, v. 54, p. 2602-2608, 2011.

[17] PUC-RIO- Certificação Digital N° 0025043/CA, "História do aço," Rio de Janeiro.

[18] "World steel association," 2021. [Online]. Available: https://worldsteel.org/steel-by-topic/statistics/annual-production-steel-data/P1_crude_steel_total_pub/WORLD_ALL. [Acesso em 07 04 2022].

- [19] “World Steel Association,” World Steel, abril 2022. [Online]. Available: <https://worldsteel.org/steel-by-topic/statistics/steel-data-viewer/>. [Acesso em 01 Junho 2022].
- [20] Companhia Souza Cruz, “A importância do aço,” Conjuntura setorial, p. 131-132, 1984.
- [21] H. Colpaert, Metalografia dos produtos siderúrgicos comuns, São Paulo: Editora Blucher, 2008.
- [22] Society of Automotive Engineers. Chemical Compositions of SAE Alloy Steels. Warrendale; PA: SAE, 2009.
- [23] R. F. Moreno, Estudo das variáveis que afetam a tenacidade e a ductilidade do aço PH13-8Mo utilizado na indústria aeronáutica, São Carlos, 2013.
- [24] A. L. V. d. C. e. Silva, The effects of non-metallic inclusions on properties relevant to the performance of steel in structural and mechanical applications, **Journal of Materials Research and Technology**, v. 8, 2019.
- [25] G. Wranglen, “Pitting and sulfide inclusions in steel.,” **Corrosion Science**, v. 14, p. 331-349, 1974.
- [26] K. Hashimoto, TakeshiFujimatsu, N. Tsunekage, K. Hiraoka, K. Kida e E. CostaSantos, “Study of rolling contact fatigue of bearing steels in relation to various oxide inclusions,” **Elsevier**, v. 7, p. 102612, 2011.
- [27] Y. E. K. L. S. M. M. L. Z. Pooya Alavian, The (α, β) -Precise Estimates of MTBF and MTTR: Definition, Calculation, and Observation Time, **IEEE Transactions on automation science and engineering**, v.18, p. 1469-1477, 2020.

[28] DIN 50602, Microscopic examination of special steels using standard diagrams to assess the content of non-metallic inclusions, **DEUTSCHE NORM**, Berlin, 1985.

[29] H.S. MANSUR, “Técnicas de Caracterização de Materiais,” Biomateriais , Belo Horizonte, 2012.

[30] M. C. J. Murta, Análise de inclusões não metálicas em aço por espectrometria de emissão óptica associada ao sistema Spark-DAT, Belo Horizonte, 2013.

[31] D. M. I. T. D.J. Livingstone, Data modelling with neural networks: Advantages and limitations, **Portsmouth: Journal of Computer-Aided Molecular Design**, v. 11, p.135-142, 1996.

[32] S. Haykin, Neural networks and learning machines, Hamilton, Ontario, Canada: Pearson prentice hall, 2009.

[33] C. Ozcan, (26 de July de 2022). Fonte: ozeninc: <https://blog.ozeninc.com/resources/where-is-step-function-in-ansys-mechanical>.
[Acesso em 30 07 2022]

[34] F. K. M. S. Başarslana. Sentiment Analysis with Machine . ADCAIJ: **Advances in Distributed Computing and Artificial Intelligence Journal**, v. 9, p. 5-15, 2020.

[35] E. L. S. d. Silva, Aplicação das redes neurais no reconhecimento de descontinuidade em cordões de solda, Rio de Janeiro, Universidade Federal do Rio de Janeiro, 2021.

[36] S.-Ali-Ahmadi. (05 de Jun de 2022).Fonte:Github:https://github.com/Seyed-Ali-Ahmadi/Digital-ImageProcessing/blob/master/DIP_course_6th_session_Spatial_Filtering.ipynb.
[Acesso em 30 07 2022]

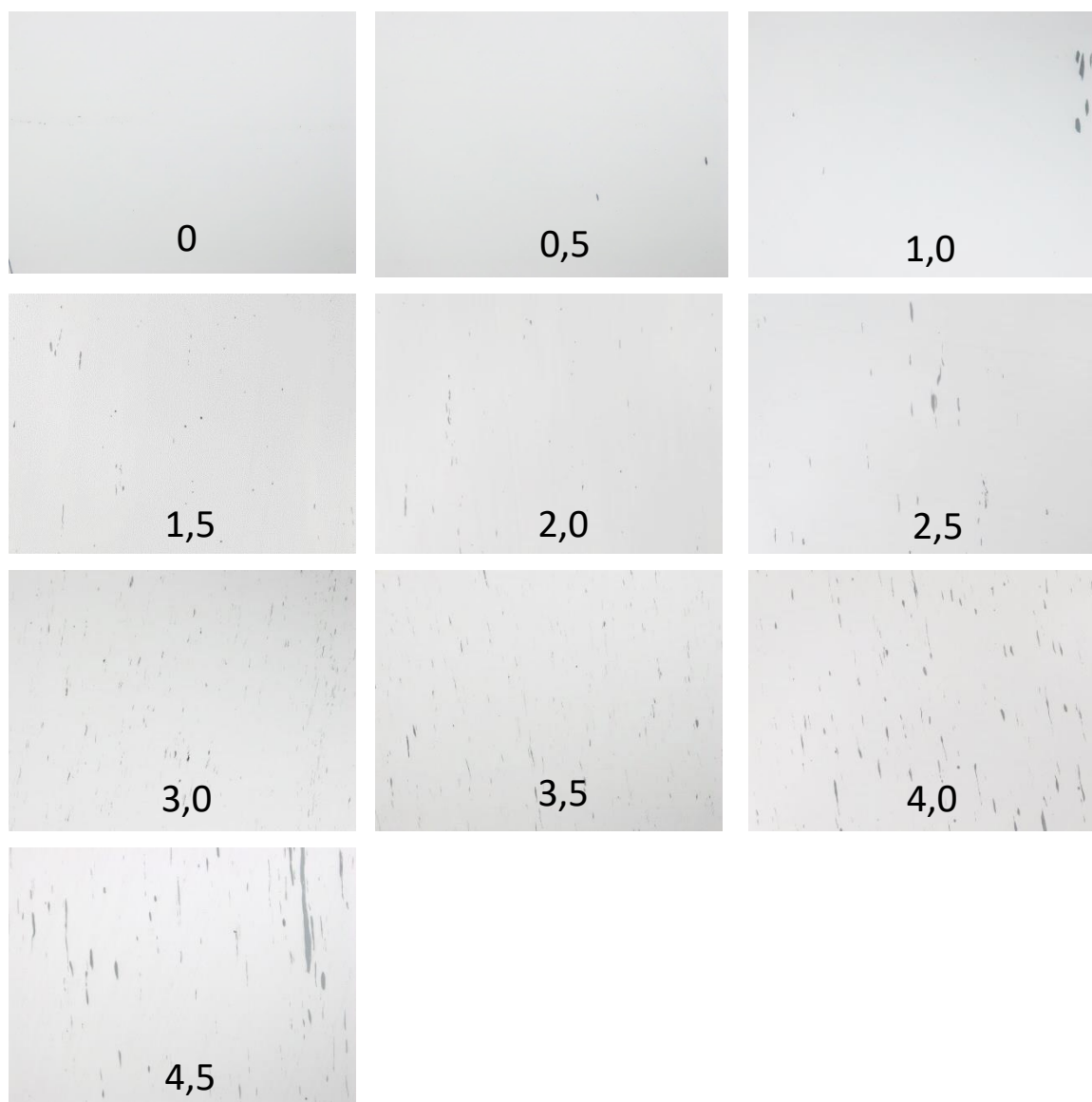
- [37] FirelordPhoenix. (26 de February de 2018). Fonte: paperswithcode: <https://paperswithcode.com/method/max-pooling>. [Acesso em 30 07 2022]
- [38] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, & R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research** **15**, v.15, p. 1929-1958, 2014.
- [39] V. K. C. R. K. E. J. S. A. C. H. S. S. Ignacio Arganda-Carreras, "Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification," **Bioimage informatics**, v. 33, p.2424-2426, 2017.
- [40] S. K. Suresh, Machine Learning Based Ultra High Carbon Steel Image, University of South Carolina, 2019.
- [41] J. A. d. C. L. A. F. d. C. Wanessa Pereira dos Reis, "MODELO DE PREVISÃO DE DEMANDA DE FERROLIGAS," junho 2019.
- [42] R. A. Rohde, Metalografia preparação de amostras, **LEMM Laboratório de Ensaio Mecânicos e Materiais**, v.3, 2010.
- [43] J. William D. Callister, Fundamentos da ciência e engenharia de materiais, Rio de Janeiro: LTC, 2010
- [44] K. A. R. V. V. U. L. S. Alan Santos, "Uma Abordagem de Classificação de Imagens Dermatoscópicas Utilizando Aprendizado Profundo com Redes Neurais Convolucionais," **XXXVII Congresso da Sociedade Brasileira de Computação**, 2010.
- [45] A. S. P. D. Farhana Sultana, "Advancements in Image Classification Using Convolutional Neural Network," **2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)**, p. 122-129, 2018.

[46] I. V. Rizzo; R. L. C. CANATO, Inteligência Artificial: Funções de ativação, **Prospectus**; v. 2, p. 51-65, 2020.

[47] G. H. A. K. I. S. R. S. Nitish Shivastava, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," **Journal of Machine Learning Research** **15**, v.15, p. 1929-1958, 2014.

**APÊNDICE A – EXEMPLOS DE IMAGENS ÓPTICAS QUE COMPÕEM O
BANCO DE DADOS**

SEVERIDADES



APÊNDICE B – CÓDIGOS DE PROGRAMAÇÃO UTILIZADOS

Código para ajuste e configuração das imagens

```
from PIL import Image
import matplotlib.pyplot as plt
import os

def adjust_images():
    __loc__ =
os.path.realpath(os.path.join(os.getcwd(),os.path.dirname(__file__)))
    for root, dirs, files in list(os.walk(__loc__)):
        for f in files:
            if(f.endswith('.jpg')):
                filepath = os.path.join(root,f)
                print(f)
                img = Image.open(filepath)
                img2 = img.resize((640,480),Image.ANTIALIAS)
                img2.save(filepath)

if __name__ == "__main__":
    adjust_images()
```

Código utilizado para aumentar o número de imagem girando as imagens em 90°

```
import os
import cv2
import matplotlib.pyplot as plt
if __name__ == "__main__":
    __loc__ =
os.path.realpath(os.path.join(os.getcwd(),os.path.dirname(__file__)))
    #folderpath = os.path.join(__loc__, 'Pierre - Fotos')

    for root, dirs, files in os.walk(__loc__):
        for f in list(files):
            if(f.endswith('.jpg')):
                filepath = os.path.join(root,f)
                print(f)
                image = cv2.imread(filepath)
                flipped_horizontally = cv2.flip(image,1)
                flipped_vertically = cv2.flip(image,0)
                flipped_horizontallyvertically = cv2.flip(image,-1)

                cv2.imwrite(os.path.join(root,f) + '_flip1.jpg'
,flipped_horizontally)
                cv2.imwrite(os.path.join(root,f) + '_flip2.jpg'
,flipped_vertically)
                cv2.imwrite(os.path.join(root,f) + '_flip3.jpg'
,flipped_horizontallyvertically)
```

Código utilizado para gerar os resultados

```
import os
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

__loc__ = os.path.realpath(os.path.join(os.getcwd(),os.path.dirname(__file__)))

def plot_history(history):
    acc = history['accuracy']
    val_acc = history['val_accuracy']

    loss=history['loss']
    val_loss=history['val_loss']

    epochs_range = range(len(acc))
    '''
    logdir = __loc__
    tensorboard_callback = keras.callbacks.TensorBoard(log_dir = logdir,
    histogram_freq = 1)
    file_writer_cm = tf.summary.create_file_writer(logdir + '/cm')
    '''

    #saver.save(sess,os.path.join(__loc__, 'Model'))

    plt.figure(figsize=(8, 8))
    plt.subplot(1, 2, 1)
    plt.plot(epochs_range, acc, label='Acurácia de Treinamento')
    plt.plot(epochs_range, val_acc, label='Acurácia de Validação')
    plt.legend(loc='lower right')
    plt.xlabel('Época')
    plt.title('Acurácia')

    plt.subplot(1, 2, 2)
    plt.plot(epochs_range, loss, label='Função de Perda (Treinamento)')
    plt.plot(epochs_range, val_loss, label='Função de Perda (Validação)')
    plt.legend(loc='upper right')
    plt.xlabel('Época')
    plt.title('Função de Perda')
    plt.show()

if __name__ == '__main__':
    history=np.load(os.path.join(__loc__, 'Model_4x_1a_vez..numpy'),allow_pickle='TRUE
    ').item()
    plot_history(history)
```

Código da rede neural que foi utilizado para treinamento, validação e teste

```
import os
import matplotlib.pyplot as plt
import numpy as np
import keras
import tensorflow as tf
from keras.preprocessing.image import ImageDataGenerator
from keras.models import Sequential
from keras.layers.core import Dense, Flatten, Dropout
from keras.layers.convolutional import Conv2D, MaxPooling2D
#from keras.layers import
Conv2D,Dense,Input,InputLayer,Dropout,Flatten,MaxPool2D,BatchNormalization
from keras.optimizers import Adam, SGD
#from keras.layers.normalization import BatchNormalization deprecated
from tensorflow.keras.layers import BatchNormalization
from sklearn.model_selection import train_test_split
import sklearn
import cv2
import itertools
'''
def plot_to_image(figure):
    """
    Converts the matplotlib plot specified by 'figure' to a PNG image and
    returns it. The supplied figure is closed and inaccessible after this call.
    """

    buf = io.BytesIO()

    # Use plt.savefig to save the plot to a PNG in memory.
    plt.savefig(buf, format='png')

    # Closing the figure prevents it from being displayed directly inside
    # the notebook.
    plt.close(figure)
    buf.seek(0)

    # Use tf.image.decode_png to convert the PNG buffer
    # to a TF image. Make sure you use 4 channels.
    image = tf.image.decode_png(buf.getvalue(), channels=4)

    # Use tf.expand_dims to add the batch dimension
    image = tf.expand_dims(image, 0)

    return image

def plot_confusion_matrix(cm, class_names):
    """
    Returns a matplotlib figure containing the plotted confusion matrix.

    Args:
        cm (array, shape = [n, n]): a confusion matrix of integer classes
        class_names (array, shape = [n]): String names of the integer classes
    """

    figure = plt.figure(figsize=(8, 8))
    plt.imshow(cm, interpolation='nearest', cmap=plt.cm.Blues)
```

```

plt.title("Confusion matrix")
plt.colorbar()
tick_marks = np.arange(len(class_names))
plt.xticks(tick_marks, class_names, rotation=45)
plt.yticks(tick_marks, class_names)

# Normalize the confusion matrix.
cm = np.around(cm.astype('float') / cm.sum(axis=1)[:, np.newaxis],
decimals=2)

# Use white text if squares are dark; otherwise black.
threshold = cm.max() / 2.

for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
    color = "white" if cm[i, j] > threshold else "black"
    plt.text(j, i, cm[i, j], horizontalalignment="center", color=color)

plt.tight_layout()
plt.ylabel('True label')
plt.xlabel('Predicted label')
return figure

def log_confusion_matrix(epoch, logs):

    # Use the model to predict the values from the test_images.
    test_pred_raw = model.predict(test_images)

    test_pred = np.argmax(test_pred_raw, axis=1)

    # Calculate the confusion matrix using sklearn.metrics
    cm = sklearn.metrics.confusion_matrix(test_labels, test_pred)

    figure = plot_confusion_matrix(cm, class_names=class_names)
    cm_image = plot_to_image(figure)

    # Log the confusion matrix as an image summary.
    with file_writer_cm.as_default():
        tf.summary.image("Confusion Matrix", cm_image, step=epoch)
    ...

if __name__ == "__main__":
    __loc__ =
os.path.realpath(os.path.join(os.getcwd(),os.path.dirname(__file__)))

    train_dir = os.path.join(__loc__, 'Training')
    batch_size = 16
    epochs = 200
    IMG_HEIGHT = 480
    IMG_WIDTH = 640
    number_of_classifications = 10

    train_datagen = ImageDataGenerator(rescale=1./255)
    train_generator = train_datagen.flow_from_directory(
        directory=os.path.join(train_dir, 'Train'),
        target_size=(IMG_HEIGHT, IMG_WIDTH),
        color_mode='grayscale',
        batch_size=batch_size,

```



```

        class_mode='categorical',
        shuffle=True,
        seed=42
    )
    validation_datagen = ImageDataGenerator(rescale=1./255)
    validation_generator = validation_datagen.flow_from_directory(
        directory=os.path.join(train_dir,'Valid'),
        target_size=(IMG_HEIGHT,IMG_WIDTH),
        color_mode='grayscale',
        batch_size=batch_size,
        class_mode='categorical',
        shuffle=True,
        seed=42
    )
    tf.compat.v1.disable_eager_execution()
    print("GPUs : " +
str(len(tf.compat.v1.config.list_physical_devices('GPU'))))

    with tf.compat.v1.Session() as sess:

        STEP_SIZE_TRAIN = train_generator.n//train_generator.batch_size
        STEP_SIZE_VALID =
validation_generator.n//validation_generator.batch_size
        print('Building Neural Network...')
        model = Sequential()

#model.add(Dense(50,input_shape=[IMG_HEIGHT,IMG_WIDTH,1],activation='sigmoid'))
        #model.add(Dense(2,activation='softmax'))
        #model.add(InputLayer(input_shape=[IMG_HEIGHT,IMG_WIDTH,1]))

#model.add(Conv2D(filters=32,kernel_size=5,strides=1,padding='same',activation=
'relu'))
        #model.add(BatchNormalization())

#model.add(Conv2D(input_shape=[IMG_HEIGHT,IMG_WIDTH,1],filters=32,kernel_size=5
,strides=1,padding='same',activation='relu'))

#model.add(Conv2D(input_shape=[IMG_HEIGHT,IMG_WIDTH,1],filters=32,kernel_size=(
3,3),activation='relu'))

model.add(Conv2D(32,input_shape=[IMG_HEIGHT,IMG_WIDTH,1],kernel_size=(3,3),acti
vation='relu'))
        #model.add(MaxPool2D(pool_size=5,padding='same'))
        model.add(MaxPooling2D(pool_size=(2,2)))
        #model.add(Dropout(rate=0.25))
        model.add(BatchNormalization())
        model.add(Conv2D(64,kernel_size=(3,3),activation='relu'))
        model.add(MaxPooling2D(pool_size=(2,2)))
        model.add(Dropout(0.25))
        model.add(BatchNormalization())
        model.add(Conv2D(128,kernel_size=(3,3),activation='relu'))
        model.add(MaxPooling2D(pool_size=(2,2)))
        model.add(Dropout(0.25))
        #model.add(BatchNormalization())

#model.add(Conv2D(filters=50,kernel_size=5,strides=1,padding='same',activation=

```

```

'relu'))
    #model.add(MaxPool2D(pool_size=5,padding='same'))
    #model.add(BatchNormalization())

#model.add(Conv2D(filters=80,kernel_size=5, strides=1,padding='same',activation=
'relu'))
    #model.add(MaxPool2D(pool_size=5,padding='same'))
    #model.add(Dropout(0.25))

    model.add(Flatten())
    model.add(BatchNormalization())
    model.add(Dense(256,activation='relu'))
    model.add(Dropout(rate=0.4))
    model.add(BatchNormalization())
    model.add(Dense(120,activation='relu'))
    model.add(Dropout(rate=0.4))
    model.add(BatchNormalization())
    model.add(Dense(120,activation='sigmoid'))
    model.add(Dropout(rate=0.4))
    model.add(BatchNormalization())
    model.add(Dropout(rate=0.5))
    #model.add(BatchNormalization())
    model.add(Dense(number_of_classifications,activation='softmax'))
    #optimizer = Adam(lr=1e-2)
    optimizer = Adam(lr=1e-4)
    #optimizer = SGD(lr=0.00001)
    #optimizer = SGD(lr=0.0001, decay=1e-2, momentum=0.9)
    #optimizer = SGD(lr=0.01, decay=0.0, momentum=0.0,nesterov=False)
    print('Saver...')
    saver = tf.compat.v1.train.Saver()
    sess.run(tf.compat.v1.global_variables_initializer())

model.compile(optimizer=optimizer,loss='categorical_crossentropy',metrics=['acc
uracy'])

model.evaluate_generator(generator=train_generator,steps=STEP_SIZE_TRAIN)

    model.summary()
    history = model.fit_generator(
        train_generator,
        steps_per_epoch=STEP_SIZE_TRAIN,
        validation_data=validation_generator,
        validation_steps=STEP_SIZE_VALID,
        epochs=epochs
    )
    print('hey')

    ...
    logdir = __loc__
    tensorboard_callback = keras.callbacks.TensorBoard(log_dir = logdir,
histogram_freq = 1)
    file_writer_cm = tf.summary.create_file_writer(logdir + '/cm')
    ...

    model.save('Model_4x_1a_vez.h5')
    np.save('Model_4x_1a_vez..np',history.history)

```

```
#saver.save(sess,os.path.join(__loc__,'Model'))
'''
acc = history.history['acc']
val_acc = history.history['val_acc']

loss=history.history['loss']
val_loss=history.history['val_loss']

epochs_range = range(epochs)
plt.figure(figsize=(8, 8))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, acc, label='Acurácia de Treinamento')
plt.plot(epochs_range, val_acc, label='Acurácia de Validação')
plt.legend(loc='lower right')
plt.xlabel('Época')
plt.title('Acurácia')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, loss, label='Função de Perda (Treinamento)')
plt.plot(epochs_range, val_loss, label='Função de Perda (Validação)')
plt.legend(loc='upper right')
plt.xlabel('Época')
plt.title('Função de Perda')
plt.show()
'''
```