



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2020년 2월
석사학위논문

기계학습을 이용한 Al-Si 주조 합금의 미세조직 분류 및 분석

조선대학교 대학원

첨단소재공학과

정 상 준

기계학습을 이용한 Al-Si 주조 합금의 미세조직 분류 및 분석

Classification and Analysis of Microstructure of Al-Si
Casting Alloy Using Machine Learning

2020년 2월 25일

조선대학교 대학원

첨단소재공학과

정 상 준

기계학습을 이용한 Al-Si 주조 합금의 미세조직 분류 및 분석

지도교수 김 희 수

이 논문을 공학 석사학위신청 논문으로 제출함.

2019년 10월

조선대학교 대학원

첨단소재공학과

정 상 준

정상준의 석사학위 논문을 인준함

위원장 조선대학교 교수 김선중 (인)

위 원 조선대학교 교수 장희진 (인)

위 원 조선대학교 교수 김희수 (인)

2019년 11월

조선대학교 대학원

목 차

| | |
|-----------------------|-----|
| LIST OF TABLES | iii |
| LIST OF FIGURES | iv |
| ABSTRACT | vii |
| | |
| 제 1 장 서 론 | 1 |
| | |
| 제 2 장 이론적 배경 | 3 |
| 제 1 절 관련연구 | 3 |
| 1. 인공신경망 | 3 |
| 2. 딥 러닝 | 6 |
| 3. 합성곱 신경망 | 7 |
| 3.1 LeNet | 13 |
| 3.2 AlexNet | 14 |
| 3.3 VGGNet | 15 |
| 3.4 GoogleNet | 16 |
| 4. AI-Si alloy | 18 |

| | |
|--|----|
| 제 3 장 실험방법 | 20 |
| 1. Al-Si 주조시편 제작 | 20 |
| 2. 이미지 전처리 | 22 |
| 3. 모델 선정 및 알고리즘 설계 | 23 |
| | |
| 제 4 장 결과 및 고찰 | 26 |
| 4.1 기계학습을 이용한 합금성분의 역산출 | 26 |
| 4.2 미세조직 이미지의 원본과 변환된 이미지 유사도 비교 | 34 |
| 4.3 데이터 셋 범위를 벗어난 조성의 시편 분석 | 45 |
| 4.4 신경망 모델 비교 | 47 |
| | |
| 제 5 장 결 론 | 48 |
| | |
| 참 고 문 헌 | 49 |

LIST OF TABLES

| | |
|------------------------|----|
| Table 2.1 VGGNet | 15 |
|------------------------|----|

LIST OF FIGURES

| | |
|---|----|
| Fig. 2.1 Neuron | 4 |
| Fig. 2.2 Schematic of Perceptron | 4 |
| Fig. 2.3 incipient Convolutional Neural Network (LeNet) | 7 |
| Fig. 2.4 Neural Network, Convolutional Neural Network | 8 |
| Fig. 2.5 Padding | 9 |
| Fig. 2.6 Stride | 10 |
| Fig. 2.7 Pooling | 11 |
| Fig. 2.8 Batch Normalization | 12 |
| Fig. 2.9 Overfitting | 12 |
| Fig. 2.10 AlexNet | 14 |
| Fig. 2.11 GoogleNet | 16 |
| Fig. 2.12 Module structure of inception | 16 |
| Fig. 2.13 Mini-network replacing 5x5 | 17 |
| Fig. 2.14 Al-Si phase Diagram | 19 |

Fig. 3.1 Al-Si microstructure 20

Fig. 3.2 Microstructure Image Preprocessing 22

Fig. 3.3 Experimental schematic 24

Fig. 4.1 Similarity distribution evolved with training step for (a) Al-1 wt%Si, (b) Al-4 wt%Si, (c) Al-7wt%Si, and (d) Al-10 wt%Si alloys. 27

Fig. 4.2 Si content where the highest peak existed, as a function of training step. 32

Fig. 4.3 Weight-averaged Si content of the alloys. These values approached to the accurate value as train step increased. 33

Fig. 4.4 Similarity distribution evolved with training step for (a) preprocessing + reverse, (b) preprocessing + reverse + half, (c) preprocessing + half Al-1wt%Si alloys. 35

Fig. 4.5 Similarity distribution evolved with training step for (a) preprocessing + reverse, (b) preprocessing + reverse + half, (c) preprocessing + half ,Al-7wt%Si alloys.38

Fig. 4.6 Comparison of original and converted images for (a) Al-1wt%si, (b) Al-7wt%si alloys. (A - original, B - preprocessing, C - preprocessing + reverse, D - preprocessing + reverse + half, E - preprocessing + half) 41

Fig. 4.7 Accuracy comparison with or without scale bar for (a) Al-4 wt%Si,(b) Al-6 wt%Si alloys. 43

Fig. 4.8 accuracy distributions of the samples with 2 and 9 wt%Si, when the dataset has the range of 46

Fig. 4.9 Comparison of Inception Model, LeNet Model 47

Abstract

Classification and Analysis of Microstructure of Al-Si Casting Casting Alloy Using Machine Learning

Sang-Jun Jeong

Advisor: Prof. Hee-soo Kim, Ph. D.

Dept. of Advanced Materials Engineering

Graduate School of Chosun University

In this study, we analyzed the chemical composition of Al-Si cast alloys from microstructure images, using image recognition and machine learning. Binary Al-Si alloys of Si = 1~10 wt% were cast and prepared as reference images in the dataset used for machine learning. Repeated training to relate the microstructure images to their chemical composition was carried out, for up to 10,000 steps, to increase the reliability of the analysis. The peaks of similarity existed in the dataset with chemical compositions corresponding to the known target composition. The heights of the peaks became higher and the distribution of similarity became sharper with further training steps. This means that the weighted average of the chemical composition approached the target composition with increasing training steps. The correctness of the analysis increased with training steps up to 10,000, then was saturated. It was found that the chemical composition outside the dataset range could not be analyzed correctly. Analysis of the compositions between the datasets showed incorrect but reasonable results. The reliability of the chemical composition analysis using machine learning and image recognition developed in this study will increase when a vast range of reference images are collected and verified.

제 1 장 서론

딥 러닝(Deep learning)은 기계학습(Machine learning)의 분야로 최근 많은 관심과 주목을 받고 있다. 과거 사람이 직접 데이터를 뽑고 그를 주축으로 오차율을 비교하였을 때에 비해 높은 성과를 보이기 때문이다. 비록 정확하고 많은 데이터베이스가 필요하지만 전보다 오차율과 정확도가 많이 향상되었기에 각 분야에 응용하기 위한 개발이 진행되고 있다.

4차 산업혁명으로 의학, 농업, 등 많은 분야에 응용되고 있지만 대표적인 성공 사례로 스마트 팩토리(smart factory)가 있다. 스마트 팩토리란 설계와 개발, 제조 및 유통, 물류 등 생산하는 과정에 디지털 자동화 솔루션을 결합한 정보통신기술(ICT)을 적용하여 품질과 생산성, 고객만족도 등을 향상시키는 지능형 시공장을 뜻한다. 자동화는 설계부터 클라우드를 통한 데이터의 수집과 축적, 정보 추출 및 정보 분석 및 해석을 하며, 그 데이터들을 중심으로 학습하여 결함 등을 찾고 방지하며 제작하고 유통된다.[1]

Google이 신경망을 통해 이미지로부터 물체를 스스로 학습하고 분류하는 사례를 바탕으로 기계학습 프레임워크인 Tensorflow를 공개하였고, 이를 통하여 회사, 공장에서 주로 쌓이기만 하는 데이터들을 적절히 사용 및 관리하기 편하게 되었다. 딥 러닝을 통한 기계학습은 강화학습을 제외한 많은 알고리즘이 수많은 데이터베이스를 필요로 하기 때문에 빅 데이터를 통한 데이터베이스가 중요시 된다.

컴퓨터 비전은 인공지능의 일부로, 구축한 데이터베이스를 중심으로 사람이 인식하는 것처럼 영상, 혹은 사진의 특징을 추출하여 이해하는 것으로 감지, 인식, 분류, 위치파악이 주된 목표다. 주로 자율주행 자동차[3], 스마트폰의 얼굴 인식[15, 20]등 우리의 일상생활에 이미 여러 분야에서 접할 수 있다.

현재 각광받고 있는 컴퓨터 비전의 일종인 합성곱 신경망(Convolution

Neural Network, CNN)은 기계학습 모델 중 하나인 인공신경망(Artificial Neural Network, ANN)을 이용하여 이미지의 복잡한 특징들 중에서 중요한 부분을 추출하여 학습하는 것을 바탕으로 이루어진 모델로써 이미지를 분류 및 분석 할 수 있다. 합성곱 신경망을 토대로 발전한 VGG-16, ResNet, Inception 등 다양한 알고리즘들 중 성능과 속도, 적합성에 알맞은 Inception 모델을 선정하여 미세조직 사진의 분류 및 분석을 하고자 한다.

제 2 장 이론적 배경

제 1 절 관련연구

기계 학습은 데이터를 컴퓨터에 입력한 후, 분류 문제를 개발하고 다양한 데이터를 처리할 수 있는 기술로서, 데이터를 어떻게 분류할 것인지를 놓고 의사결정나무(Decision Tree)[9], 베이저안 망(Bayesian network)[16], 서포트 벡터 머신(support vector machine, SVM)[9,21,18], 인공신경망(Artificial neural network, ANN)[10]등 많은 알고리즘이 등장했다. 가장 접하기 쉬운 인공신경망(ANN)이 대표적으로 많이 쓰이고 있다. 인공신경망을 여러 층으로 쌓은 심층 신경망을 통해 기계 학습을 행하는 것이 딥 러닝이며, 지도 학습과 비지도 학습 등에 이용하게 된다.

1. 인공신경망(Artificial neural network, ANN)

인공 신경망은 인간의 뇌의 구조를 컴퓨터에 옮겨 놓는다면, 상당히 수준 높은 인공지능을 만들 수 있지 않을까 하는 발상에서 나온 신경망이다. 기계 학습을 이용하는 상위 신경망 중 하나로, 알파고로 유명세를 떨친 딥 러닝의 전신이다.

인공 신경망(ANN)에 앞서 짚고 넘어가야 할 것은 퍼셉트론(perceptron)이다. 인공 신경망은 ‘다층 신경망’ 과 ‘단층 신경망’ 으로 구분되며 단층 신경망을 일컬어 퍼셉트론이라 한다.

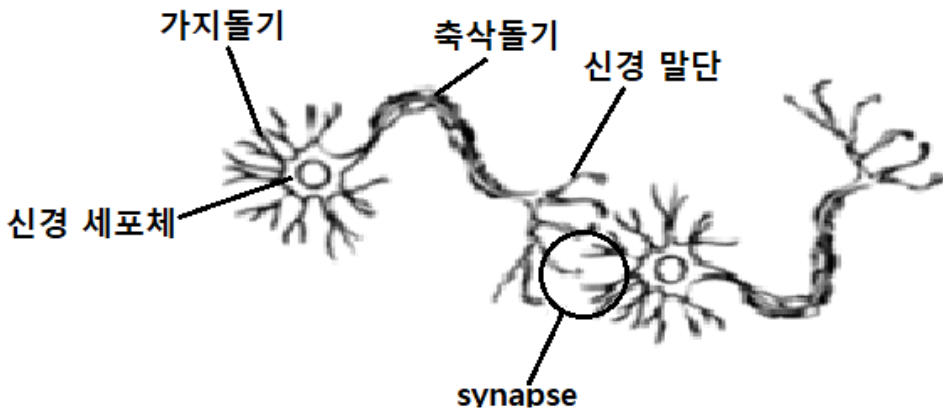


Fig. 2.1 Neuron

가지 돌기로 신호를 받은 뉴런이 신경 세포체에서 강도를 조절하고, 축삭 돌기에서 변환하여 신경 말단에서 출력한다. Fig. 2.1은 입력 값을 받고 가중치를 주고 편향을 준 뒤 출력되는 퍼셉트론을 도식화한 형태다.

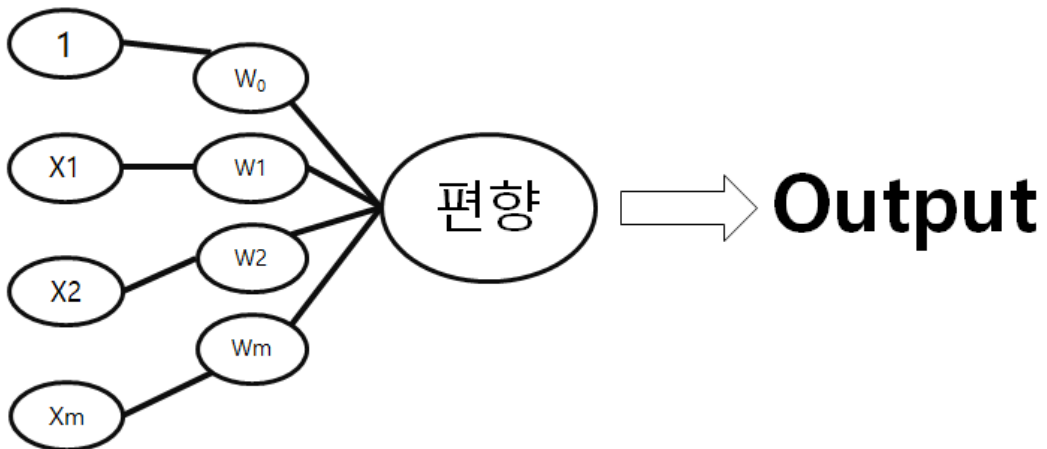


Fig. 2.2 Schematic of Perceptron

단층 퍼셉트론은 다수의 신호를 받아 하나의 신호를 출력하기에 뉴런과 유사하다. 퍼셉트론은 각 입력 신호의 세기에 따라 다른 가중치를 부여하는데 입력 신호의 합이 일정 값을 초과할 때 시냅스를 통하여 다른 뉴런으로 전달하게 되고 그렇게 다층 퍼셉트론이 되는 것이 인공 신경망이다. 인공 신경망의 각 층은 여러 개의 노드를 가질 수 있어 다양한 형태의 네트워크로 표현될 수 있다. [4, 12, 13, 19]

2. 딥 러닝(Deep Learning)

딥 러닝은 인공 신경망을 이용하여 데이터를 군집화 또는 분류하는데 사용하는 기계 학습의 일종이다. 인공 신경망이 특징 추출이 주가 되고 특징 추출과 분류기의 알고리즘이 독립된 형태라면 딥 러닝은 데이터로부터 자동으로 특징을 추출하고 분류기와 알고리즘이 통합되어 있다는 점이 장점이다. 이를 토대로 기존의 기계 학습에 응용하여 성능을 끌어올리게 된다.

딥 러닝의 핵심은 분류와 이를 통한 예측이다. 분류하는 방식은 지도 학습과 비지도 학습이 있고, 별개로 강화 학습이 있다. 지도 학습은 입력과 정답으로 학습 데이터를 사용하는 것으로 데이터 분류, 회귀 문제에 많이 사용되고 있다. 비지도 학습은 입력으로만 이루어진 학습 데이터를 사용하며, 분포의 추정과 군집화 작업에 대부분 사용되고 있다. 강화 학습이란, 보상 함수가 존재하여 어떤 행동을 취했을 때 최대한 보상을 많이 얻는 방향으로 학습하는 것이다. 강화 학습은 알파고의 모토로써 알파고가 인공지능이라면 딥 러닝은 알파고에게 배움을 주는 관계로 이 강화 학습이 사용되었다.

딥 러닝에는 많은 신경망이 존재하며 분야에 따라 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), DBN(Deep Belief Network) 등 다양한 학습 모델이 있다.[2,11,17]

3. 합성곱 신경망(Convolutional Neural Network, CNN)의 구조

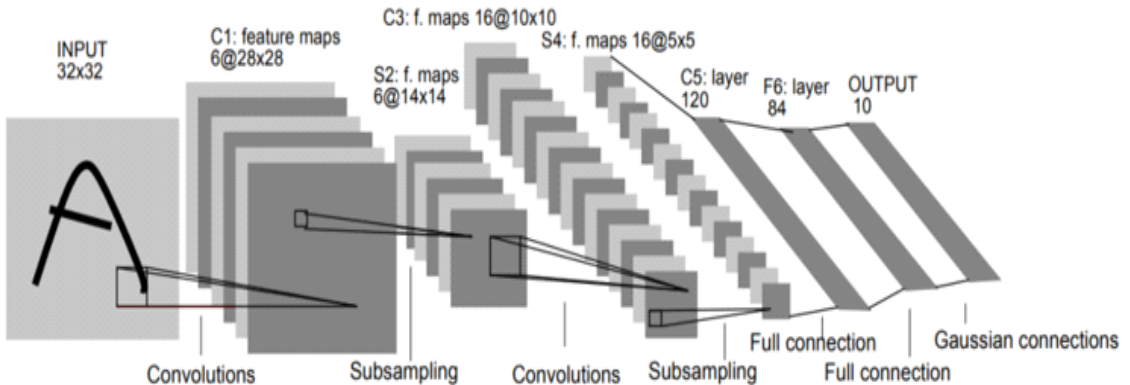


Fig. 2.3 incipient Convolutional Neural Network (LeNet)

Fig. 2.3 에서 보여주는 구조를 갖춘 기존 합성곱 신경망의 모델은 LeCun교수가 오래전 고안했던 CNN으로 LeNet 모델을 보편적인 네트워크로 하여 이미지 처리에 적합한 합성곱 신경망 모델이다. 입력 데이터를 이미지로 받고 합성곱 계층(convolution Layer)과 풀링(pooling) 계층을 거치며 계층이 깊어질수록 추출한 특징의 크기를 줄이고 커널의 수를 늘려가도록 하는 일반적인 구조다. 즉, 일반적인 완전연결 계층 신경망 모델에 여러 층의 합성곱 계층과 풀링 계층을 추가한 구조다.[7]

합성곱 계층은 전체 데이터에 가중치를 공유하기에 연결성이 좋아 합성곱 신경망의 특징이라 할 수 있다. 이미지 데이터의 경우 3차원 데이터로 가로 x 세로 x 채널 형상을 나타내고 완전 연결 계층에 입력할 때는 3차원 데이터를 1차원 데이터로 평탄화 한다. 합성곱 계층은 데이터의 형상을 유지하므로 이미지와 같은 형상을 가진 데이터를 제대로 이해할 가능성이 높다.

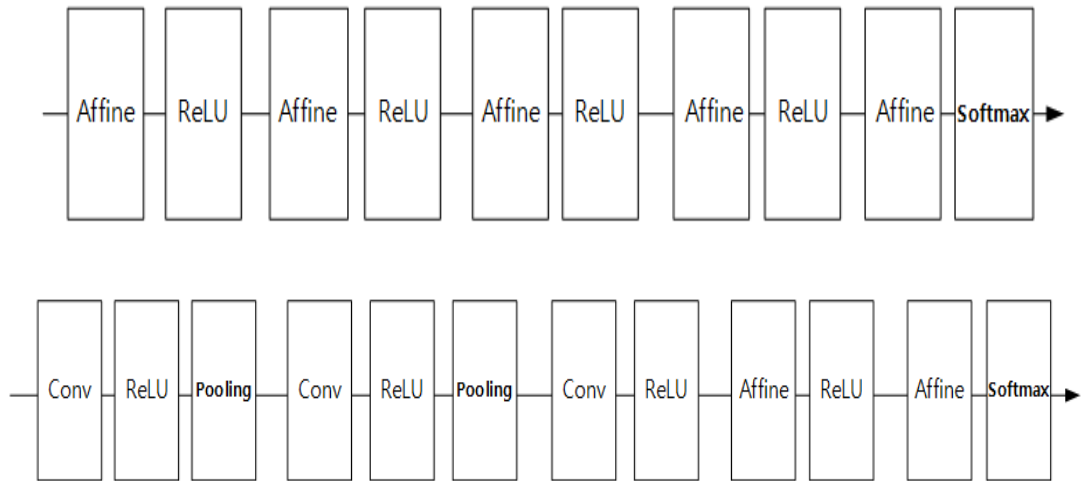


Fig. 2.4 Neural Network, Convolutional Neural Network

합성곱 신경망은 커널을 적용하여 분석하면서 파라미터의 수가 필터를 통해 줄어들어 복잡도가 단순해진다. 필터 적용의 이동위치 간격을 스트라이드(stride)라 하며 필터가 이동할 때 마다 줄어드는 출력 크기를 조절하기 위한 목적의 패딩(padding)이 있다. 패딩은 합성곱(convolution) 연산을 수행하기 전 입력 데이터 주변에 특정 값을 채워 크기를 조정한다.

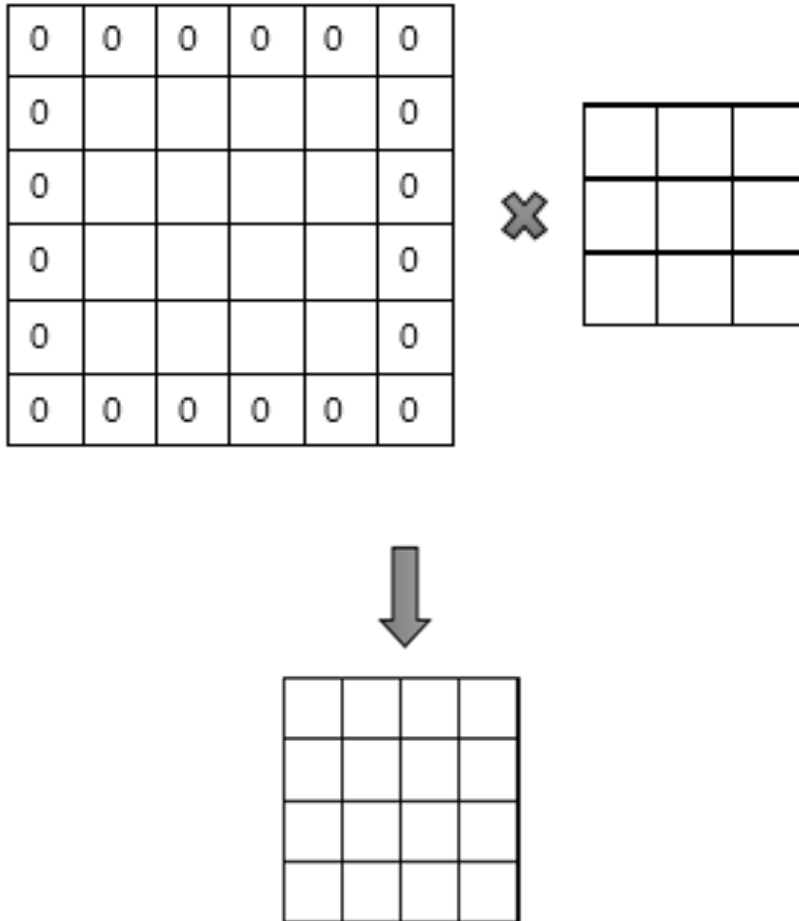


Fig. 2.5 Padding

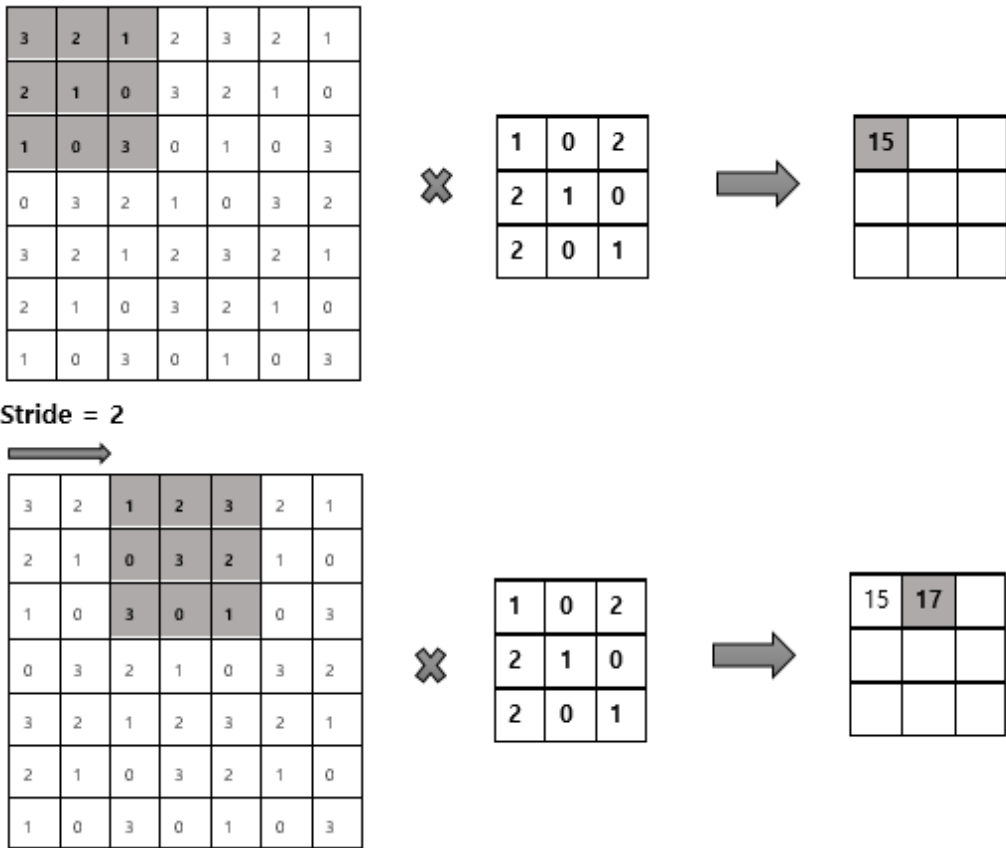


Fig. 2.6 Stride

마지막 단계인 풀링 계층은 서브 샘플링(sub-sampling)으로 합성곱 계층에서 만들어진 이미지의 차원을 축소하는 목적을 가지고 있다. 가로와 세로 방향의 공간을 줄이는 연산이며 학습해야 할 매개 변수가 없고 채널의 수가 변하지 않는다. 또한, 입력의 변화에 영향을 적게 받는 것이 특징이다. 풀링은 많아진 뉴런의 개수 때문에 과적합(overfitting)문제를 제어하기 위해 사용된다. 필터 안의 가장 큰 값을 추출하고 윈도우와 스트라이드가 같은 최댓값 풀링(max-pooling)과 평균값을 추출하는 평균값 풀링(average pooling), 최솟값을 추출하는 최솟값 풀링(min-pooling)이 있어 사용하는 모델에 맞게 사용한다.[5,8]

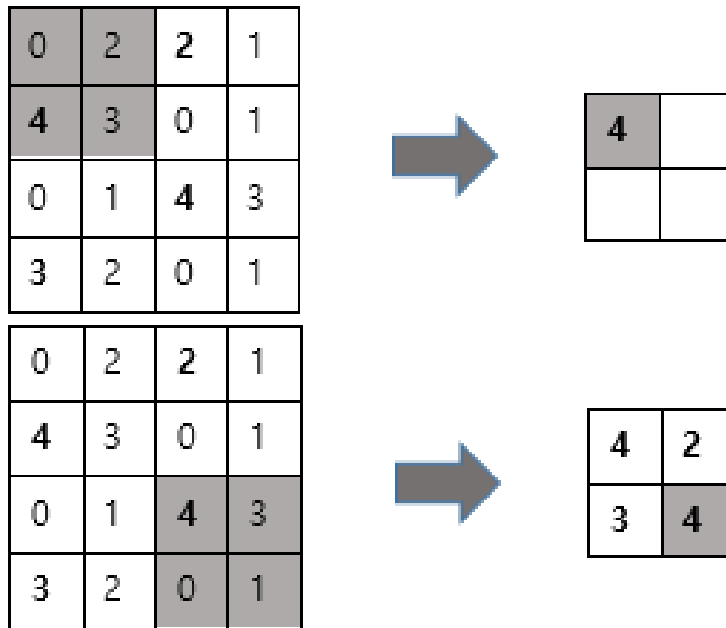


Fig. 2.7 Pooling

합성곱 신경망 모델을 이용하여 흔히 발생하는 문제가 존재하는데 바로 오버피팅, 즉, 과적합이다. 오버피팅이란, 데이터 셋에 지나치게 최적화되어 학습 데이터가 아닌 다른 범용 데이터를 입력할 때 정확도가 감소되는 문제다. 매개변수가 많고 표현력이 높은 복잡한 모델, 훈련 데이터가 적을 때 주로 발생한다. 이러한 문제를 해결하며 합성곱 신경망 모델의 성능을 향상시키는 몇 가지 추가적인 방법들이 있다.

첫 번째로, 배치정규화는 각 층에서 활성화 값을 강제로 적당히 퍼뜨리는 방법으로 학습 속도가 증가하며, 초기 값의 영향이 적어지고 과적합을 억제할 수 있다. 내부의 분산 변화를 감소시켜 출력 값이 작아지거나 커지는 변화를 막아준다. 그래서 학습을 안정화 시켜 속도를 향상 시키고 정확도를 올려준다.

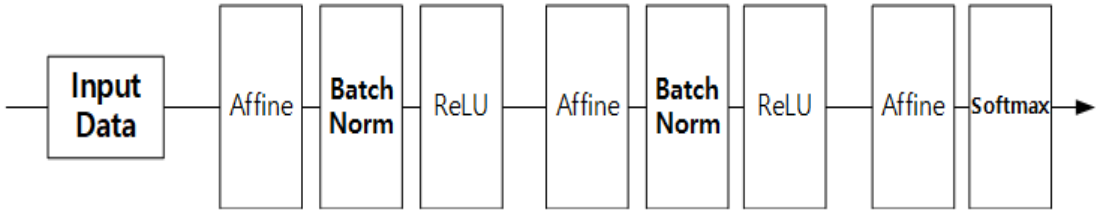


Fig. 2.8 Batch Normalization

두 번째는, 드롭아웃(Dropout)으로 뉴런을 임의로 삭제하면서 학습하는 방법이다. 훈련 때 은닉 층의 뉴런으로 데이터를 보낼 때마다 무작위로 삭제하며 시험 때에는 모든 뉴런에 신호를 전달한다.

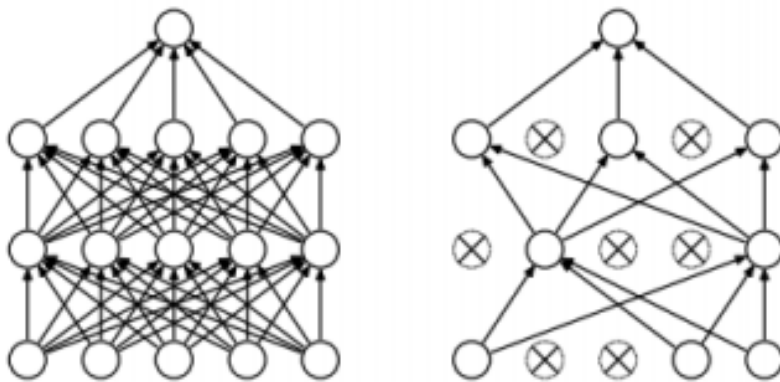


Fig. 2.9 Overfitting

세 번째는 가중치 감소다. 학습 과정 중 큰 가중치에 대해서 그에 상응하는 큰 패널티를 부과하여 오버피팅을 억제한다. 이 외에도 결과의 비선형성을 유지시켜 주어 0보다 큰 값이 출력되면 무조건 1을 반환하여 기울기가 소실되는 것을 방지 해주는 활성화 함수가 있다. 기본적으로 전부터 사용되던 시그모이드 함수(sigmoid)와 하이퍼볼릭 탄젠트(Hyperbolic tangent, tanh), 렐루 함수(Rectifide Linear Unit, ReLu)가 있다.

3.1 LeNet

앞에서 언급한 LeNet 모델은 최초 합성곱 신경망의 모델이다. 첫 컨볼루션 단계에서 합성곱을 연산하여 28x28 사이즈의 6개의 특징 맵(feature map)을 생성하고 두 번째 서브 샘플링 단계에서 특징 맵을 14x14로 사이즈를 줄인다. LeNet은 평균 값 풀링(Average pooling)을 수행하고 sigmoid 활성화 함수를 사용한다. 세 번째 합성곱 단계에서 5x5 합성곱을 연산하여 10x10 사이즈로 줄이고 16개의 특징 맵을 생성한다. 네 번째 서브 샘플링에서는 모든 특징 맵이 다섯 번째 합성곱 계층의 특징 맵과 연결되어 full-connection 단계에서 결과를 unit에 연결시켜 결과를 도출해 낸다.[6]

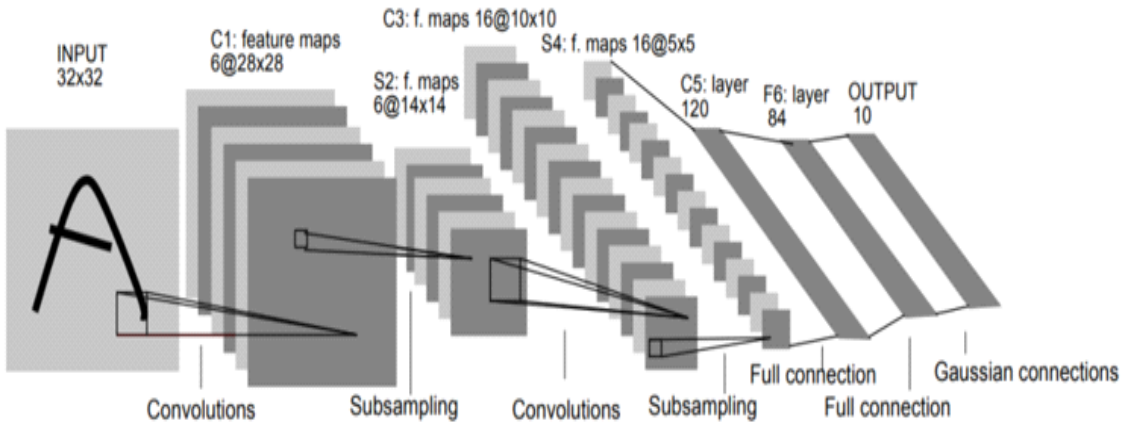


Fig. 2.3 incipient Convolutional Neural Network (LeNet)

3.2 AlexNet

AlexNet 모델은 LeNet에서 2개의 GPU를 기반으로 한 병렬 구조인 점과 비슷하다. 5개의 합성곱 계층과 3개의 full-connection 계층으로 구성되어 있고 속도를 개선시키기 위해 softmax함수를 사용한다. LeNet모델과 각 층의 연산 과정은 동일하지만 227x227의 3개의 채널 이미지를 사용하여 합성곱 계층과 커널의 크기, Stride의 차이가 있다. 또한 크기가 막대한 신경망인 만큼 속도의 문제점을 해결하기 위해 ReLu활성화 함수를 사용하며 과적합 문제의 해결을 위한 또 다른 방법으로 256x256 이미지에서 랜덤으로 227x227 이미지로 자른다. 또한 겹침 폴링(overlapping)으로 조금씩 겹치게 하여 과적합 현상을 개선했다.

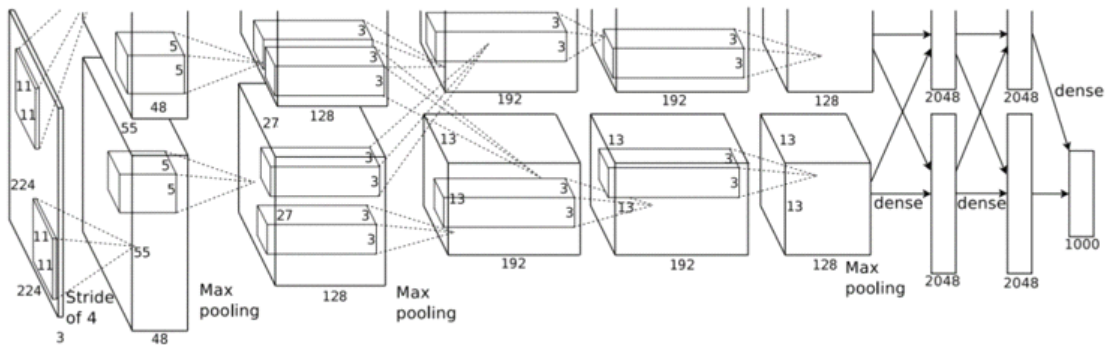


Fig. 2.10 AlexNet

3.3 VGGNet

| ConvNet Configuration | | | | | |
|-------------------------------------|------------------------|-------------------------------|--|--|---|
| A | A-LRN | B | C | D | E |
| 11 weight layers | 11 weight layers | 13 weight layers | 16 weight layers | 16 weight layers | 19 weight layers |
| input (224×224 RGB image) | | | | | |
| conv3-64 | conv3-64 LRN | conv3-64 conv3-64 | conv3-64 conv3-64 | conv3-64 conv3-64 | conv3-64 conv3-64 |
| maxpool | | | | | |
| conv3-128 | conv3-128 | conv3-128 conv3-128 | conv3-128 conv3-128 | conv3-128 conv3-128 | conv3-128 conv3-128 |
| maxpool | | | | | |
| conv3-256 conv3-256 | conv3-256 conv3-256 | conv3-256 conv3-256 | conv3-256 conv3-256 conv1-256 | conv3-256 conv3-256 conv3-256 | conv3-256 conv3-256 conv3-256 conv3-256 |
| maxpool | | | | | |
| conv3-512 conv3-512 | conv3-512 conv3-512 | conv3-512 conv3-512 | conv3-512 conv3-512 conv1-512 | conv3-512 conv3-512 conv3-512 | conv3-512 conv3-512 conv3-512 conv3-512 |
| maxpool | | | | | |
| conv3-512 conv3-512 | conv3-512 conv3-512 | conv3-512 conv3-512 | conv3-512 conv3-512 conv1-512 | conv3-512 conv3-512 conv3-512 | conv3-512 conv3-512 conv3-512 conv3-512 |
| maxpool | | | | | |
| FC-4096 | | | | | |
| FC-4096 | | | | | |
| FC-1000 | | | | | |
| soft-max | | | | | |

Table. 2.1 VGGNet

VGGNet 모델은 다른 모델들에 비해 간단한 구조를 가지고 좋은 성능을 보여 최근 많은 분야에서 쓰이고 있다. 큰 이미지를 학습할 때 커널이 큰 AlexNet과 달리 작은 합성곱 필터로 바꾸어 구성되었다. 3x3 필터 두 개를 사용하는 것이 5x5, 7x7 필터를 사용하는 것과 비슷한 성능을 보여주어 두 개의 필터를 사용함으로써 보다 많은 ReLU함수를 사용할 수 있게 되어 특징을 결정하는 학습이 더 잘되고, 파라미터(parameter)의 수도 줄어들어 속도도 더 빨라진다.[22]

3.4 GoogLeNet

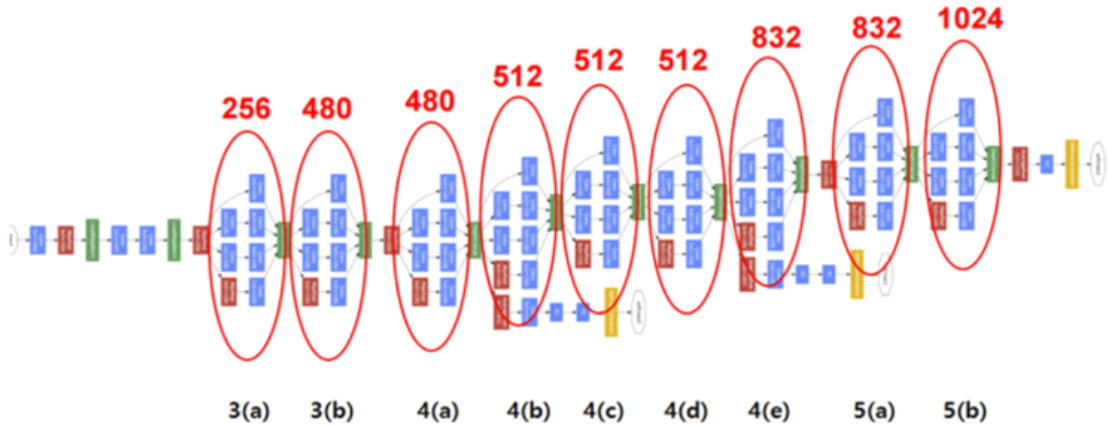


Fig. 2.11 GoogLeNet

딥 러닝은 신경망이 깊고 레이어가 넓을수록 성능이 좋지만, 과적합 등의 문제로 적용하기가 힘들다. 이 문제를 해결하며 나온 모델이 GoogLeNet이다. 이 모델은 Inception모듈 9개를 적용한 형태로 7x7 등의 하나의 합성곱 필터를 사용하는 일반적인 모델들과 달리 작은 합성곱 계층 여러 개를 한 층에서 구성한다.



Fig. 2.12 Module structure of inception

Inception모듈은 더 다양한 특징들을 추출하기 위해 여러 개의 합성곱을 병렬로 사용하여 다른 크기의 특징들을 추출하기 적합하다. 3x3, 5x5의 합성곱은 연산량이 크기 때문에 앞에 1x1 합성곱을 두어 특징 맵의 개수를 줄여 연산량을 조절할 수 있다. 즉, 여러 층의 표현을 좀 더 함축하여 표현할 수 있다.

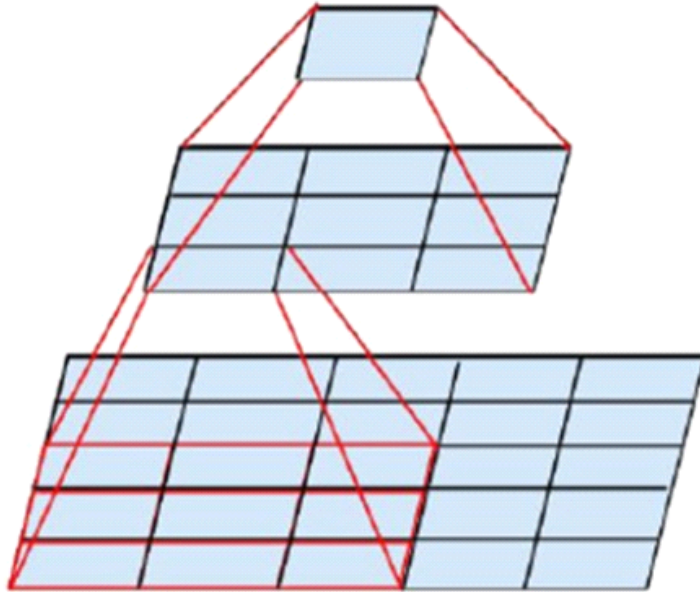


Fig. 2.13 Mini-network replacing 5x5

Fig. 2.13 처럼 큰 5x5 필터 크기를 갖는 커널을 인수분해하면, 작은 커널 여러 개로 구성된 깊은 네트워크를 만들 수 있고, 파라미터가 적어지면서 신경망은 깊어지는 효과를 얻을 수 있다. 5x5 합성곱은 3x3에 비해 넓은 영역의 특징을 1번에 추출 가능 하지만 연산량이 막대하여 3x3 합성곱이 더 효과적이다.

기존의 Inception 모듈을 최적화 하면서 v2, v3 등의 모델이 나오게 되고, v3 이전의 모델들은 학습이 장시간에 걸쳐 이뤄지게 되면 과적합과 구배가 사라지는 현상(gradient vanishing)이 종종 일어나게 되는데 v3 모델은 최적화와 연계성을 바탕으로 보완되었다.

4. Al-Si alloy

Al합금은 가벼우면서 내식성과 비강도, 열전도성 등의 우수한 성질을 가지고 있다. 그 중 Al-Si계 합금은 가장 일반적인 주조용 합금계로 산업용 주조의 상당한 비중을 차지한다.

Al-Si 합금의 공정 온도는 577°C 이며 공정 조성에서 Si를 형성한다. 형성된 Si의 함유량은 약 12.6 wt% 이고 Al이 주성분이 되는 α 상은 공정 온도에서 최대 고용률이 1.65 wt% 지만 Si상 쪽 Al의 고용률이 전무하여 Si상은 Al을 거의 포함하지 않는 순수한 상이다.

과공정 Al-Si 합금은 높은 주조성과 높은 강도, 내고온성 등의 특징을 가지지만, 연신율과 피로특성이 낮다. 반면 아공정 Al-Si 합금은 강도가 높지 않지만 높은 연신율과 피로특성을 가지고 있다.

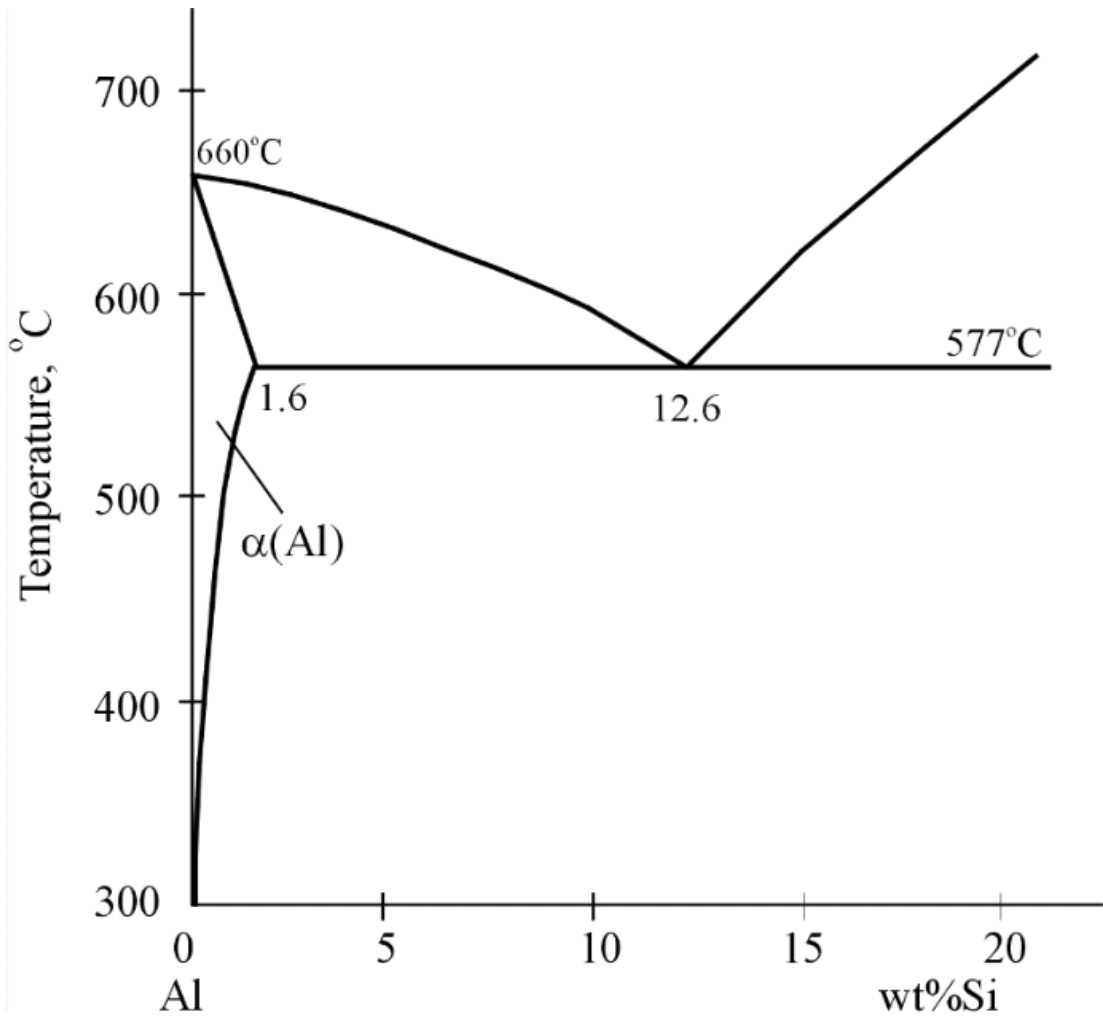
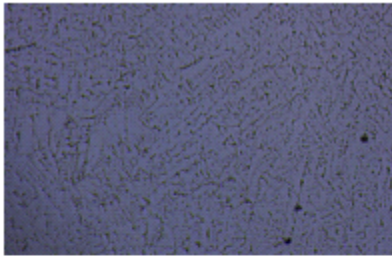


Fig. 2.14 Al-Si phase Diagram

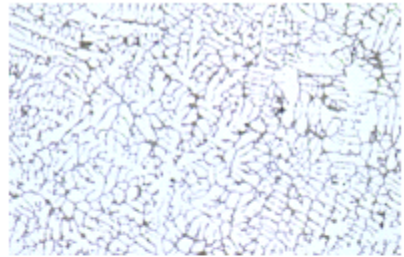
제 3 장 실험 방법

3.1 Al-Si 주조시편 제작

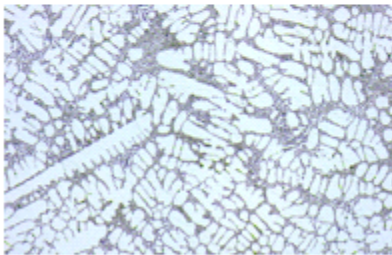
본 실험에서는 Al-Si 이원계 합금으로 Si 함량을 1~10%까지, 1 wt% 간격으로 제작하였다. 공정 조성이 약 12.6 wt% 지만 Si 함유량이 공정 조성보다 적은 아공정 합금으로 α 상이 응고 초기에 발생 후 수지상정 형상으로 성장한다. 냉각 정보가 다른, 다양한 시편을 얻기 위해 폭이 다른 계단형 주형을 사용하였다. 약 200° C로 주형을 예열하였고, 용탕의 주입 온도는 약 750° C로 일정하게 설정하였다. 응고가 완료된 시편을 연마와 미세 연마 후 켈러시약(Keller's reagent: 95 % H₂O, 2.5% HNO₃, 1.5% HCl, 1.0% HF)으로 에칭(Etching)하여 광학현미경을 통해 미세조직을 관찰하였다.



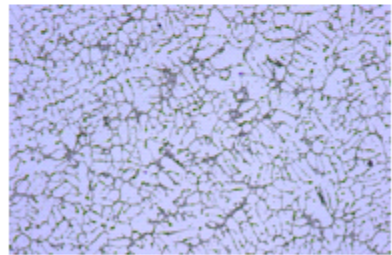
Al-1%Si



Al-4%Si



Al-7%Si



Al-10%Si

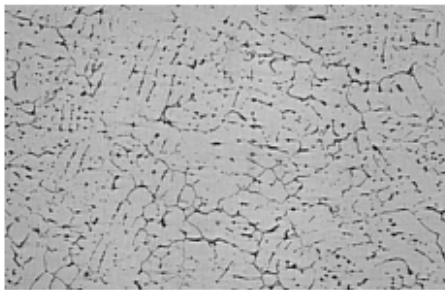
Fig. 3.1 Al-Si microstructure

이와 같이 관찰한 Al-Si 합금의 대표적인 일부 사진들을 Fig. 3에 나타내었다. 제작 한 시편이 아공정 합금으로 Si 함량 증가에 따라 α 상 수지

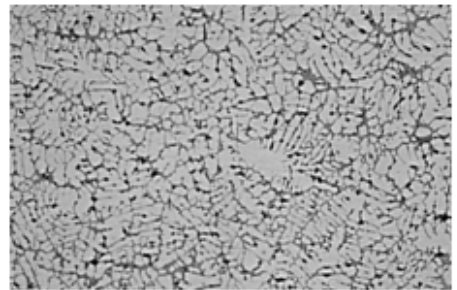
상정의 2차 수지성정 간격이 증가함을 볼 수 있다.

3.2 이미지 전처리

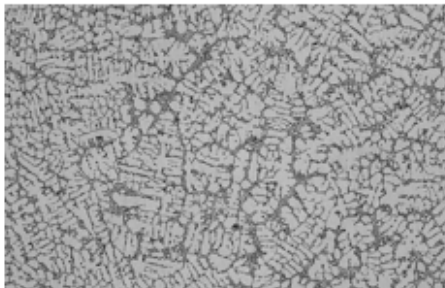
광학현미경으로 관찰한 미세조직을 $\times 50$, $\times 100$, $\times 200$ 의 배율로, Si 함량마다 각 100장씩 사진을 촬영하였다. Scale bar를 통해 배율 정보를 남기면 이미지 인식에 영향을 줄 것이라 판단하여 배율 정보를 배제하고, 컬러 사진들을 흑백으로 전환하며 명암을 최대한 유사하게 조정하여 인식하기 쉽게 사진들의 정보를 수정하였다. 즉, 조성들의 각 이미지들은 배율과 냉각 조건, 촬영 부분이 모두 다르게 존재해있다.



Al-1%Si



Al-4%Si



Al-7%Si



Al-10%Si

Fig. 3.2 Microstructure Image Preprocessing

3.3 모델 선정 및 알고리즘 설계

각 모델들은 구조(Architecture)가 다른 만큼 특징을 가지고 있어 활용 방안 에 따라 각기 다른 모델들을 사용한다.

1) LeNet 모델은 손으로 쓴 글씨를 32 x 32 픽셀 이미지로 디지털화 한 것을 인식하는 신경망으로 합성곱 계층과 풀링 계층을 반복하는 모델이다. 이미지와 관련된 신경망 분야에서 처음 나온 모델이기에 고화질 이미지를 분석하기 위해서는 더 많은 신경망과 계층을 추가적으로 확대해야 한다. 즉, 저해상도 이미지인 Text 기반의 문자 분류에 최적화 되어있다.

2) AlexNet 모델의 경우, GPU 사용을 고려한 최초의 모델로써 두 개의 병렬 네트워크로 구성되어 약 0.4%~1.4% 정도의 정확도 향상과 대량의 데이터 연산이 가능한 모델이지만, 많은 수의 파라미터가 필요하게 되어 초기에 비해 많이 쓰이지 않고 있다.

3) VGGNet 모델은 간단한 구조와 단일 네트워크의 사용으로 성능이 우수해 여러 컴퓨터 비전에 응용되고 있다. 큰 필터 하나를 사용하는 것에 비해 3 x 3필터 두 개를 사용하여 파라미터 수를 줄여 연산량을 대폭 감소시켜 계산이 빨라진다.

4) GoogleNet(Inception) 모델은 효과적으로 차원을 줄이고 네트워크는 더 깊게 만든 구조를 가지고 있다. 또한, 1 x 1 합성곱 계층을 사용함으로써 특징 맵을 생성하여 더 비슷한 성질을 갖는 특정 맵들을 묶어주고 차원축소 효과를 보여 깊은 네트워크임에도 불구하고 적은 파라미터 수를 가지게 되며 연산량이 줄어 빠른 계산 속도를 보인다.

이처럼 많은 계층 수로 어려움을 감소시키는 데 중점을 둔 모델, 빠른 처리속도에 기반을 둔 모델, 대량의 데이터를 처리하는데 적합한 모델 등이 존재하며, 쓰임에 따라 각기 다른 모델들을 사용하거나 여러 모델들을 혼합해서 사용하기도 한다.

본 논문에서는 미세조직 이미지를 사용하기에 특징 맵을 생성하여, 더 정교한

특징 추출을 기반으로 하는 GoogleNet 모델의 Inception으로 알고리즘을 응용하게 되었고, 처리속도 향상을 위해 학습하는 training 코드와 분류, 예측을 위한 predict 코드를 나누어 한번 학습 시키고 여러 이미지를 테스트 할 수 있게 하였다. training 코드에는 훈련을 진행할 때 일정 조건으로 훈련할 수 있도록 이미지에 대한 용량, 명암, 사이즈를 일정하게 맞추어 주는 항목을 추가하였다. 또한 predict 코드에 상대적이지만, 백분율로 유사도를 나타내었다.

실험하기 사용될 DataBase를 횟수(epoch)를 설정하여 미리 훈련시키고 알고자 하는 Test Image를 Input data로 입력을 하면, predict 코드에 의해 Si 함량 몇%에 해당하는 미세조직 이미지인지 구분을 해주게 된다.

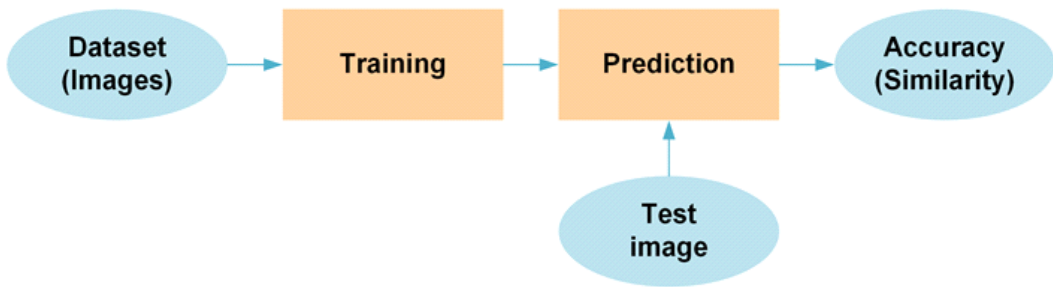


Fig. 3.3 Experimental schematic

훈련에 사용할 미세조직 이미지로는 광학현미경으로 촬영 후, 아무런 작업도 하지 않은 원본 이미지와 Scale bar를 삭제한 이미지, 앞서 언급한 전처리 이미지, 전처리 이미지의 반전과 절반으로 나눈 이미지들을 준비한다.

- 1) 원본 이미지 각 100장 = 1000장
- 2) Scale bar 가지고 있는 이미지 각 100장 = 1000장
- 3) 전처리 이미지 각 50장 + 반전 이미지 각 50장 = 1000장
- 4) 전처리 이미지 50장 → 반쪽 이미지 각 100장 = 1000장

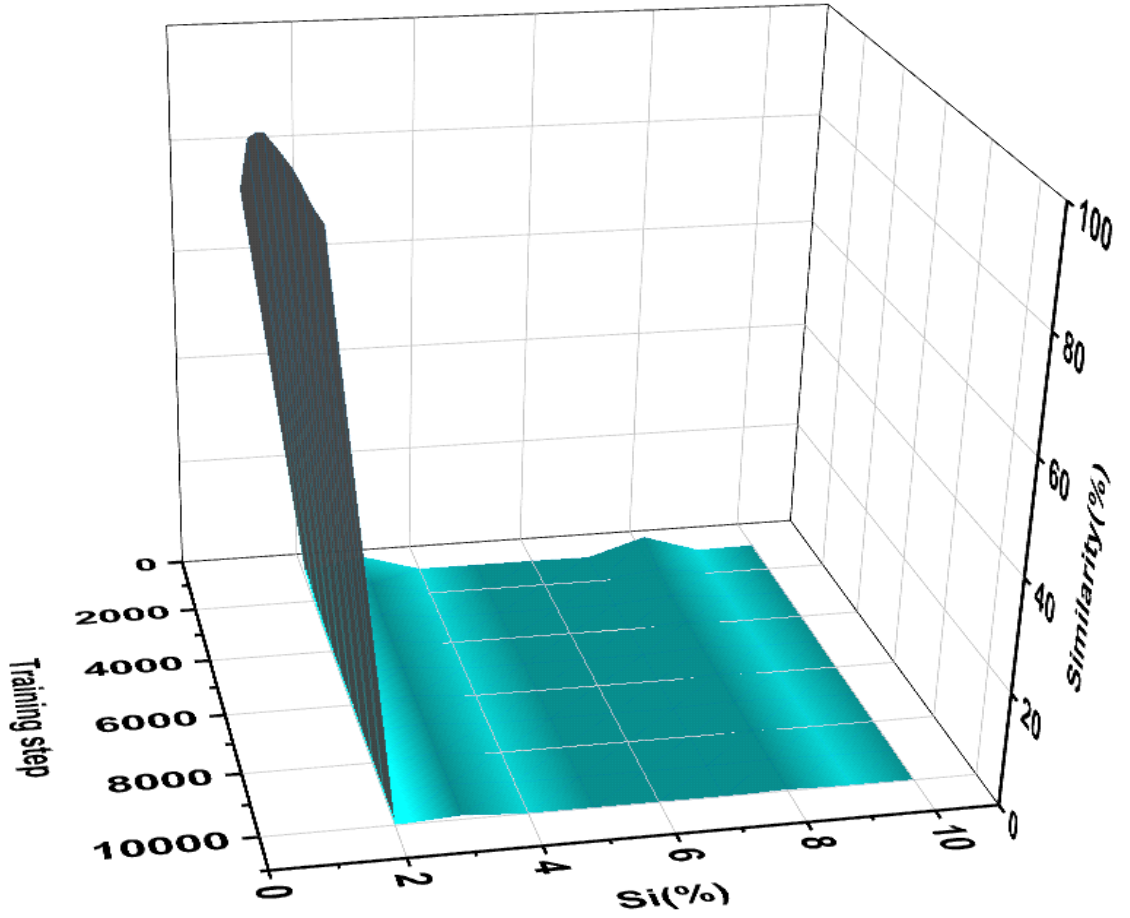
Training 코드에 이미지를 통일시켜 주는 항목이 있기에 원본 이미지와 전처리

한 이미지의 정확도를 비교해보고자 하고, Scale bar가 존재하는 이미지와 배재한 이미지의 정확도, 데이터베이스 증가에 이미지의 수정이 미치는 영향을 알아보기 위해 다른 조건의 이미지들을 준비하고 각각 1000번부터 10,000번까지 학습을 진행한다. 10,000번을 넘어서는 경우 정확도 증가의 동향이 미세하기에 최대 10,000번까지 학습을 하게 되었으며, 데이터 셋이 수십만 장에 이른다면 더 정확한 유사도를 얻기 위해 학습 횟수도 그에 따라 증가하게 될 것으로 보인다.

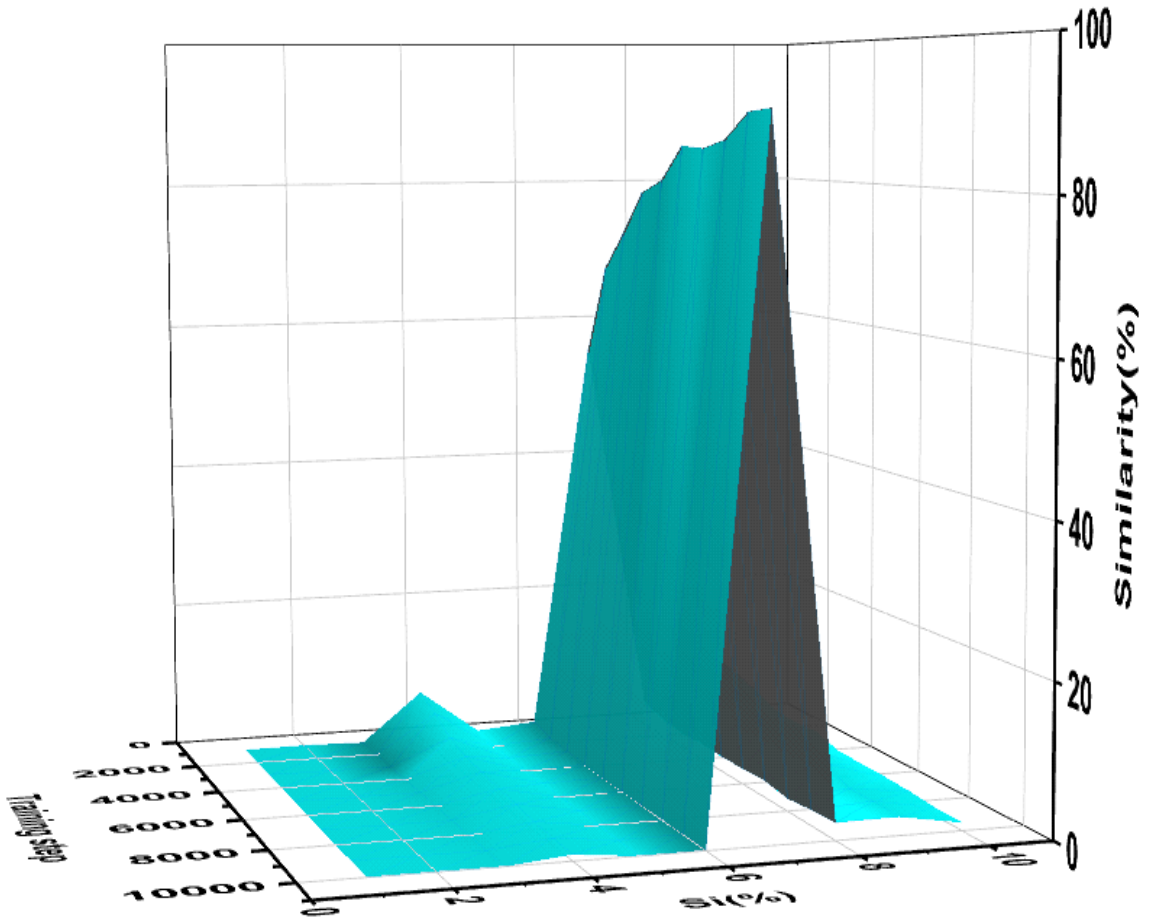
제 4 장 결과 및 고찰

4.1 기계학습을 이용한 합금성분의 역산출

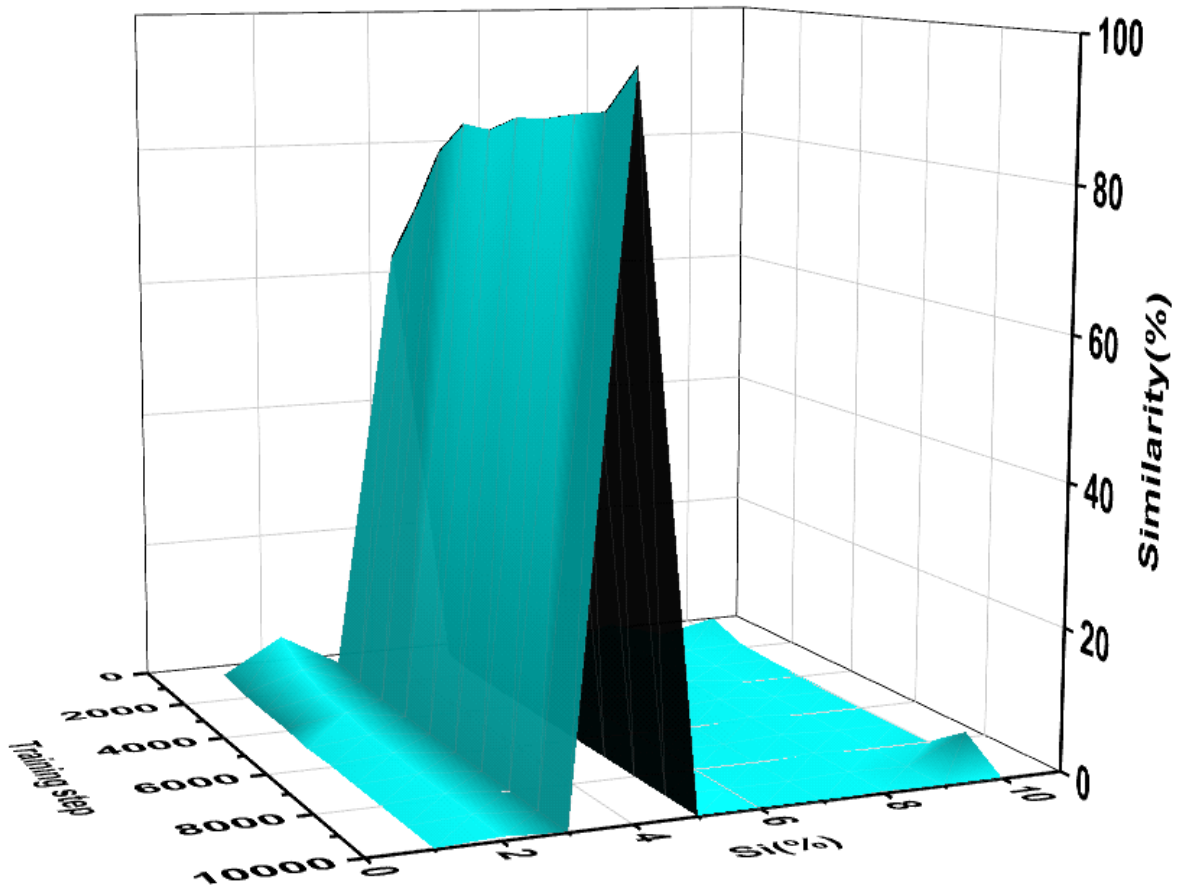
이미지에 대한 기계학습은 반복학습이 요구되는데, 그 영향을 알아보기 위하여 10,000번까지 1000번 단위로 학습하고 실험을 진행하였다. 조성예측 단계에서는, 학습된 데이터 셋에 포함되지 않은 조성이 알려진 이미지를 입력하였고, 그 결과의 일부를 Fig. 4.1에 나타내었다.



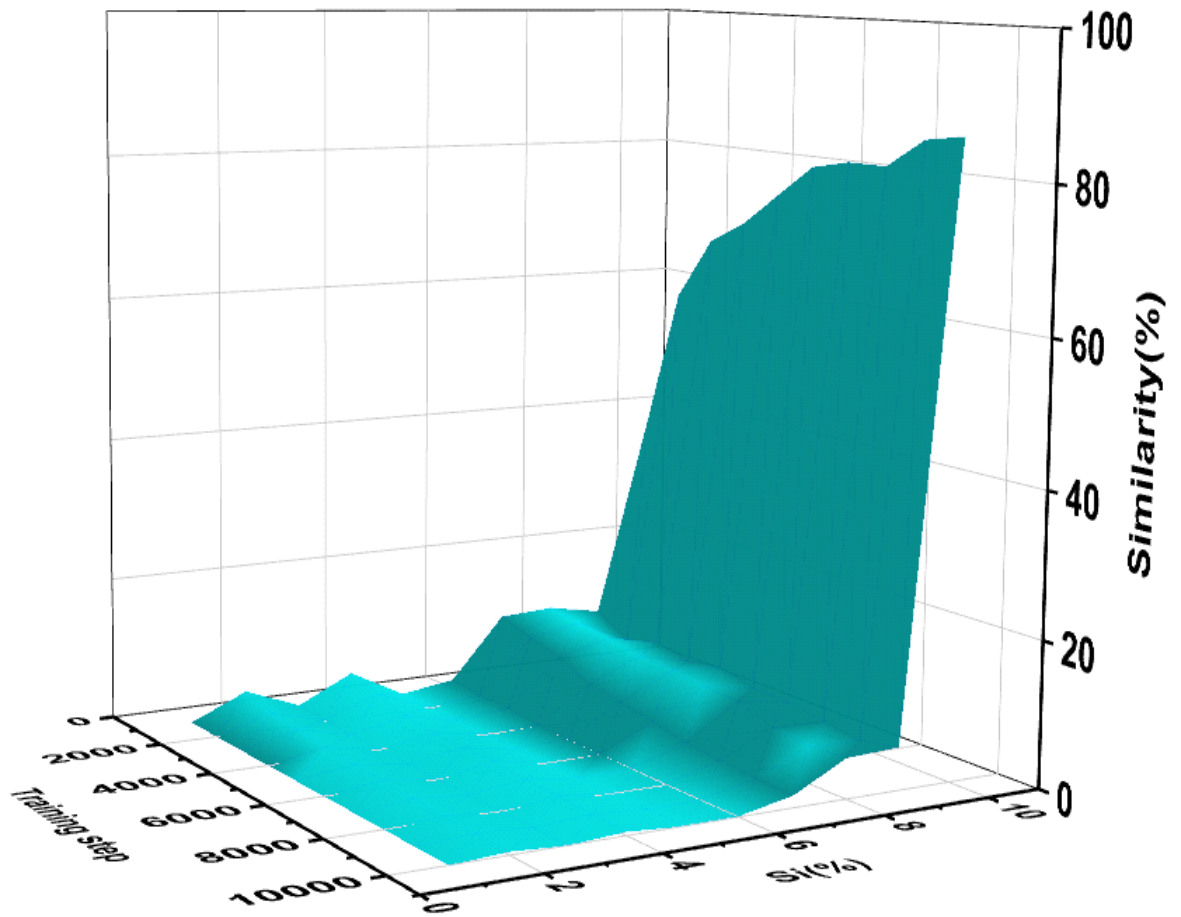
(a)



(b)



(c)



(d)

Fig. 4.1 Similarity distribution evolved with training step for (a) Al-1 wt%Si, (b) Al-4 wt%Si, (c) Al-7wt%Si, and (d) Al-10 wt%Si alloys.

그림에는 대표적으로 Si 함량 1%, 4%, 7%, 10%를 나타내었고, 학습 횟수와 관계없이 어느 조성에서나 해당 조성의 피크가 관찰되었다. 해당 조성이 아닌 다른 조성에도 낮은 피크가 보이는데 이러한 현상은 학습 횟수가 증가함에 따라 피크의 크기가 점차 감소하였으며, 해당 조성은 피크의 폭이 좁아지고 유사도는 증가하는 경향을 나타내었다. 이는 학습 횟수가 증가할수록 기계 학습의 신뢰도가 높아짐을 보여준다.

나머지 각각의 조성들도 Fig.4.2의 그림을 통해 해당 조성의 피크가 가장 높음을 나타내었고, 각 조성들의 유사도를 가중치로 Si의 평균 함량을 계산하여 정량적인 결과를 Fig. 4.3에 나타내었다. 학습 횟수가 적을 때에는 해당 조성들에서 벗어나는 경향이 있지만, 학습 횟수가 증가함에 따라 목표 조성에 근접함을 알 수 있다. 학습 초반에 이러한 경향이 나타나는 것은 데이터 셋에 사용된 미세조직 이미지들의 배율이 각기 다른 점, 조성은 같지만 냉각속도가 다르기에 일어나는 2차 수지상 간격, 등의 이유가 있을 것으로 판단된다. 하지만 학습 횟수의 증가에 따라 이러한 현상들이 결과에는 크게 미치지 않는 것으로 보아 이미지의 유사도에는 크게 영향이 없는 것으로 판단된다.

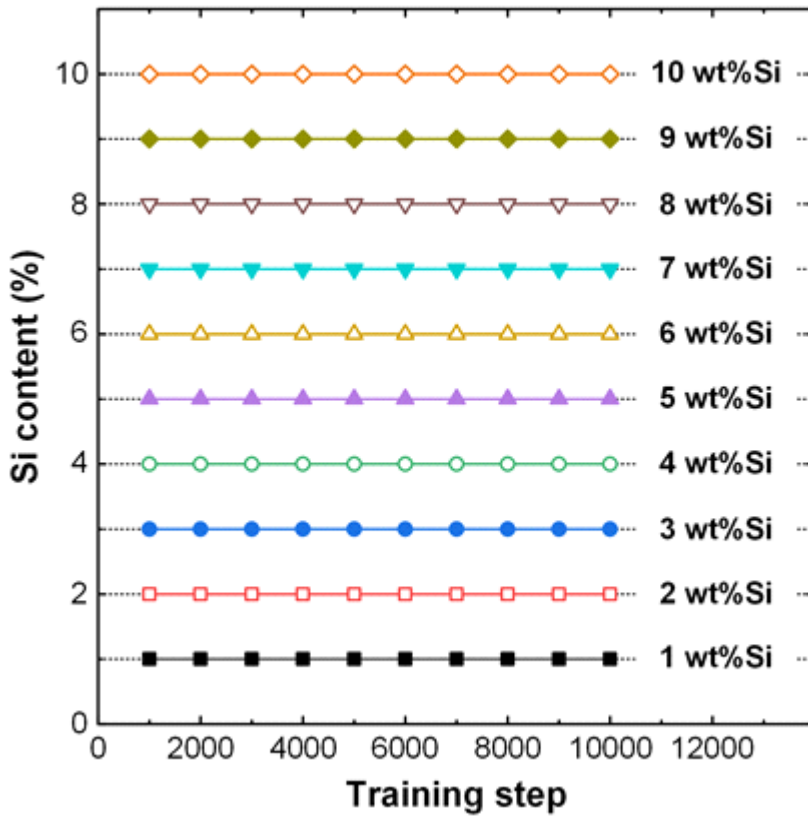


Fig. 4.2 Si content where the highest peak existed, as a function of training step.

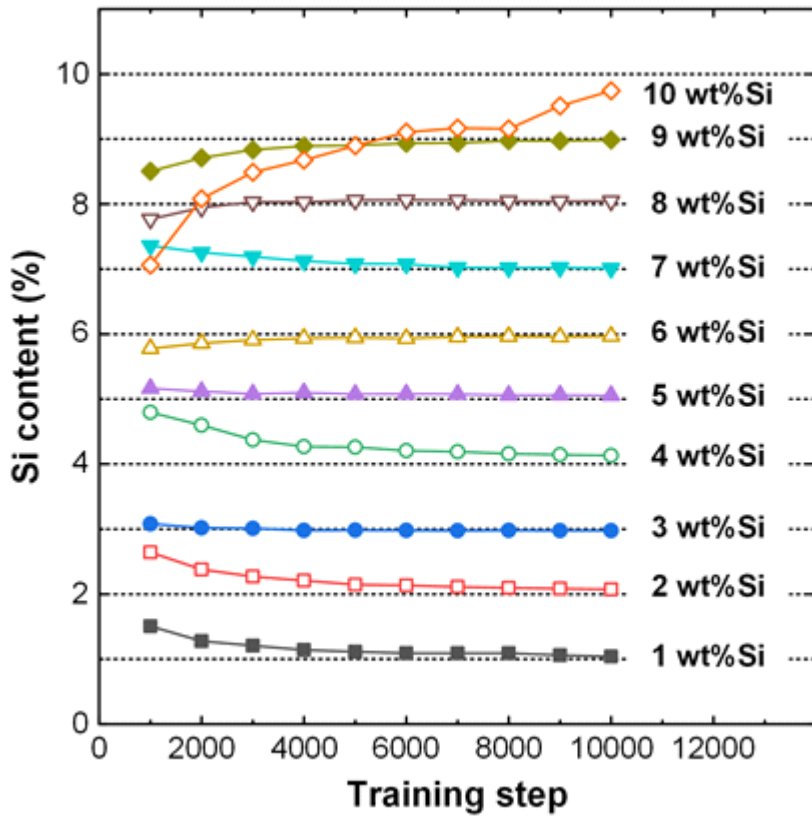
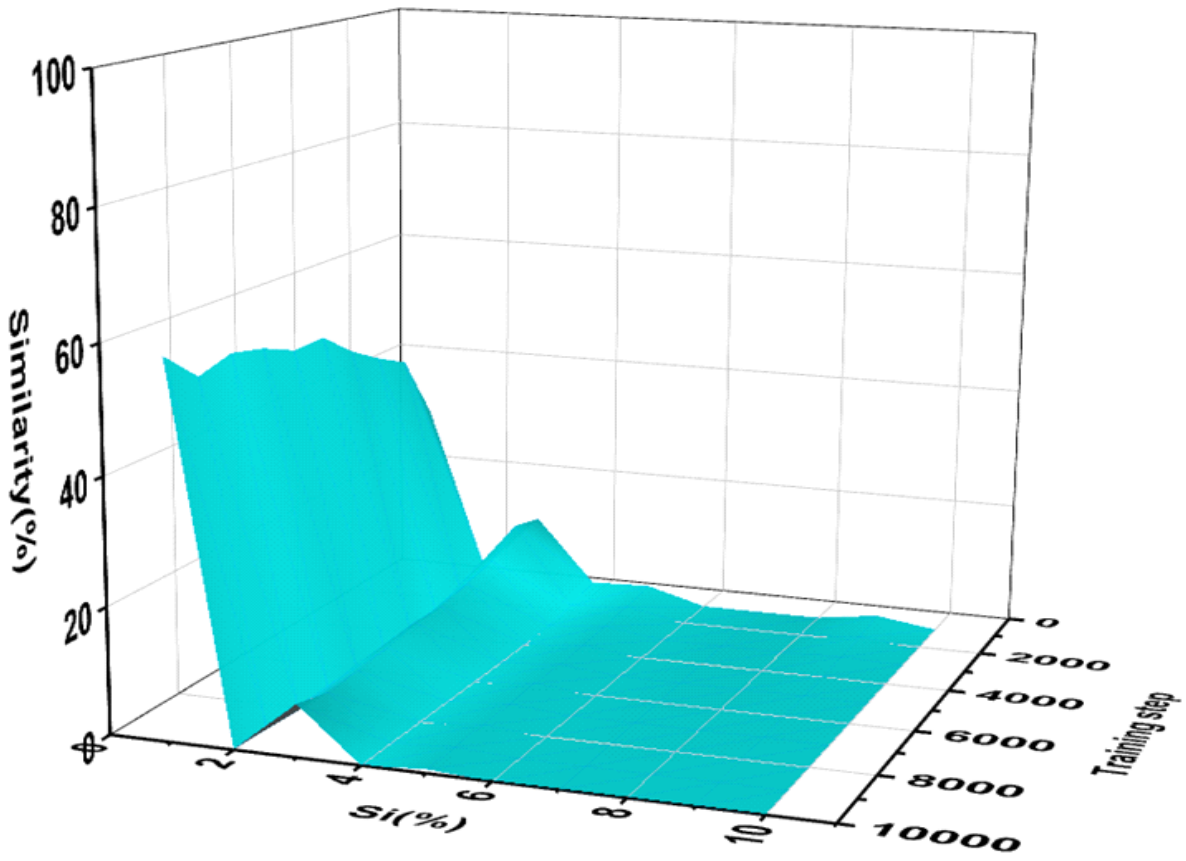


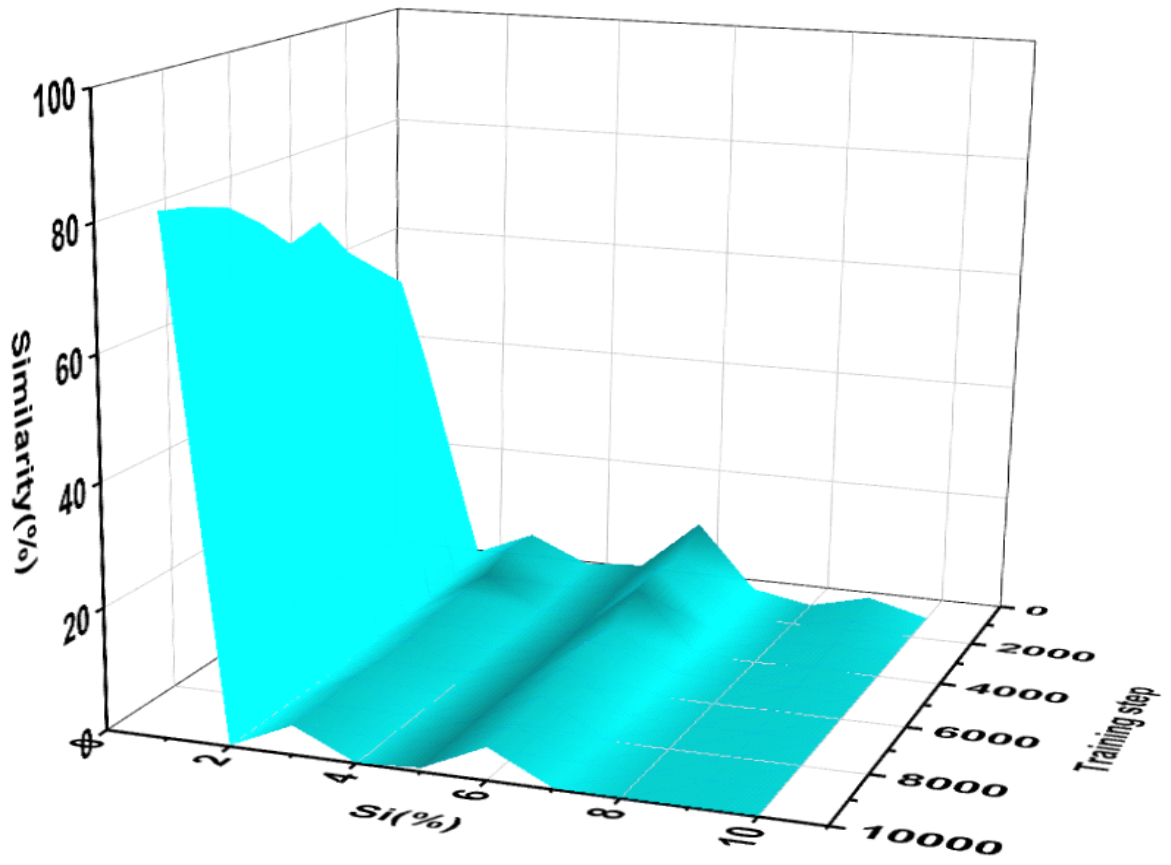
Fig. 4.3 Weight-averaged Si content of the alloys. These values approached to the accurate value as train step increased.

4.2 미세조직 이미지 원본과 변환된 이미지 유사도 비교

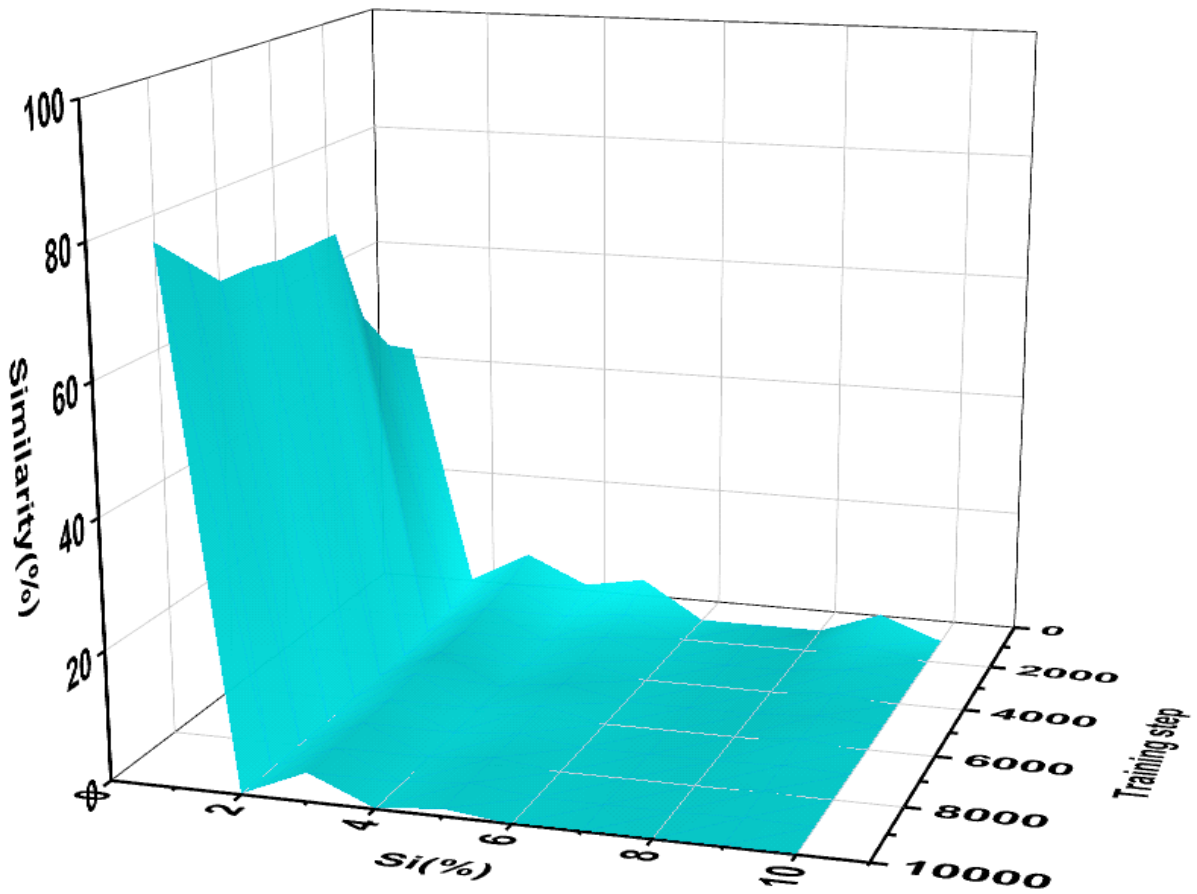
신뢰도 높은 데이터 셋을 구축하기 위해서는 많은 수의 데이터가 요구된다. 하지만 현실적으로 시편의 수 백 장 이상의 이미지 데이터를 얻기에는 어려움이 많다. 그래서 원본의 이미지 데이터를 변환하여 데이터베이스 증가에 영향을 주는 지, 명암 등이 다른 이미지들을 최대한 비슷하게 통일하여 실험을 하였을 때와, 원본 그대로 실험하였을 경우 상대적이지만 어떤 데이터 셋이 유사도가 더 높게 나오는지 실험을 진행하였다. 끝에 위치하는 조성인 Si의 조성 1%와 중간에 위치하는 7%의 조성을 대표적으로 나타내었고, 축척표시의 유무에 따른 정확도를 중간 값에 속하는 4% 조성과 6% 조성으로 비교하였다.



(a)

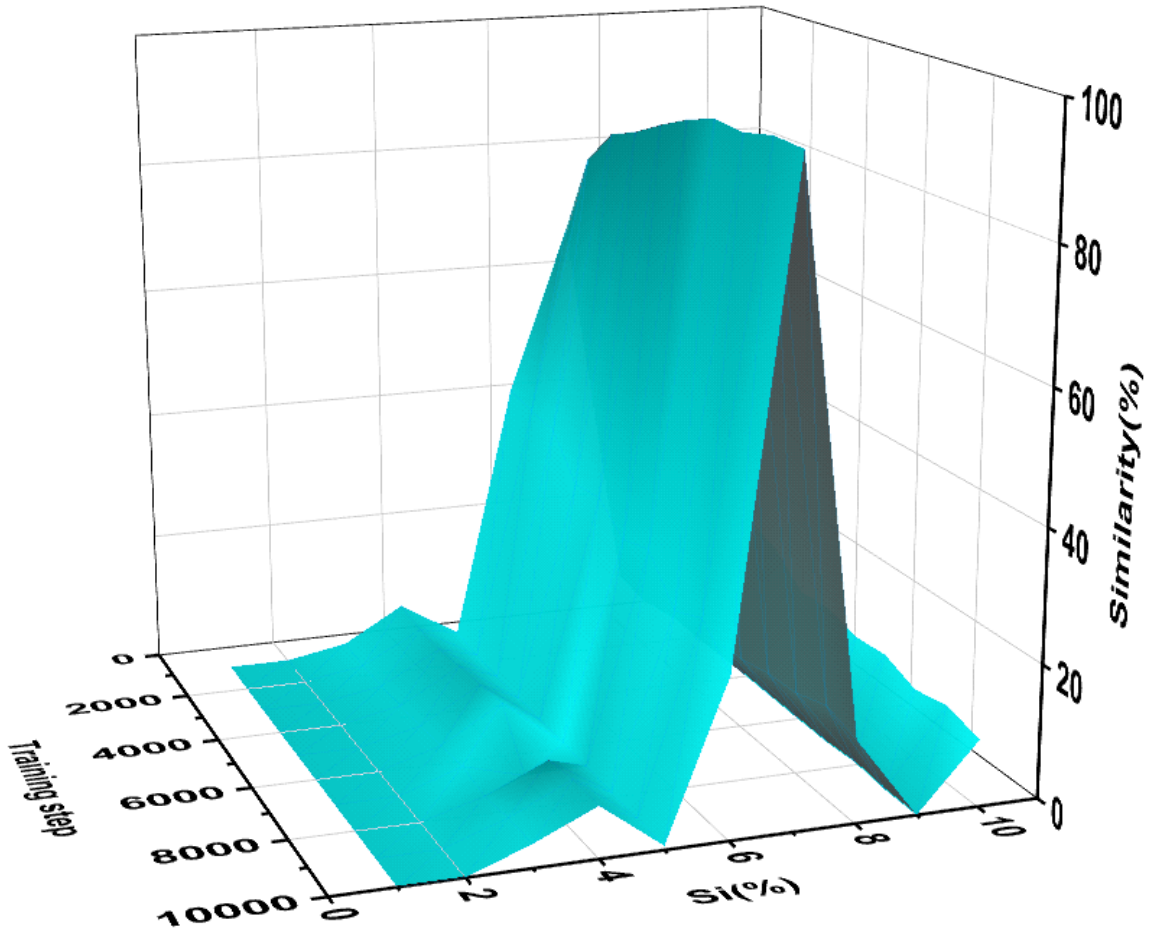


(b)

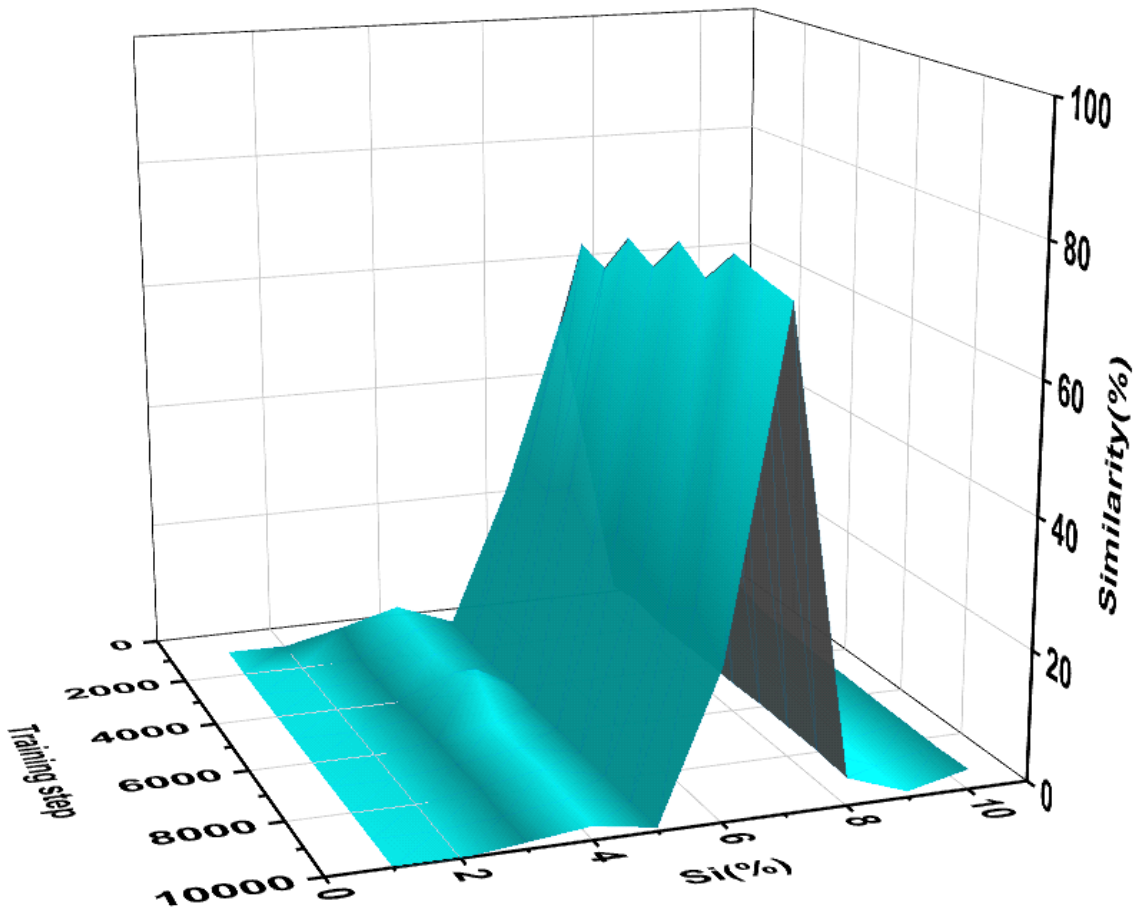


(c)

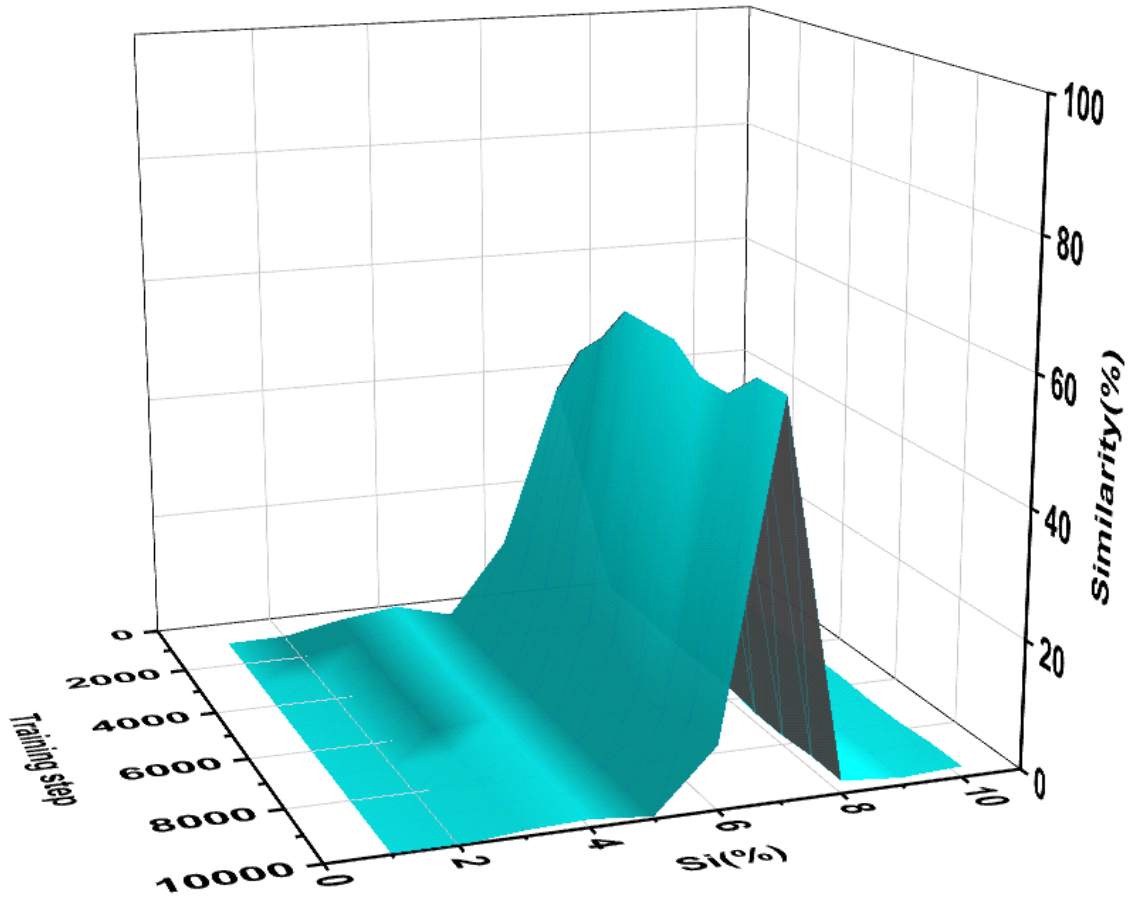
Fig. 4.4 Similarity distribution evolved with training step for (a) preprocessing + reverse, (b) preprocessing + reverse + half, (c) preprocessing + half Al-1wt%Si alloys.



(a)

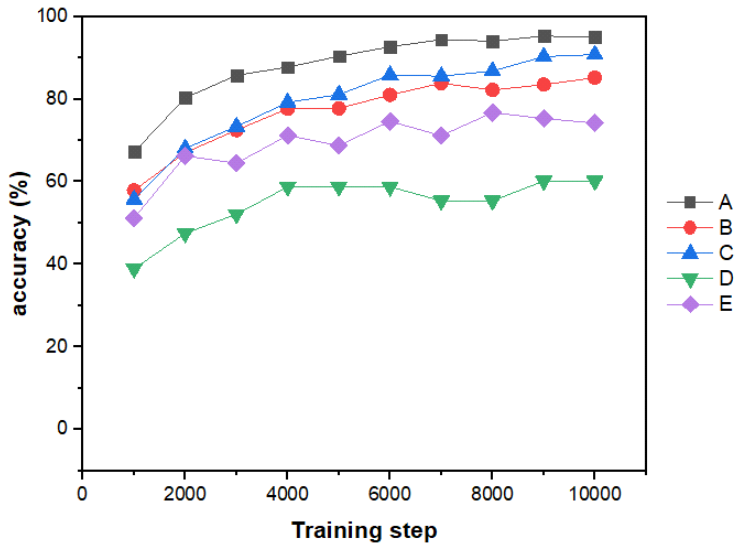


(b)

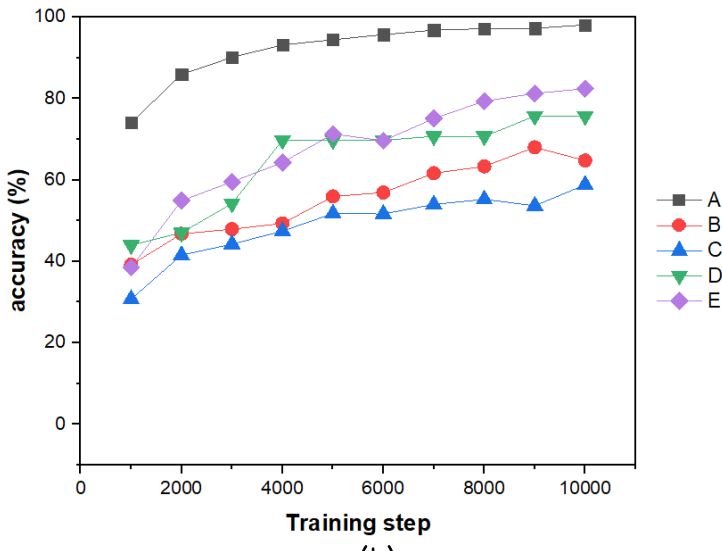


(c)

Fig. 4.5 Similarity distribution evolved with training step for (a) preprocessing + reverse, (b) preprocessing + reverse + half, (c) preprocessing + half ,Al-7wt%Si alloys.



(a)

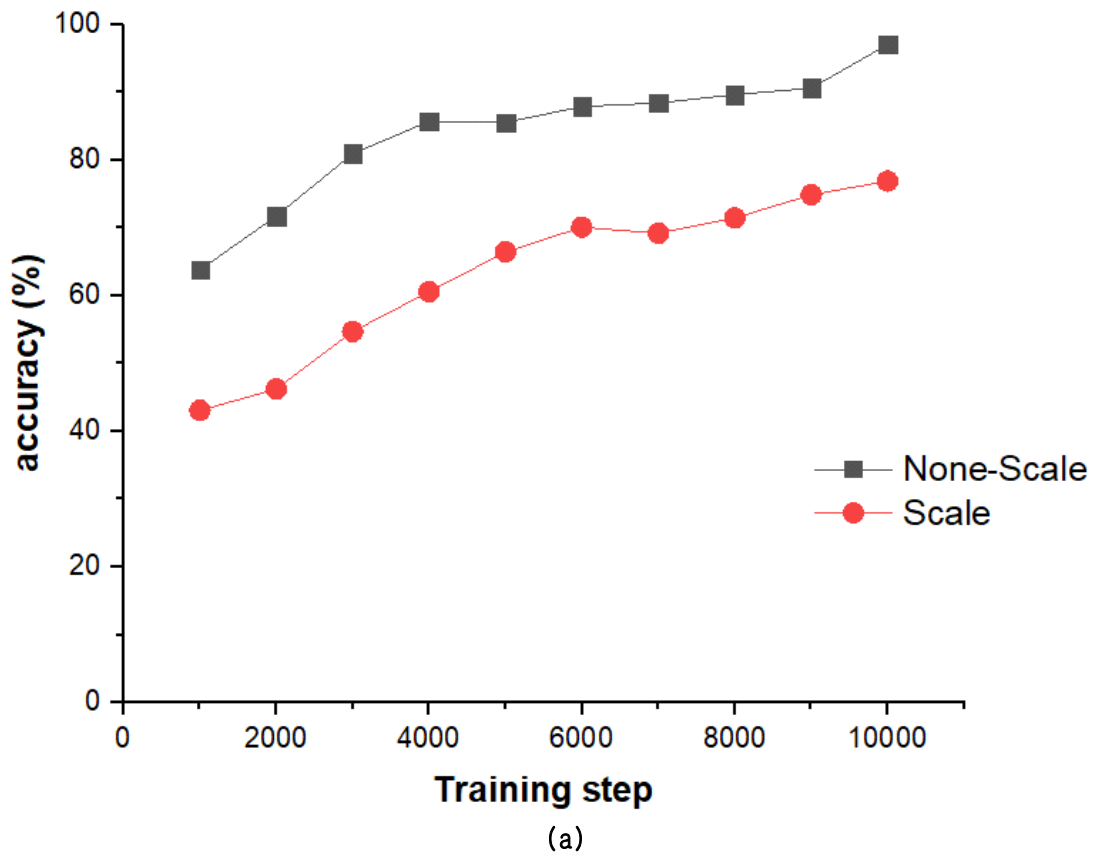


(b)

Fig. 4.6 Comparison of original and converted images for (a) Al-1wt%Si, (b) Al-7wt%Si alloys. (A - original, B - preprocessing, C - preprocessing + reverse, D - preprocessing + reverse + half, E - preprocessing + half)

Fig. 4.4~4.6을 통하여 원본 이미지 데이터의 정확도가 가장 높게 나옴을 알 수 있었다. 데이터베이스가 많을수록, 학습 횟수가 많을수록 신뢰성 높은 정확도를 나타내야 하지만, 데이터베이스를 증가 시키고자 이미지를 변환시킨 데이터 셋은 오히려 낮은 정확도를 나타내어 미세조직 데이터에 한하여 상대적으로지만 이미지 변환의 불필요성을 나타내었다. 또한, 데이터의 통일을 위해 명암과 크기 등 전처리한 이미지도 원본 이미지에 비해 정확도가 낮게 나타났는데, training 코드에 앞서 언급한 이미지의 용량, 명암, 크기를 일정하게 해주는 항목으로 인해 이미지의 전처리도 불필요한 것으로 판단된다.

이미지의 배율 정보가 이미지 인식에 영향을 주는지에 대하여 Fig.4.7을 통해 나타내었다. 예상한 바와 같이 Scale bar가 차지한 영역에 대하여 컴퓨터가 받아들이는 데이터 특징의 양에 영향이 있다고 판단되며, Si 조성 4%에서는 확연하게 나타나고, 6%의 조성에서는 학습 초기에는 경향이 다르지만 학습 횟수가 증가함에 따라 배율 정보가 없는 이미지에 대한 정확도가 상대적으로 더 높아짐을 알 수 있다.



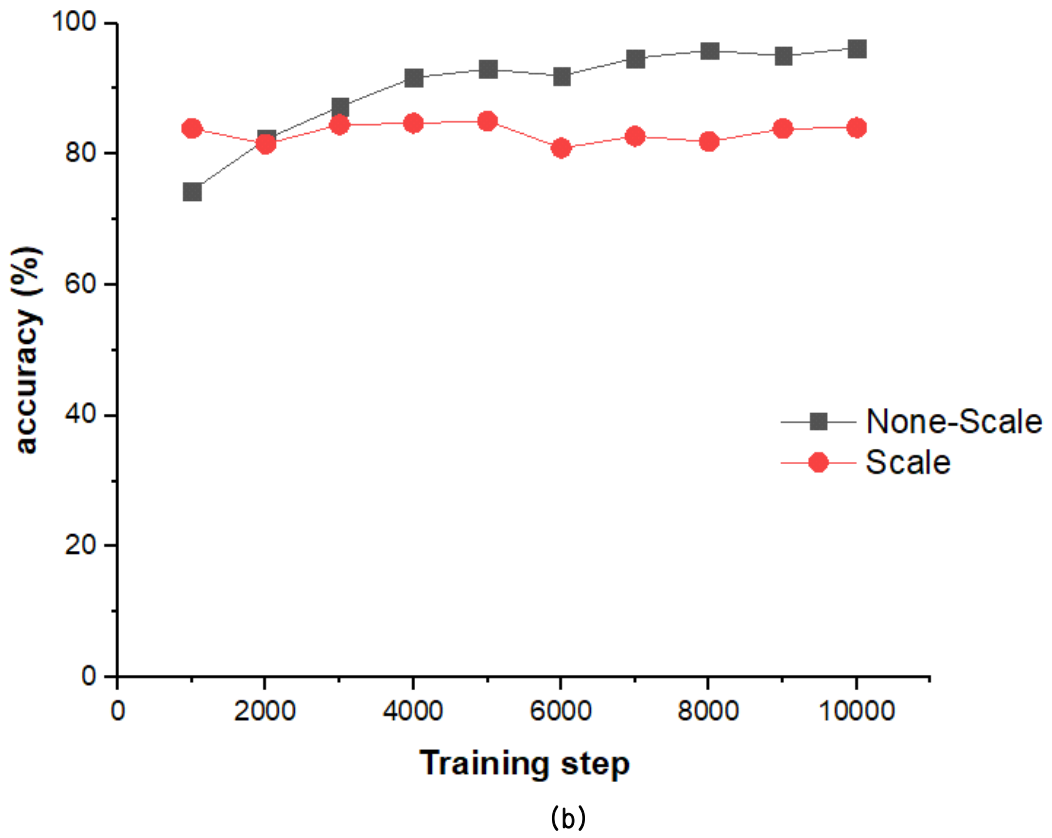


Fig. 4.7 Accuracy comparison with or without scale bar for (a) Al-4 wt%Si, (b) Al-6 wt%Si alloys.

4.3 데이터 셋 범위를 벗어난 조성의 시편 분석

기존의 데이터 셋의 범위를 벗어나는 바깥의 조성을 입력하게 되면 결과가 어떻게 나타날지 판단하기 위하여, Si의 조성 3 wt%부터 8 wt%까지 10,000번 학습하여 실험을 진행하였다. 그 후 학습 범위에 존재하지 않는 2 wt%와 9 wt%의 조성을 입력하여 조성 예측을 시도하였다. Fig. 4.8에 나타난 바와 같이, 2 wt%조성의 경우 그래프 조성의 하한선인 3 wt%조성에 피크가 발생하였고, 마찬가지로 9 wt%의 조성도 그래프 조성의 상한선인 8 wt%조성에서 피크가 발생하였다. 이를 토대로 보아, 데이터 셋의 범위에 존재하지 않을 경우 범위 중 가장 유사한 패턴을 보이는, 합금 조성이 가장 비슷한 데이터 셋에서 높은 정확도를 얻을 수 있다는 사실을 알 수 있다. 그러므로 목표 조성이 데이터 셋의 범위에 있지 않으면, 정확한 조성을 예측할 수 없기에 먼저, 범위 내에 목표 조성이 존재하는지를 확인하는 것이 중요하다.

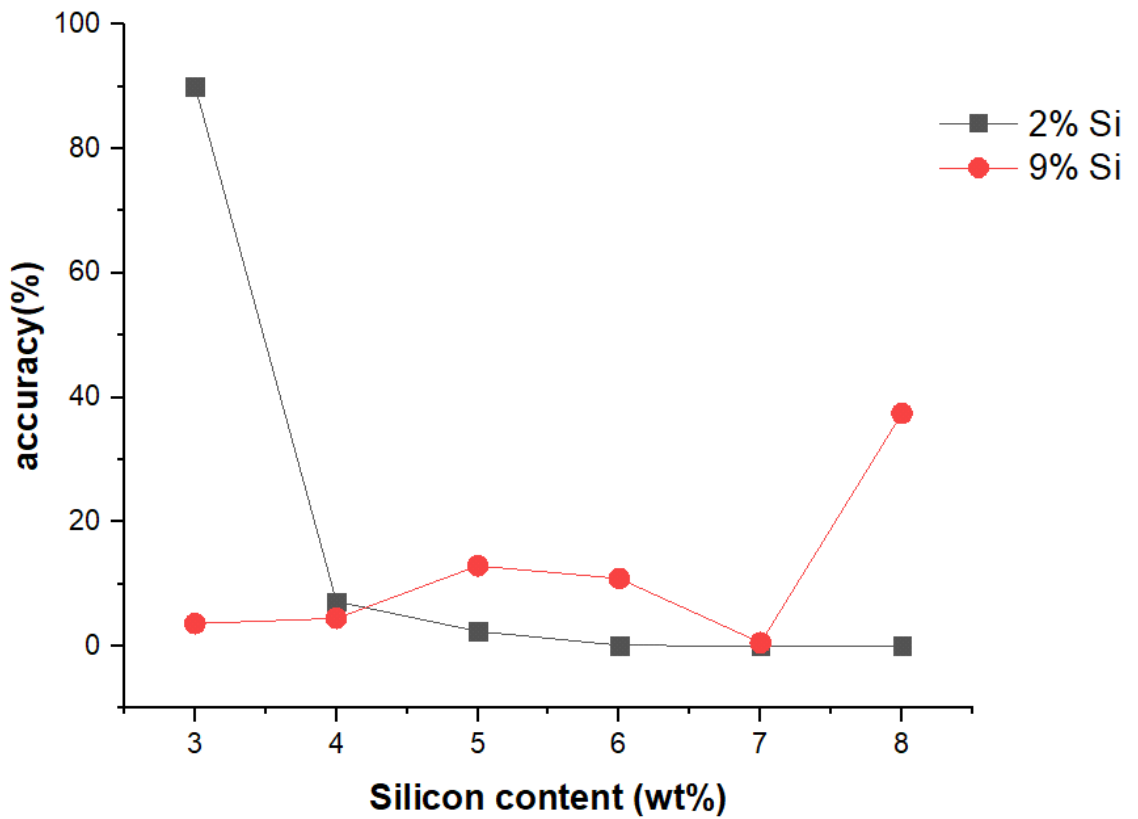


Fig. 4.8 accuracy distributions of the samples with 2 and 9 wt%Si,
 when the dataset has the range of 3-8 wt%Si.

4.4 신경망 모델 비교

실험에 선정된 Inception모델과 성능 비교를 위해 초기에 개발된 LeNet모델을 같은 조건으로 Si의 조성 1 wt% ~ 10 wt% 조성까지 10,000번 학습 후 기존의 실험 방법과 같이 학습에 사용하지 않은 확인 된 각 조성의 이미지를 입력한 결과 값을 Fig. 4.9에 나타내었다. 초기에 개발된 LeNet모델은 손 글씨 전용의 저해상도에 최적화된 모델인 만큼 각 조성들의 정확도가 매우 불규칙하게 나타났으며, Inception모델의 경우 높은 정확도 값에 고른 분포를 보여줌을 알 수 있다. LeNet 모델의 경우 학습 시간 또한 약 140분으로, 97분이 소요된 Inception모델에 비해 확연한 차이를 보였다.

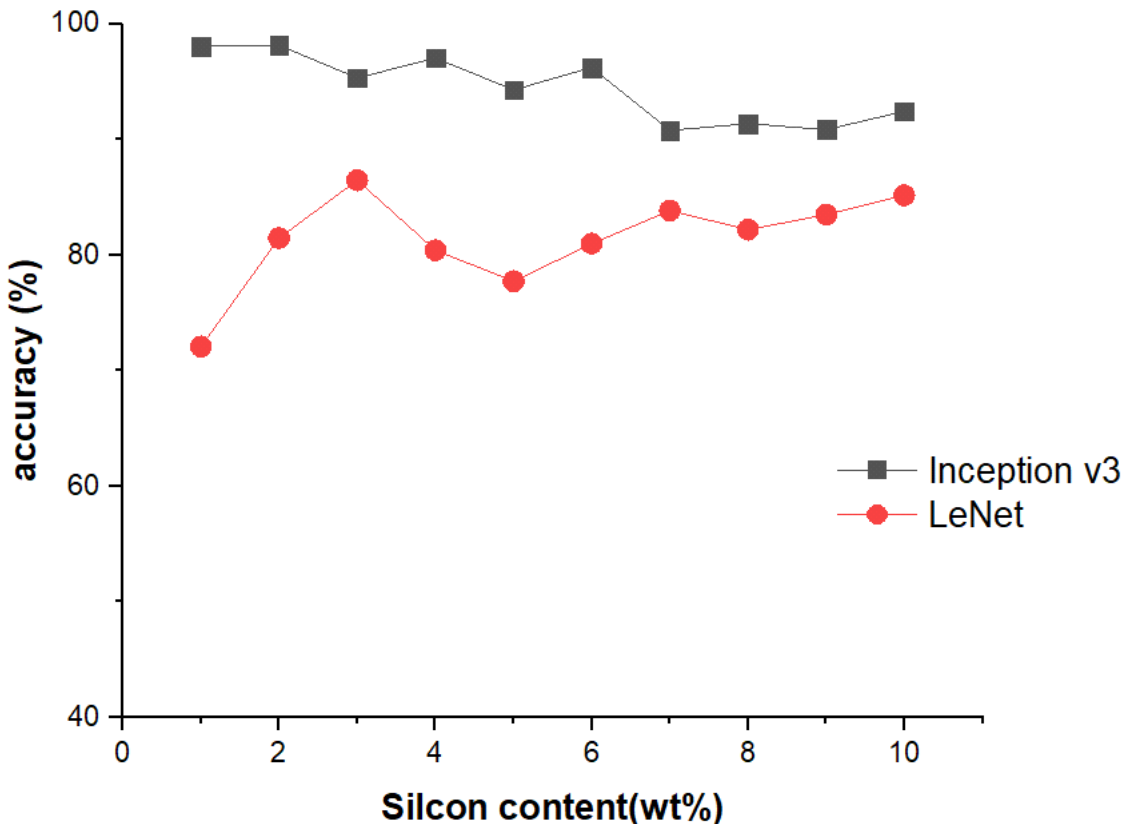


Fig. 4.9 Comparison of Inception Model, LeNet Model

제 5 장 결론

본 연구에서는 Al-Si 주조합금의 미세조직 이미지를 통한 합금 조성의 분석을 위하여 GTX 1060 그래픽 카드와 16GB Ram의 사양으로 실험을 진행하였으며, Tensorflow 프레임워크와 Python, Inception V3 모델을 이용하여 Si함량 1~10 wt%으로 1 wt% 간격으로 데이터베이스를 구축하여 학습을 진행하였다. 기계 학습을 통한 조성 분석을 통하여 다음과 같은 결론을 얻을 수 있었다.

1) 분석하고자 하는 시편을 입력하였을 때, 해당 목표 조성에 맞는 정확도를 확인할 수 있었고, 반복 학습을 통한 학습 횟수가 증가할수록 더 높은 신뢰성을 얻는 데이터를 얻을 수 있었다.

2) 이미지 데이터의 확보를 위해 기존 이미지 데이터의 변환을 통한 이미지 데이터 증가는 예측 결과에 큰 영향을 미치지 않고, 축척 표시의 배율 정보를 제거할 경우 컴퓨터가 많은 특징 정보를 입력받아 더 높은 정확도를 나타낸다.

3) 데이터 셋의 범위를 벗어나는 경우 해당 목표 조성에 맞는 예측과 분석은 가능하지 않았고, 그에 가장 유사한 범위 내의 조성으로 나타나게 되었다. 그러므로 알고자 하는 조성이 범위 내에 존재하는지의 확인이 우선시된다. 더욱 세밀한 분석을 위해서는 데이터의 확보가 중요하며 양에 따라 학습 횟수의 증가가 필요하다고 판단된다.

4) 분석하고자 하는 데이터가 이미지인지, 텍스트인지, 정확도가 우선시 되어야 하는 데이터인지, 속도가 우선시 되어야 하는지에 따라 필요한 모델이 다르게 될 것으로 판단된다.

참고 문헌

- [1] 진성옥(SungOk Jin) 외 (예술인문사회융합멀티미디어논문지, Vol.8 No.4, [2018])[KCI등재] Smart Factory 구축과 성과에 대한 연구 = A Research on Constructing Smart Factory and Its Performance-Focused on A Company-
- [2] Deng L, Yu D. Deep learning: methods and applications[J]. Foundations and Trends® in Signal Processing, 2014, 7(3-4): 197-387.
- [3] Fritz, H. (INTELLIGENT AUTONOMOUS VEHICLES, Vol.2 No.-, [1995]) Neural Speed Control for Autonomous Road Vehicles
- [4] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network[J]. arXiv preprint arXiv:1503.02531, 2015.
Texture, Dense- and Sparse-Stereo Reconstruction
- [5] IOFFE, Sergey; SZEGEDY, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.
- [6] Karen Simonyan& Andrew Zisserman,VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION,2014
- [7] LECUN,Yann, 1995,3361.10: 1995/KARPATHY, Andrej, 2014. P.1725-1732
- [8] LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. nature, 2015, 521.7553: 436.
- [9] Li, X.-r.; Zhao, G.-z.; Zhao, L.-y.; (CONTROL AND DECISION, Vol.23 No.7, [2008]) [SCOPUS] Design of decision-tree-based support vector machines multi-class classifier based on vector projection
- [10] Moonasar, V.;Venayagamoorthy, G.; (INTELLIGENT ENGINEERING SYSTEMS THROUGH ARTIFICIAL, Vol.10 No.-, [2000]) Artificial Neural Network Based Automatic Speaker Recognition Using a Hybrid Technique for Feature Extraction
- [11] Ngiam J, Khosla A, Kim M, et al. Multimodal deep learning[C]//Proceedings of the 28th international conference on machine learning (ICML-11). 2011: 689-696.
- [12] Pomerleau D A. Neural network perception for mobile robot guidance[M]. Springer Science & Business Media, 2012.
- [13] Pradhan B, Lee S. Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia[J]. Landslides, 2010, 7(1): 13- 30.
- [14] Ratanamahatana, C. A., & Keogh, E. (2004, April). Making time-series

- classification more accurate using learned constraints. In Proceedings of the 2004 SIAM International Conference on Data Mining. pp. 11-22. Society for Industrial and Applied Mathematics
- [15] Rara, H.M.; Ali, A.A.; Elhabian, S.Y.; Starr, T.L.; (International Conference on Pattern Recognition, Vol.2 No.-, [2010]) Face Recognition at-a-Distance Using
- [16] Sandberg, A.; Lansner, A.; Petersson, K. M.; Ekebe; (Network, Vol.13 No.2, 2002) [SCI,SCIE,SCOPUS] A Bayesian attractor network with incremental learning
- [17] Schmidhuber J. Deep learning in neural networks: An overview[J]. Neural networks, 2015, 61: 85-117.
- [18] Srivastava, Chandan, Lap Lambert Academic Publishing,[2011] Support vector data description : means of detecting artificially-induced postural degradation, machine learning, biomedical signal processing
- [19] Taormina R, Chau K W, Sethi R. Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon[J]. Engineering Applications of Artificial Intelligence, 2012, 25(8): 1670-1676.
- [20] (Yong Bin Gao) (Journal of information processing systems, Vol.11 No.4, 2015)[KCI등재] Viewpoint Unconstrained Face Recognition Based on Affine Local Descriptors and Probabilistic Similarity = Viewpoint Unconstrained Face Recognition Based on Affine Local Descriptors and Probabilistic Similarity
- [21] Zou, B.; Li, L.-q.; (JOURNAL- HUBEI UNIVERSITY NATURAL SCIENCE EDITION, Vol.26 No.1, [2004]) The relation of solution between support vector machine regression and nu-support vector machine classification