International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 1918-1924 www.ijaem.net ISSN: 2395-5252

Deep Learning Techniques: A Review

Meena Siwach[1], Manya Garg[2], Supriya Rai[3], Shruti Sagar[4]
[1] Assistant Professor, MaharajaSurajmalInstituteofTechnology,GGSIPU,Delhi
[2]Student, MaharajaSurajmalInstituteofTechnology,GGSIPU,Delhi
[3]Student, MaharajaSurajmalInstituteofTechnology,GGSIPU,Delhi
[4]Student, MaharajaSurajmalInstituteofTechnology,GGSIPU,Delhi

Submitted: 15-05-2022 Revised: 25-05-2022 Accepted: 28-05-2022

ABSTRACT: The tremendous advancements indeeplearningalgorithms withinvarious domainshaveprovidedmuchcontributioninartificialintelligence. It has secured a good place in patternrecognition and machine learning. The given paperillustratesthevariousdeeplearningstrategiesutilizedforvariousmodels.Afterstudyingandanalysing various techniques, comparing their prosandcons,itbecomeseasiertodecidewhichtechnique is best for a particular model. This articlealsodiscussesvariousadvantagesanddisadvantagesof eachtechnique.

KEYWORDS:RNN,CNN,gradientdescent,backpropagation,Boltzmannmachine,deepreinforcement.

I. INTRODUCTION

The high-performance computing facility of deeplearningtechniqueshavemadethembecomepopular. The main advantage of deep learning is itsabilitytoprocesslargenumberofdataeveninunstructureddata.Deeplearningmaybeimplementedusingvarioustechniqueslike Recurrentneuralnetwork,convolutionalneuralnetwork,deepreinforcement,gradientdescent,back propagation, which are described in sectionsIII.SectionIVlistsvariousadvantagesanddisadvantages.

II. MACHINE LEARNING VS. DEEPLEARNING

Machine learning is a part of artificial intelligent(AI) that has the capability to think and act likeshumans. Once structured data are fed, it can takeindefinitely new data, acting and sorting on its ownwithout the support of humans. It requires smallamount of data to make predictions. It depends onlowendmachines.

Deep learning is a subpart of machine learning. Itconsists of more algorithms than machine learning.Thesenetworksofalgorithmsareknownasartificialneuralnetworks.Ituseslargeamountsof datatomakepredictions.Itworksonhighendmachines.

III. TYPES OF DEEPLEARNING

ThedifferentDeeplearningtechniquesarediscussed below:

1. RECURRENTNEURALNETWORK

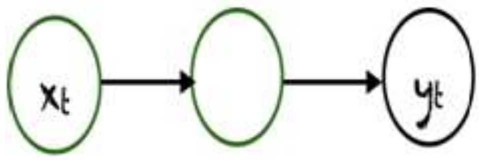
Whenwetalkaboutsequentialdata,theoldnetworks don't seem to perform great in terms oflearningandpredictionofvariousdatasets.Thereis a need of such a network that can colab with thepastdataefficientlyandthenpredictthenewtestingdatawithhighaccuracyscore.Suchanetwork is Recurrent Neural Network or in shortRNN.JurgenSchmidhuberweretheonewhostartedworkingwithRNNandmadea researchteam. The major advancement after the use of RNNisitsabilitytolinkthenodesinpreviouslayerswith the futurelayers.

THEACTIVATIONFUNCTION[1]:

ThecommonactivationfunctionsusedinRNNare:
i) Sigmoidfunction:$1/(1+e^{-x})$
ii) Tanh function:$(e^{x}-e^{-x})/(e^{x}+e^{-x})$
iii) Relu function:max(0,x)

TYPESOFRNN:

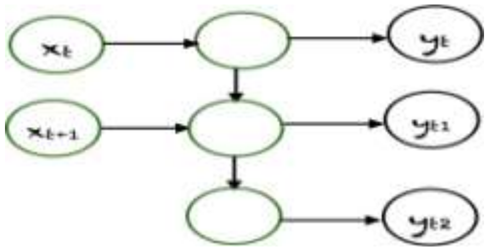
Differenttypesofrecurrentneuralnetworkswithdifferent architecturesare:
OneTo One:



One to one

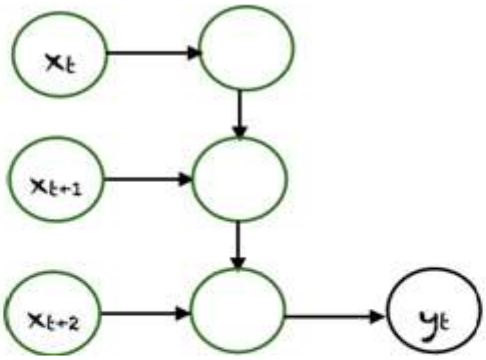
ManyToMany:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 1918-1924 www.ijaem.net ISSN: 2395-5252



Many to many

ManyTo One:



Many to one

ARCHITECTURESOFRNN:

1) FRNN(FullyRecurrentNeuralNetwork): It connects all of a neuron'soutputsto allofaneuron'sinputs.

2) ElmanandJordanNetworks:

It is a 3 layered (x, y, z) network witha set of context units. They are alsoknownasSimpleRecurrentNetworks.

3) LongShort-TermMemory: LSTM isa deep learning system that resolvestheGradientDescentproblem.Itprevents backpropagated errors fromvanishing.

4) GratedRecurrentUnit:

GRU'sareagatingmechanismin RNN.Itsmainapplicationisinspeechrecognitionandmusicmodelling.TheyhavefewerargumentsthanLSTM.

2. CONVOLUTIONALNEURALNETWORK

ACNNisatypeofartificialneuralnetwork usedtoevaluatevirtualpicturesindeeplearning.CNNs are

also known as Shift Invariant Artificial NeuralNetworks (SINNs) or Space Invariant ArtificialNeural Networks (SINNs) (SIANN). Instead ofmatrix multiplication, it employs a mathematicaltechniqueknownasconvolution.

ARCHITECTUREOFCNN:

ACNNconsistsofaconvolutionlayer,poolinglayer andfully connectedlayers.

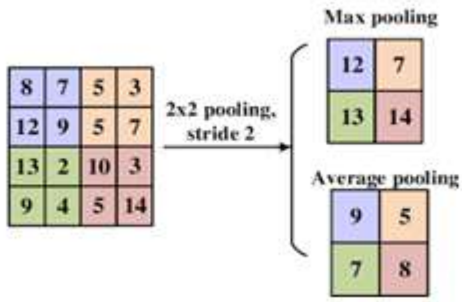
CONVOLUTIONLAYER[2]:

Similar to a neuron's reaction to a specificstimulus,convolutionallayersconvolvethe inputs and transmit the findings to thenextlayer.Convolutionreducesthenumber of free parameters in the network,making it deeper. Also, for data with grid-liketopologies,suchasphotographs,convolutionalneuralnetworksaresuperior.

POOLINGLAYER[2]:

By merging the output of a neuron clusteratonelayerintoasingleneuronatthenextlayer,poolinglayersreducesthedimensionality of data. In the CNN, therearetwoformsof pooling:

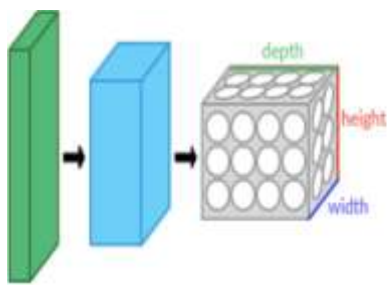
i) Max Pooling
ii) Average Pooling



Layers of CNN

FULLYCONNECTEDLAYERS:

All of the neurons in one layer are coupledto all of the neurons in another layer in afully connected layer. It works in the samewayasamulti-layerperceptronneuralnetwork (MLP). For picture classification,the flattened matrix passes through a fullyconnected layer.



CNNlayersarranged in3dimensions

APPLICATIONSOFCNN:

1. Recognition ofImages
2. ExaminationofVideos

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 1918-1924 www.ijaem.net ISSN: 2395-5252

3. NaturalLanguageProcessing
4. DetectionofAnomaly
5. DrugDevelopmentand Discovery
6. Riskassessmentfor health
7. Agerelated biomarkers
8. Checkersgame
9. Forecastingoftimeseries
10. Culturalheritageand3ddatasets

3. GRADIENTDESCENT

Gradient Descent [5] is an algorithm that solvesoptimizationproblemsindeeplearningandmachine learning models by operating iteratively tofind the best values for the parameters (weights andbiases in deep learning) of model's cost functionthatminimizeit.

WORKING

Basically,Gradientdescentfindsthepointoflocalminimaof thecostfunction[3].
Itisdrivenbytheintuitionthatthefunction[J(w)]atitsoptimalpointswillhaveahorizontalslopeandifthefunctionisconvex,itwillbeminimum.In the beginning, a random parameter value (Wo) isprovided(whichisupdatedateachiteration)totheGradientDescent,and thelearningrateisdefined.

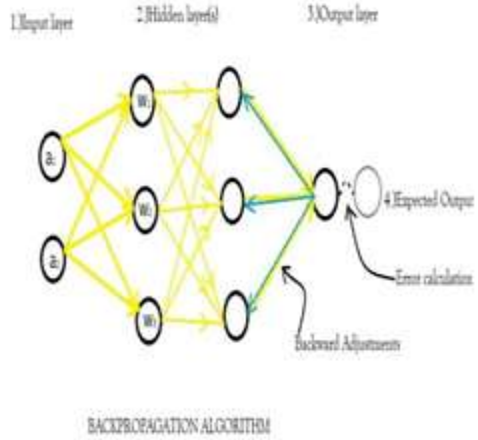
Wn= Wo - l.r * d/dw [J(w)]Where,
Wn= updatedparametervalue
Wo=previousparametervalue
l.r =learning rate
J(w)=costfunction

Then,thealgorithmproceedsbycheckingthepartialderivatives(calledgradientinGradientDescent)atthegivenparametervalues(Wo).
SimultaneouslyWnandWoareupdatedonthebasisoflearningrateandobtainedpartialderivative value. Above steps are iterated and an optimal parametervalueisobtainedfor which the values of the cost function J(w) is significantlydecreased.

CostFunctionDerivative(d/dw[J(w)])

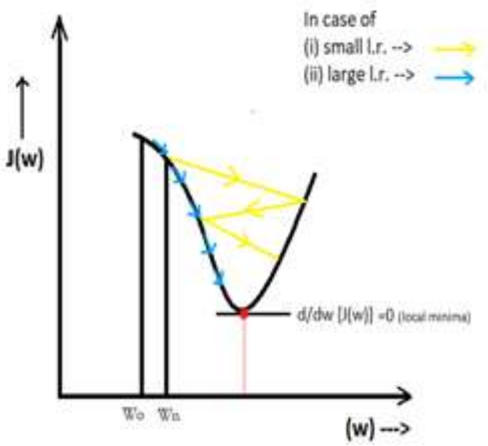
Sincethecostfunctionimplies the error rate oftraining models,minimizingitwouldgivebetterpredictingvalues.
So, thealgorithmisiterateduntiltheCostFunction derivative (d/dw [J(w)]) is minimum, atwhichthecorrespondingparametervaluesminimizethe costfunctionmost.
Derivative (slope) indicates the direction as well inwhich the coefficient is to be moved in the nextiteration toget the lowercostvalue.
That'swhytheDerivativeoftheCostFunction isused in GradientDescent.

LEARNING RATE It is a constant that defines thesizeofthestepsandthepaceatwhichthealgorithm. IftheLRistoolarge,thealgorithmmayjumpover/skip the local minima. Therefore, small valuesofLR areusedtakingintoaccountthatLRisnottoo



Backpropagation Algorithm

smallasthatwillincreasethenumberofstepsmakingthe algorithmtooslow.Using theright learning rate is important as the efficiency ofGradientDescentisdependentonit.



Efficiency of gradient descent

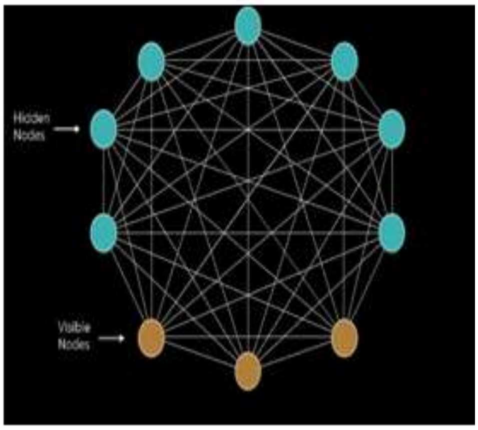
4. BACKPROPAGATION

In order to minimize the cost function, weights andbiases are updated by computing its gradient withtheuseofBack Propagation [6]Algorithm.
It is a standard method, especially used to traindeepneuralnetworksassociatedwitherrorsusceptibleprojectslikeimageandspeechprocessing.
Thealgorithmishighlyefficient,simpleandconvenientlyprogrammable.

Itisaflexiblewaywithverylittleprerequisiteknowledgeneeded.

WORKING:
Gradientofthelossfunctioniscalculatedwithrespecttoeachweightinthenetworkas partialderivativeusingchainruleofdifferentiation.
Trainingisdonewithclassifieddatasets(inwhichoutputsfortheinputsarealready known).
Followingistheexplanation:
i.)Data(a) isprovidedto theinput layer[7].
ii.)Inputisframed usingweights'w'thatinitially arerandomly selected.
iii.) Output for each neuron at everylayer, from input to output via the hidden ones iscalculated.
iv.)After comparing the obtained and the expectedoutputs,error iscomputed.
v.)Finally,thealgorithmtravelsbackfromtheoutput layerto theinputlayeradjustingthemodel'sparametersbasedonthecomputederrorvalueinweights and biases for the purpose of optimization.Algorithmrunsiterativelyandrecursivelyuntilthemodelistrainedcompletely.

5.BOLTZMANNMACHINE
BoltzmannMachine[8]isaneuralnetworkwithbidirectionallyconnectednetworksofstochasticprocessing units.
It falls under the category of unsuperviseddeep learning.
Boltzmann Machine has only two types ofnodes–hiddenandvisiblenodes.Allnodes are connected to each other and itallows them to share information amongthemselves.



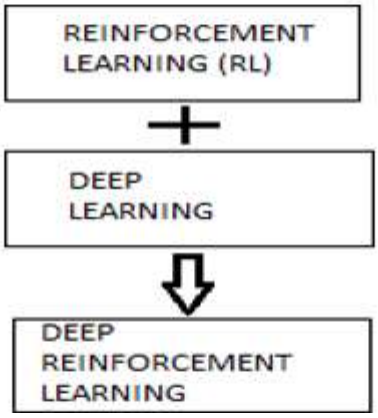
Nodes of Backpropagation

Boltzmann Machine is made up of neuralnetworks with several layers of input andtheseneuralnetworksareconnectedtoneurons.Theseneuronsgenerateinformation.

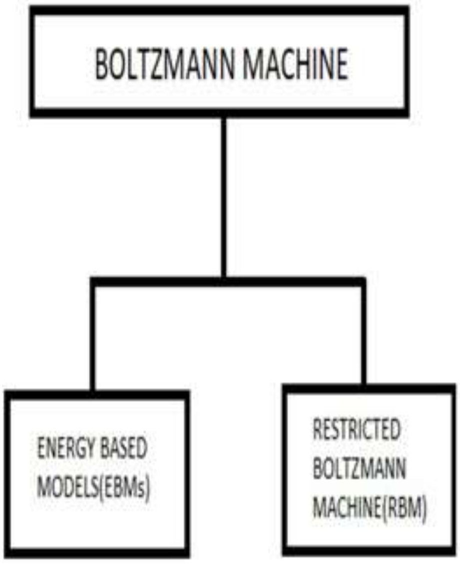
USESOFBOLTZMANNMACHINE:
1. Identify underlying structurewithin data
2. Optimizesquantitiesandweight
3. RBM [8] is used in imageprocessing

APPLICATIONS:
1.Dimensionalityreduction
2.Recommendersystem
3.Topicmodelling

BMISCATEGORIESAS:



Sequenceof Reinforcement Learning



Boltzman Machine Flow Graph

5. DEEPREINFORCEMENTLEARNING
DeepReinforcementLearning[9]isasubdivisionofmachinelearningandartificialintelligence.

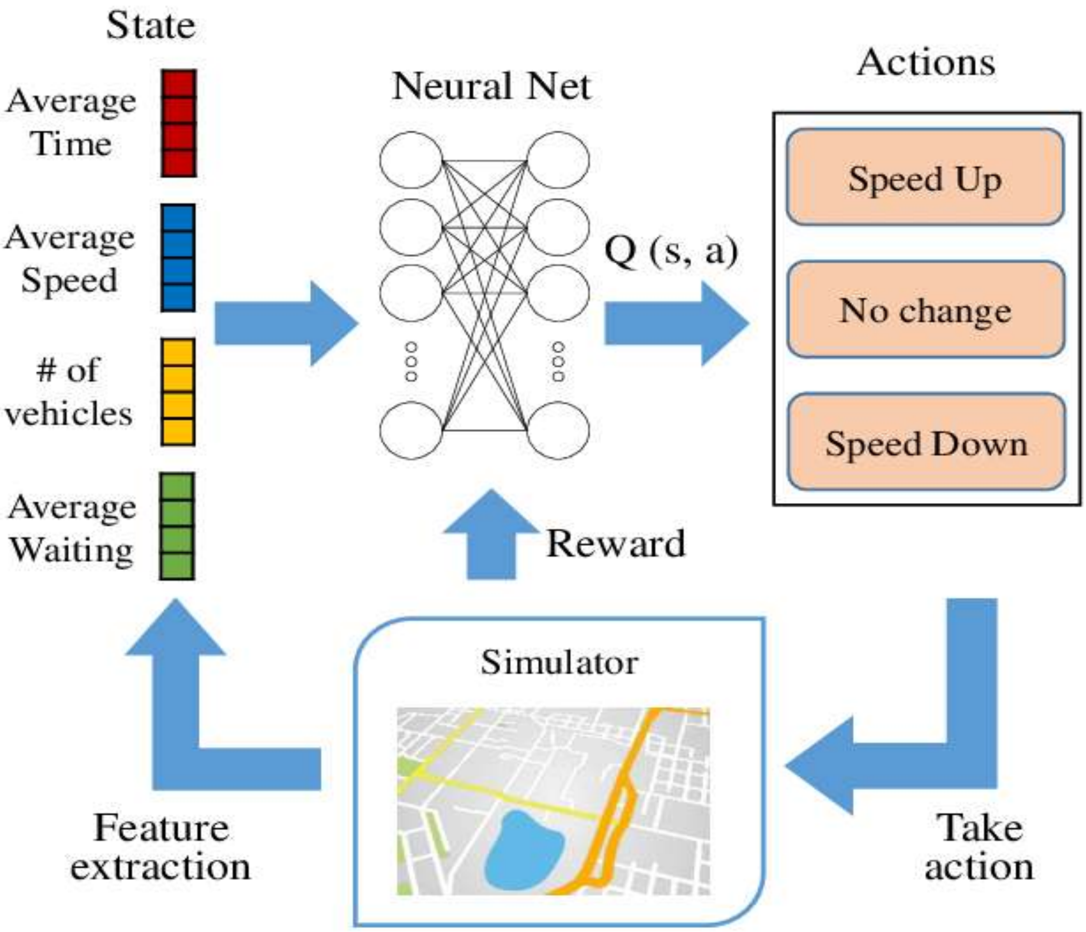
DeepRListhecombinationofreinforcementlearninga

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 1918-1924 www.ijaem.net ISSN: 2395-5252

nddeeplearning.

This field has been able to solve immensenumberofcomplexproblemsthatwereforme routofreachfora machine. Deeplearningalgorithmsarebasedonartificialneuralnetworks.Theseneuralnetworks make every effort to behave likehumanbrainthoughfarfromitsabilitythus allowing it to learn from large amountof data.

Deeplearningtechnologiesarebehindmanyeverydayproductslikedigitalassistants, voice enable TV remote as wellasself-drivingcars. Reinforcementlearning(RL)[10]isafield of machine learning that deals withconsecutive decision making. RL methodworksoninteractingwiththeenvironment.



Reinforcement Learning Implications

IV. ADVANTAGES AND DISADVANTAGES

Sr.No	Techniques	Pros	Cons
1.	RecurrentNeuralNetwork	Itcan handleordereddata.	Thenumerationcanbeveryslow.
2.	ConvolutionalNeuralNetwork	Ithashighaccuracyinimage recognitionproblems.	Needlargetrainingdata.
3.	GradientDescent	It is simpleand the kind of input that itgives is easy to understand.This makes it a popularalgorithm.	Finds local minima instead ofrequired global minima. In caseof multiple output of thealgorithmdependsontheinitialinput (parameter) provided i.e.for a different value of initialguess varied minima might begiven asoutput.
4.	Backpropagation	Useful in error proneprojects.	Itissensitivefornoisydata.
5.	BoltzmannMachine	Ability to produce newexamplesofdatavectors.	Adjustmentofweight.
6.	DeepReinforcementLearning	In manydifferentapplications, same neuralnetwork approach can beperformed.	Inordertoperformbetterlargeamountof data isrequired.

V. CONCLUSIONS

Looking back at the study that has been presented inthis paper, six efficient deep learning algorithms viz.RNN,CNN,GRADIENTDESCENT,BACK PROPAGATION have been thoroughly introduced.Working of these algorithms, common challenges thatare encountered while deploying them and suitableareasof applicationshave beenexplained.Suitablealgorithmfordeploymentcan beinferredfromthe above study according to the nature of the objectiveand data set associated with them.When required to deal with ordered data sets RNNcould be used and Boltzmann machine to produce newexamplesofthedatavector.CNNisaviableoptionfortthe projects involving high accuracy in imageprocessing and backpropagation for the ones that arepronetoerror. GradientDescentservesthepurposeofoptimizationindeep learningmodels.Deep Reinforcement learning is deployed in theprojects involving consecutive decision making byinteracting with theenvironment.

REFERENCES

[1]. HojjatSalehinejad, Sharan Sankar, Joseph Barfett,ErrolColak,andShahrokhValaee,RecentAdvancesinRecurrentNeuralNetworks,2018.

[2]. SakshiIndolia,AnilKumarGoswami,S.P.Mishra,andPoojaAsopa,ConceptualUnderstanding of Convolutional Neural Network-A Deep Learning Approach, Procedia ComputerScience,Volume132,pp.679-688,2018.

[3]. FZhouandG.Cong,OntheconvergencepropertiesofaK-stepaveragingstochasticgradientdescentalgorithm, 2017.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 5 May 2022, pp: 1918-1924 www.ijaem.net ISSN: 2395-5252

[4]. ZHuoandH.Huang,Asynchronousmini-batchgradient descent with variance reduction for non-convexoptimization[C].ProceedingsoftheAAAIConference onArtificial Intelligence, vol.31,no. 1,2017.
[5]. Xin Wang, Liting Yan, Qizhi Zhang, Research onthe Application of Gradient Descent Algorithm inMachineLearning.2021InternationalConferenceonComputerNetwork,Electronicand Automation (ICCNEA),2021.
[6]. Esser, S., Appuswamy, R., Merolla, P., Arthur, J.,andModha,Backpropagationforenergy-efficientneuromorphiccomputing.AdvancesinNeuralInformationProcessingSystems,1117–1125,2016.
[7]. MassimoBuscema,BackPropagationNeuralNetworks. Substance Use & Misuse 33(2):233-70,1998.
[8]. ZhengWangandQingbiaoWu,ShapeCompletionUsingDeepBoltzmannMachine.ComputationalIntelligenceandNeuroscience,vol2017, 2017.
[9]. PieterAbbeelandJohnSchulman.DeepReinforcementLearningthroughPolicyOptimization, Tutorial atNIPS, 2016.
[10]. KaiArulkumaran,MarcPeterDeisenroth,MilesBrundage,andAnilAnthonyBharath.ABrief Survey of Deep Reinforcement Learning,IEEESIGNALPROCESSINGMAGAZINE,SPECIAL ISSUE ON DEEP LEARNING FORIMAGEUNDERSTANDING(ARXIVEXTENDEDVERSION), 2017.
[11]. Meena Siwach, Suman Mann, Anomaly detection for web log data: A Survey,IEEE Conference,2022.
[12]. Meena Siwach, Suman Mann, Anomaly detection for web log data analysis using improved PCA Technique, Journal of information and optimization Science. 131-141, DOI: 10.1080/02522667.2022.2037283.