tionoverZ doesnotsignilicantly impacttheapproach described s0far,buttheuseof

latentvariables isSOubiquitous thatit deserves specialtreatment.

In particular,someinsightcanbegainedbyviewingtheinferenceprocessinthe

presence oflatentvariables asasimultaneous minimization overYandZ:

y" a argmlliygyzzbi41.d) (43)

Latentvariables canbeviewedasintermediate resultsonthewayto findingthebest

outputY. Atthispoint,onecouldarguethatthereisnoconceptual difference between

theZ andy variables: Z couldsimplybefoldedintoY. Thedistinction arisesduring

training:wearegiventhecorrectvalueof Y foranumberof trainingsamples, butwe

arenevergiventhecorrectvalueof Z.

Latentvariables areveryuscfulin situationswherea hiddencharacteristic ofthe

processbeingmodeled canbeinferredfromobservations, butcannotbepredicted di-

rectly.Onesuchexample isin recognition problems, Forexample, infacerecognition

thegenderofapersonortheorientation ofthefacecouldbealatentvariable,Knowing

thesevalueswouldmaketherecognition taskmucheasier.Likewiseininvariantobject

recognition theposeparameters oftheobject(location,orientation, scale)ortheillumi-

nationcouldbelatentvariables. Theyplayacrucialrolein problemswheresegmenta-

tionofthesequential datamustbeperformed simultaneously withtherecognition task.

A goodexample is speechrecognition, in whichthesegmentation of sentences into

wordsandwordsintophonemes musttakeplacesimultancously withrecognition, yet

thecorrectsegmentation intophonemes israrelyavailableduringtraining.Similarly,in

handwriting recognition, thesegmentation of wordsintocharacters shouldtakeplace

simultaneously withtherecognition. Theuseoflatentvariables infacerecognition

discussed inthissection,andSection7.3describes a latentvariablearchitecture for

handwriting recognition.

4.1 An Exampleof Latent Variable Architecture

Toillustratetheconceptof latentvanables, weconsiderthetaskof facedetection,

beginning withthesimpleproblemofdetermining whetherafaceispresentornotin

asmallimage.ImaginethatweareprovidedwithafacedetectingfunctionGface(X)

whichtakesa smallimagewindowasinputandproduces ascalaroutput.It outputs

asmallvaluewhenahumanfacefillstheinputimage,andalargevalueif nofaceis

present(orif onlya pieceofafaceoratinyfaceis present).Anenergy-based face

detector builtaroundthisfunctionisshowninFigure8(a).ThevariableY controlsthe

positionofabinaryswitch(1= "face",0= "non-face").Theoutputenergyis equal

toGface(X)whenY = 1,andtoafixedthreshold valueT whenY = 0:

E(Y,X) 1 YGaata)+(-YyT.

ThevalueofY thatminimizes thisenergyfunctionis 1(face)if Glace(X)< Tand0

(non-lace)othervisc.

Letusnowconsider themorecomplextaskofdetectingandlocatingasingleface

inalargeimage.WecanapplyourGface(X)function tomultiplewindowsinthelarge

image,computewlhichwindowproduces thelowestvalueof Gince(X),anddetecta

21